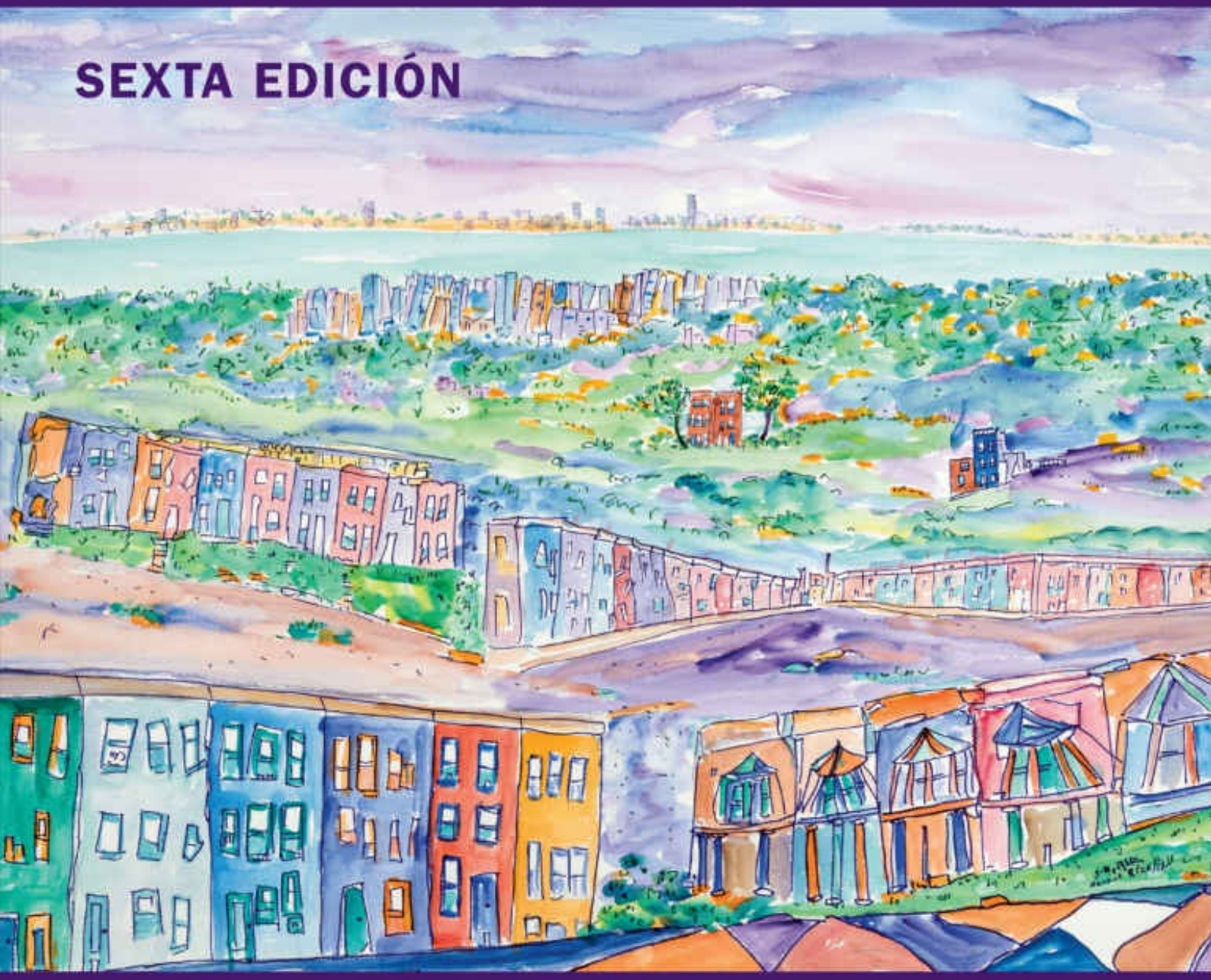
DAVID D. CELENTANO • MOYSES SZKLO

Incluye VERSIÓN DIGITAL en inglés

GORDIS
EPIDEMIOLOGÍA

SEXTA EDICIÓN



ELSEVIER

Gordis. Epidemiología

6.ª EDICIÓN

David D. Celentano, ScD, MHS

Dr. Charles Armstrong Chair and Professor
Department of Epidemiology
Johns Hopkins Bloomberg School of Public Health
Baltimore, Maryland

Moyses Szklo, MD, MPH, DrPH

University Distinguished Professor
Department of Epidemiology
Johns Hopkins Bloomberg School of Public Health
Baltimore, Maryland



Índice de capítulos

Instrucciones para el acceso en línea

Cubierta

Portada

Página de créditos

Dedicatoria

Prefacio

Agradecimientos

Sección 1: Abordaje Epidemiológico de la Enfermedad y la Intervención

Introducción

Capítulo 1: Introducción

- ¿Qué es la epidemiología?
- Objetivos de la epidemiología
- Epidemiología y prevención
- Epidemiología y práctica clínica

Planteamiento epidemiológico

De observaciones a acciones preventivas

Cuando la frecuencia de una enfermedad disminuye, ¿de quién es el mérito?

Integración de la prevención y el tratamiento

Conclusión

Capítulo 2: Dinámica de la transmisión de las enfermedades

Modos de transmisión

Enfermedad clínica y subclínica

Estado de portador

Endémico, epidémico y pandémico

Brotes de enfermedad

Inmunidad y susceptibilidad

Inmunidad grupal

Periodo de incubación

Tasa de ataque

Exploración de la ocurrencia de la enfermedad

Investigación del brote

Conclusión

Capítulo 3: La ocurrencia de la enfermedad: I. Vigilancia de la enfermedad y medidas de la morbilidad

Vigilancia

Etapas de la enfermedad en un individuo y en una población

Medidas de la morbilidad

Conclusión

Capítulo 4: La ocurrencia de la enfermedad: II. Mortalidad y otras medidas del impacto de la enfermedad

Medidas de la mortalidad

Comparación de la mortalidad en diferentes poblaciones

Otras medidas del impacto de la enfermedad

Conclusión

Capítulo 5: Evaluación de la validez y fiabilidad de las pruebas diagnósticas y de cribado

Variación biológica de las poblaciones humanas

Validez de las pruebas de cribado

Uso de pruebas múltiples

Valor predictivo de una prueba

Fiabilidad (repetibilidad) de las pruebas

Relación entre validez y fiabilidad

Conclusión

Apéndices del capítulo 5

Capítulo 6: La historia natural de la enfermedad: formas de expresar el pronóstico

Tasa de letalidad

Personas-años

Supervivencia a cinco años

Supervivencia observada

El método Kaplan-Meier

Suposiciones empleadas al usar tablas de vida y el método de Kaplan-Meier

Efectos aparentes sobre el pronóstico de las mejoras diagnósticas

Mediana de supervivencia

Supervivencia relativa

Generalización de los datos de supervivencia

Conclusión

Sección 2: Empleo de la Epidemiología Para Identificar las Causas de las Enfermedades

---

Introducción

Capítulo 7: Estudios observacionales

Casos clínicos y series de casos

Estudios ecológicos

Estudios transversales

Estudios de casos y controles

Conclusión

Capítulo 8: Estudios de cohortes

Diseño de un estudio de cohortes

Selección de las poblaciones de estudio

Tipos de estudios de cohortes

Ejemplos de estudios de cohortes

Estudios de cohortes para investigar la salud y las enfermedades infantiles

Posibles sesgos en los estudios de cohortes

¿Cuándo se requiere un estudio de cohortes?

Estudios de casos y controles basados en una cohorte definida

Conclusión

Capítulo 9: Comparación de los estudios de casos y controles y de cohortes

Capítulo 10: Valoración de las medidas preventivas y terapéuticas: ensayos aleatorizados

Selección de sujetos

Asignación de sujetos a grupos de tratamiento sin aleatorización

Asignación de sujetos utilizando aleatorización

Recogida de datos sobre los sujetos

Diseño cruzado

Diseño factorial

Falta de cumplimiento

Conclusión

Capítulo 11: Ensayos aleatorizados: algunos aspectos adicionales

Tamaño de la muestra

Reclutamiento y retención de participantes del estudio

Formas de expresar los resultados de los ensayos aleatorizados

Interpretación de los resultados de los ensayos aleatorizados

Las cuatro fases para probar nuevos fármacos en Estados Unidos

Cinco ensayos aleatorizados importantes en Estados Unidos

Ensayos aleatorizados para evaluar intervenciones ampliamente aceptadas

Registro de ensayos clínicos

Consideraciones éticas

Conclusión

Epílogo

Capítulo 12: Estimación del riesgo: ¿existe una asociación?

Riesgo absoluto

¿Cómo se puede determinar si una enfermedad concreta se asocia a una exposición específica?

Riesgo relativo

Cociente de posibilidades (posibilidades relativas)

Conclusión

Apéndice del capítulo 12

Capítulo 13: Conceptos adicionales sobre el riesgo: estimación de las posibilidades de prevención

Riesgo atribuible

Comparación de riesgo relativo y riesgo atribuible

Conclusión

Preguntas de revisión del capítulo 13

Apéndice del capítulo 13: fórmula de Levin para calcular el riesgo atribuible para la población total

Capítulo 14: De la asociación a la causalidad: derivación de inferencias a partir de estudios epidemiológicos

Estrategias para el estudio de la etiología de las enfermedades

Tipos de asociaciones

Tipos de relaciones causales

Evidencia de una relación causal

Directrices para juzgar si una asociación observada es causal

Derivación de inferencias causales: dos ejemplos

Modificaciones de las directrices para las inferencias de causalidad

Conclusión

Preguntas de repaso del capítulo 14

Capítulo 15: Más sobre inferencias causales: sesgos, confusión e interacción

Sesgo

Confusión

Interacción

Conclusión

Capítulo 16: Identificación de los factores genéticos y ambientales en la causalidad de la enfermedad

Genética tradicional

Enfermedades complejas

Tendencias temporales en la incidencia de la enfermedad

Análisis de vinculación en estudios familiares

Interacción entre los factores genéticos y los factores de riesgo ambientales

Medicina de precisión

Perspectivas de futuro

Conclusión

Glosario de los términos genéticos del capítulo 16

Sección 3: Aplicación de la Epidemiología a la Evaluación y al Desarrollo de Políticas Sanitarias

Introducción

Capítulo 17: Uso de la epidemiología para evaluar los servicios de salud

- Estudios de proceso y de resultado
- Eficacia, efectividad y eficiencia
- Medidas de resultado
- Comparación entre los estudios epidemiológicos sobre la etiología de la enfermedad y la investigación epidemiológica que evalúa la efectividad de los servicios de salud
- Evaluación mediante el uso de datos de grupo
- Evaluación utilizando datos individuales
- Conclusión

Capítulo 18: Enfoque epidemiológico en la evaluación de los programas de cribado

- Historia natural de la enfermedad
- Patrón de progresión de la enfermedad
- Problemas metodológicos
- Diseños de estudio para la evaluación del cribado: estudios no aleatorizados y aleatorizados
- Problemas en la evaluación de la sensibilidad y la especificidad de las pruebas de cribado
- Interpretación de los resultados de estudios que no muestran beneficios del cribado
- Análisis de rentabilidad del cribado
- Conclusión

Capítulo 19: Epidemiología y políticas de salud pública

- Epidemiología y prevención

Estrategias poblacionales frente a estrategias de alto riesgo para la prevención

Epidemiología y medicina clínica: tratamiento hormonal sustitutivo en mujeres posmenopáusicas

Evaluación del riesgo

Metaanálisis

Sesgo de publicación

Epidemiología en la vía judicial

Fuentes y efecto de la incertidumbre

Aspectos políticos referentes al riesgo: ¿cuáles deberían ser los objetivos?

Conclusión

Capítulo 20: Aspectos éticos y profesionales en epidemiología

Problemas éticos de la epidemiología

Obligaciones de los investigadores con los sujetos de estudio

Protección de la privacidad y la confidencialidad

Acceso a los datos

Raza y etnicidad en los estudios epidemiológicos

Conflicto de intereses

Interpretación de los resultados

Conclusión

Respuestas a las preguntas de repaso

Índice alfabético

Página de créditos



Avda. Josep Tarradellas, 20-30, 1.°, 08029, Barcelona, España

Gordis Epidemiology, 6th edition
Copyright © 2019 by Elsevier Inc. All rights reserved.
Previous editions copyrighted 2014, 2009, 2004, 2000, 1996 by Saunders, an imprint of Elsevier Inc.
ISBN: 978-0-323-55229-5

This translation of *Gordis Epidemiology*, 6th ed., by David D. Celentano and Moyses Szklo, was undertaken by Elsevier España, S.L.U. and is published by arrangement with Elsevier, Inc.

Esta traducción de *Gordis Epidemiology*, 6.ª ed., de David D. Celentano y Moyses Szklo, ha sido llevada a cabo por Elsevier España, S.L.U. y se publica con el permiso de Elsevier, Inc.

Gordis. Epidemiología, 6.ª ed., de David D. Celentano y Moyses Szklo
© 2020 Elsevier España, S.L.U., 2014, 2005
ISBN: 978-84-9113-536-4
eISBN: 978-84-9113-633-0

Todos los derechos reservados.

Reserva de derechos de libros
Cualquier forma de reproducción, distribución, comunicación pública o transformación de esta obra solo puede ser realizada con la autorización de sus titulares, salvo excepción prevista por la ley. Diríjase a CEDRO (Centro Español de Derechos Reprográficos) si necesita fotocopiar o escanear algún fragmento de esta obra (www.conlicencia.com; 91 702 19 70/93 272 04 45).

Advertencia

Esta traducción ha sido llevada a cabo por Elsevier España, S.L.U. bajo su única responsabilidad. Facultativos e investigadores deben siempre contrastar con su propia experiencia y conocimientos el uso de cualquier información, método, compuesto o experimento descrito aquí. Los rápidos avances en medicina requieren que los diagnósticos y las dosis de fármacos recomendadas sean siempre verificados personalmente por el facultativo. Con todo el alcance de la ley, ni Elsevier, ni los autores, los editores o los colaboradores asumen responsabilidad alguna por la traducción ni por los daños que pudieran ocasionarse a personas o propiedades por el uso de productos defectuosos o negligencia, o como consecuencia de la aplicación de métodos, productos, instrucciones o ideas contenidos en esta obra.

Revisión científica:
Dr. Vicente Monge Jodra
Jefe de Servicio de Medicina Preventiva
Hospital Ramón y Cajal
Profesor Agregado de Cátedra en el Área de Conocimiento de Medicina Preventiva y Salud Pública
Universidad San Pablo-CEU

Servicios editoriales: DRK Edición

Depósito legal: B 19263-2019
Impreso en Polonia

Dedicatoria

In memoriam

Leon Gordis, MD, MPH, DrPH

1934–2015



Prefacio

La epidemiología es una de las disciplinas fundamentales en las que se sustenta la salud pública. La investigación clínica se basa en gran medida en métodos epidemiológicos y en la investigación actual sobre la atención médica, en particular en los estudios de eficacia comparada y en enfoques estadísticos de *big data* (como en el uso de la historia clínica electrónica para estudios sobre salud). Tal y como el Dr. Leon Gordis escribió en su prólogo a la quinta edición: "La epidemiología es la ciencia básica de la prevención de enfermedades y desempeña un papel destacado en el desarrollo y la evaluación de políticas públicas relacionadas con la salud y las cuestiones sociales y jurídicas". En la actualidad, la epidemiología tiene numerosas aplicaciones. La mayor parte de la investigación epidemiológica se centra en establecer asociaciones etiológicas entre los supuestos riesgos y los resultados de salud. Sin embargo, la epidemiología también se usa ampliamente en la evaluación de programas de prevención primaria y secundaria, las comparaciones de intervenciones y la evaluación de políticas a nivel poblacional. Los hallazgos epidemiológicos suelen acabar publicados en los medios de comunicación, lo que ofrece a la opinión pública y a los responsables de elaborar políticas datos para guiar las decisiones personales sobre su conducta. Cada vez es más frecuente que el estudio centrado en la epidemiología pueda causar cierto malestar en los investigadores y médicos, pues la interpretación de los principios epidemiológicos básicos puede estar sujeta a un error considerable. Nuestra tarea es hacer que el pensamiento en el que se basa la epidemiología sea transparente.

Esta obra tiene el propósito de convertirse en una introducción básica a las definiciones, la lógica y el uso del método epidemiológico para dilucidar los factores que influyen en la salud y la enfermedad. Hemos intentado ilustrar los principios con ejemplos de las aplicaciones de la epidemiología en el mundo real. Incluyen tanto "ejemplos clásicos" de los primeros días del desarrollo de la disciplina de la epidemiología como ejemplos actuales. Allí donde ha sido apropiado, nos basamos en ejemplos de la práctica clínica.

Tras el fallecimiento del Dr. Gordis en 2015, la sexta edición de este libro ha sido revisada por dos nuevos autores. Ambos trabajaron con y a las órdenes del profesor Gordis y han participado activamente en la docencia de la epidemiología en la Johns Hopkins durante más de cuatro décadas. Por lo general, se ha conservado la estructura y la organización de la edición anterior. En la quinta edición se insertaron objetivos de aprendizaje en la mayoría de los capítulos, que se han revisado, y los ejemplos de todos ellos se

han actualizado. Se han añadido nuevas preguntas de repaso en la mayoría de los capítulos. Se ha realizado un cambio significativo de la presentación y del orden de los métodos en epidemiología, que antes aparecían al final de la sección 1 y más extensamente en la sección 2. En lugar de comenzar con el ensayo aleatorizado (o el diseño "experimental") y después comparar el diseño del estudio observacional con el patrón oro, se ha organizado la presentación de los métodos epidemiológicos a lo largo de un continuum de estudio desde la observación clínica a las series de casos, el uso de estudios ecológicos y, después, hasta las investigaciones transversales, como enfoque fundamental del desarrollo de hipótesis epidemiológicas. A continuación, se sigue con los diseños de casos y controles y los de cohortes, para llegar al ensayo aleatorizado. En nuestra opinión, esto sigue de manera más orgánica el modo en el que las observaciones e hipótesis epidemiológicas se desarrollan en la práctica diaria de *hacer* epidemiología.

Esta sexta edición, al igual que la previa, consta de tres secciones. La sección 1 aborda cómo se utiliza la epidemiología para comprender la salud y el desarrollo de enfermedades en las poblaciones, así como los fundamentos de las intervenciones para influir en la historia natural de la enfermedad. Los primeros capítulos ofrecen el marco conceptual en el que se basa la disciplina de la epidemiología y presentan muchos de los principios básicos de la disciplina. En el capítulo 1 se ofrece una panorámica de la epidemiología, utilizando numerosos ejemplos históricos para ilustrar cómo se desarrolló la especialidad. El capítulo 2 describe el modo de transmisión de las enfermedades en la población, tanto directa (en el caso de los patógenos infecciosos) como indirectamente (p. ej., a través de un vector como un mosquito o del aire contaminado). Se presentan los términos básicos utilizados en las epidemias y se ilustran para mostrar al estudiante cómo usar estos principios y términos. El capítulo 3 trata de la vigilancia de la enfermedad y de cómo se puede medir la morbilidad en las poblaciones, mientras que el capítulo 4 se centra en los aspectos de la mortalidad y los parámetros del impacto de la enfermedad en las poblaciones. El capítulo 5 se centra en las formas de detectar la enfermedad en las poblaciones, comparando diferentes métodos para distinguir a las personas que tienen una enfermedad de las que están sanas y explicando cómo las pruebas de cribado pueden ajustarse para diagnosticar mejor a las personas que presentan o no la enfermedad en cuestión. Los aspectos de la fiabilidad y la validez de las pruebas de cribado tienen interés clínico tanto para los médicos como para quienes planifican los servicios sanitarios. Por último, el capítulo 6 presenta cómo la historia natural de la enfermedad puede utilizarse para expresar mejor el pronóstico de la enfermedad, utilizando ejemplos de letalidad y de supervivencia.

La sección 2 detalla los métodos utilizados por los epidemiólogos principalmente para atribuir asociaciones entre una exposición hipotética (riesgo) y un resultado de salud. El capítulo 7 describe las observaciones iniciales realizadas en la práctica clínica (el informe de caso) que llevan a la

identificación de una acumulación de casos que parecen tener puntos en común (la serie de casos). Esto se sigue de una introducción al diseño ecológico y su análisis, con las precauciones en cuanto a su interpretación. Por último, se presentan los estudios transversales (puntuales) como bases para el desarrollo de hipótesis. A continuación, el capítulo 8 proporciona una introducción a los estudios observacionales tal y como suelen utilizarse en epidemiología, presentando los estudios de casos y controles y los de cohortes, que se comparan después en el capítulo 9. En este punto, se abordan las exposiciones tal y como tienen lugar en las poblaciones, donde somos observadores de dichas exposiciones y de sus supuestos impactos sobre los resultados de salud. En el capítulo 10 se pasa a un enfoque "experimental" (ensayo aleatorizado) en el que el investigador "asigna" una exposición o intervención de salud (por lo general, al azar) para estudiar a los participantes con el fin de ver cómo esto influye en el resultado de salud. En este caso, la exposición está bajo el control del investigador, no del participante en el estudio, una diferencia crucial en el ensayo aleatorizado en comparación con el diseño de estudio de cohortes o de otro tipo observacional. En el capítulo 11 se describen varios aspectos implicados en la realización de ensayos aleatorizados, como el tamaño muestral, la potencia y la generalizabilidad; la determinación de la eficacia (frente a la efectividad); los aspectos éticos, y las fases de la Federal Drug Administration estadounidense para la evaluación de nuevos fármacos. En el capítulo 12, se presentan aspectos de la estimación del riesgo, incluidos el riesgo absoluto y relativo, y su interpretación, el cálculo e interpretación del *odds ratio* en un estudio de casos y controles y en un estudio de cohortes, así como en un estudio de casos y controles de participantes emparejados. En el capítulo 13, el concepto de riesgo se amplía para incluir el cálculo y la interpretación del riesgo atribuible, los riesgos atribuibles poblacionales y su uso en la evaluación del éxito de los programas preventivos. La inferencia causal se introduce en el capítulo 14, que se centra en el modo de derivar inferencias en las investigaciones epidemiológicas. El capítulo 15 presenta los aspectos del sesgo, los factores de confusión y la interacción en los estudios epidemiológicos y analiza cómo influyen en la inferencia causal. Por último, el capítulo 14 trata el papel de las contribuciones genéticas y ambientales en la etiología de la enfermedad y presenta nuevos métodos de investigación genética que suelen utilizarse en los estudios epidemiológicos actuales.

La sección 3 aborda las aplicaciones cotidianas de la epidemiología en la salud pública. Los últimos cuatro capítulos tratan algunos de los aspectos críticos a los que se enfrenta la especialidad en la actualidad. El capítulo 17 ilustra el modo en el que los principios y diseños epidemiológicos descritos en las Secciones I y II se utilizan en la evaluación de los servicios de salud. El capítulo 18 aborda el uso de la epidemiología para evaluar los programas de cribado, mientras que el capítulo 19 detalla cómo se puede utilizar la epidemiología para abordar los principales ámbitos de la política de salud pública. El último capítulo resume los aspectos éticos con los que nos

enfrentamos en la práctica de la epidemiología y en él se revisan algunas de las cuestiones profesionales importantes a las que se enfrenta la especialidad actualmente.

Hemos continuado aplicando el uso que el profesor Gordis hacía de las ilustraciones y ejemplos para demostrar cómo se ponen en práctica las cuestiones y principios epidemiológicos. Hemos actualizado los ejemplos ampliamente y hemos añadido otros nuevos a lo largo de todo el texto. Muchos de los capítulos previos se han editado y actualizado exhaustivamente y algunos de ellos se han ampliado en gran medida. La presentación de la epidemiología genética se ha editado y actualizado en profundidad, lo que refleja el increíble crecimiento en la investigación genómica que se ha producido en los últimos 5 años. En el capítulo 16, se ha añadido un glosario de términos de epidemiología genética para proporcionar al estudiante cierta orientación en este campo un tanto complejo. Por último, se han añadido nuevas preguntas de repaso al final de la mayoría de los capítulos.

Nuestro objetivo para este libro es permitir que el lector aprecie cómo la epidemiología puede utilizarse para responder a los problemas de salud poblacional a los que se enfrenta la sociedad hoy en día. No esperamos que el lector sea capaz de realizar una investigación epidemiológica, pero tenemos la esperanza de que logre apreciar lo que es la epidemiología, lo que son los diseños de investigación básica y de evaluación, así como el modo de interpretar los hallazgos básicos en un estudio epidemiológico. Esperamos ser capaces de transmitir al lector de este libro el entusiasmo que nosotros sentimos por los usos de la epidemiología.

La ilustración de la portada seleccionada para esta edición de *Gordis. Epidemiología* tiene un significado especial. Esta acuarela original de Haroutune Armenian, MD, DrPH, pintada en agosto de 2017, se titula "Remembering Baltimore". El Dr. Armenian fue profesor de Epidemiología, catedrático adjunto al profesor Gordis durante muchos años y catedrático provisional de 1993 a 1994, hasta que Jon Samet se convirtió en catedrático. "Remembering Baltimore" capta la esencia del paisaje urbano de Baltimore, que fue el hogar del Dr. Gordis durante 60 años. Las características casas adosadas del puerto son la quintaesencia de Baltimore, al igual que las "Damas pintadas" se identifican con San Francisco. Gran parte de la investigación del Dr. Gordis se centró en las enfermedades pediátricas en Baltimore, como se ilustra en muchos de los ejemplos de este libro. Estamos particularmente orgullosos de incluir este homenaje del Dr. Armenian al Dr. Gordis en nuestra primera revisión de esta obra de renombre mundial. En esta sexta edición hemos tenido presente a nuestro amigo y mentor.

David D. Celentano

Moyses Szklo

Agosto de 2018

Agradecimientos

Este libro refleja las colaboraciones de varias generaciones de profesores de epidemiología de la Johns Hopkins, primero en la School of Hygiene and Public Health, y más recientemente en la Bloomberg School of Public Health. El curso fue desarrollado por los profesores del Departamento de Epidemiología y se impartió inicialmente como Principios de Epidemiología por el Dr. Abraham Lilienfeld, catedrático del departamento de 1970 a 1975. El Dr. Leon Gordis se convirtió en profesor del curso después de una enfermedad aguda del Dr. Lilienfeld, a mediados de su enseñanza de la asignatura en 1975. A continuación, el Dr. Gordis fue el profesor principal durante los siguientes 30 años. Además, el Dr. Gordis enseñó epidemiología a muchas promociones de la facultad de medicina durante un periodo similar. Este libro se desarrolló a partir de esas experiencias y el Dr. Gordis fue el autor de las primeras cinco ediciones de esta famosa obra.

Los autores actuales se formaron en salud pública en la Johns Hopkins y tuvieron una participación activa como miembros del equipo docente de epidemiología durante muchos años cuando eran profesores jóvenes. El Dr. Szklo impartió el segundo curso en la secuencia de la epidemiología, Epidemiología Intermedia. Tras la jubilación del Dr. Gordis, el Dr. Celentano se convirtió en el director de Principios de Epidemiología, cuyo contenido se ha revisado recientemente y se ha renombrado como Inferencia Epidemiológica en Salud Pública 1. Su contenido se refleja en esta sexta edición de *Gordis. Epidemiología.*

Muchos colegas han hecho valiosas contribuciones a esta revisión de *Gordis. Epidemiología*. El principal de ellos fue el difunto Dr. George W. Comstock, mentor y asesor nuestro, así como un eminente científico. También queremos agradecer la ayuda de muchos colegas previos y actuales, como Haroutune Armenian, catedrático adjunto del Dr. Gordis y catedrático tras la jubilación de este. El Dr. Armenian ha contribuido con su pintura original «Remembering Baltimore» a la cubierta de este libro. También agradecemos a nuestro antiguo catedrático, Jonathan Samet, así como a Michel Ibrahim, que se unió a nosotros como profesor tras abandonar en 2002 su puesto de decano en la University of North Carolina–Chapel Hill. Otras personas que han tenido papeles destacados en el programa docente del departamento han sido Javier Nieto, Rosa Crum, Paul Whelton, Stephen Gange, Shruti Mehta y Alvaro Munoz. Entre los profesores previos del curso introductorio, queremos agradecer a Bill Moss, Elizabeth Platz y Jennifer Deal su dedicación a la hora de educar a numerosos estudiantes de salud pública en el «arte» de

la epidemiología. En particular, el Dr. Deal ha realizado contribuciones sobresalientes a nuestro curso introductorio y muchos de los ejemplos presentados en esta edición proceden de sus sugerencias, por las que estamos especialmente agradecidos. También apreciamos el apoyo de muchos decanos de la facultad, como D. A. Henderson, Al Sommer, Mike Klag y, más recientemente, Ellen MacKenzie. El curso en el que se basa este libro no existiría sin la dedicación prolongada y los conocimientos de nuestro colega Allyn Arnold, que ha actuado como puente entre los años de Gordis y la actualidad.

La preparación de la sexta edición de este libro ha sido una tarea significativa para nosotros. Nuestro objetivo era preservar la voz (y el humor) del Dr. Gordis, así como mantener el estilo del texto tanto como fuese posible. También hemos intentado actualizar los ejemplos e intercalar nuevas ilustraciones de los principios epidemiológicos que se presentan en el libro, junto con los ejemplos clásicos que se incluyeron en las ediciones previas.

Youssef Farag, MBBCh, MPH, PhD, ha sido una ayuda inestimable en la preparación de la sexta edición. Es un médico-epidemiólogo joven, lleno de talento y con espíritu trabajador a quien reclutamos para ayudarnos en este esfuerzo extraordinario. Mientras completaba su doctorado en Epidemiología en la Johns Hopkins Bloomberg School of Public Health, Youssef asumió la minuciosidad de preparar este libro, desde la actualización de las cifras de los CDC sobre morbilidad y mortalidad, hasta trabajar estrechamente con el National Cancer Institute para realizar nuevos análisis de datos con el fin de ilustrar los aspectos epidemiológicos clave, además de encontrar referencias sugeridas vagamente por nosotros. Se encargó de la reorganización significativa de los capítulos de esta sexta edición, incluida la reescritura de nuevas secciones desde cero en varios capítulos. También tomó la iniciativa de actualizar ejemplos obsoletos a partir de sus conocimientos de cuestiones médicas y de salud pública actuales, así como de sus sólidas nociones de la literatura relevante. Sus contribuciones creativas han facilitado la simplificación y clarificación de conceptos habitualmente difíciles en epidemiología. Durante un periodo de más de un año, a partir de nuestras discusiones profundas durante reuniones semanales, los numerosos correos electrónicos entre ellas y las múltiples revisiones de cada capítulo, este proyecto nunca se habría desarrollado con tanta facilidad sin su compromiso, su calma y su determinación, por lo que le estamos muy agradecidos. Creemos firmemente que llegará a convertirse en un líder en el campo de la epidemiología.

El capítulo sobre el papel de la genética en la epidemiología contemporánea está muy influido por nuestras colegas expertas en epidemiología genética, Priya Duggal y Terri Beaty. Los cambios en este campo son tan rápidos y su complejidad tecnológica para los legos es tan grande, que nos han ayudado a realizar una revisión significativa en esta sexta edición. Nunca les agradeceremos lo suficiente sus contribuciones a este capítulo.

Charlotte Gerczak fue una ayuda inestimable en la corrección de textos. Charlotte trabajó durante muchos años con Jonathan Samet y tiene una gran experiencia a la hora de colaborar con epidemiólogos en ejercicio. Su gran visión para la gramática, la estructura de las frases y el significado han convertido esta obra en un libro mejor de lo que habría sido sin su cuidadosa revisión.

La preparación de la sexta edición de *Gordis*. Epidemiología nos ha traído muchos recuerdos de Leon y de su legado en la Johns Hopkins. El departamento ha cambiado mucho desde que él dejó la cátedra en 1993. En la actualidad, somos una facultad mucho más amplia, que abarca muchas más áreas de la epidemiología y que usa herramientas inimaginables incluso hace una década. Al mismo tiempo, la disciplina sigue basándose en las ideas enunciadas inicialmente por Wade Hampton Frost en los albores de nuestra facultad en 1919. Este libro es un testamento a los líderes del pensamiento y gigantes de la epidemiología que han estudiado y enseñado en la Johns Hopkins en los últimos 100 años y esperamos que nos guíe en nuestro segundo siglo de práctica, docencia, investigación y servicio.

David D. Celentano

Moyses Szklo

SECCIÓN 1

Abordaje Epidemiológico de la Enfermedad y la Intervención

Introducción
Capítulo 1: Introducción
Capítulo 2: Dinámica de la transmisión de las enfermedades
Capítulo 3: La ocurrencia de la enfermedad: I. Vigilancia de la enfermedad y medidas de la morbilidad
Capítulo 4: La ocurrencia de la enfermedad: II. Mortalidad y otras medidas del impacto de la enfermedad
Capítulo 5: Evaluación de la validez y fiabilidad de las pruebas diagnósticas y de cribado
Capítulo 6: La historia natural de la enfermedad: formas de expresar el pronóstico

Introducción

Esta sección comienza con una visión general de los objetivos de la epidemiología, algunos de los abordajes utilizados en la epidemiología y ejemplos de las aplicaciones de la epidemiología en los problemas de salud del ser humano (cap. 1). A continuación se trata la transmisión de las enfermedades (cap. 2). Las enfermedades no surgen de la nada, se deben a la interacción entre los seres humanos y su entorno. La comprensión de los conceptos y los mecanismos que subyacen en la transmisión y la adquisición de las enfermedades resultan fundamentales para el estudio de la epidemiología de las enfermedades del ser humano y para la prevención y el control de numerosas enfermedades infecciosas.

Para analizar los conceptos epidemiológicos presentados en este libro es necesario desarrollar un lenguaje común, especialmente para describir y comparar la morbilidad y la mortalidad. Para ello, en el capítulo 3 se analiza la morbilidad y el importante papel epidemiológico en la vigilancia de las enfermedades. Más adelante en el capítulo se expone cómo se utilizan las medidas de la morbilidad tanto en la medicina clínica como en la salud pública. El capítulo 4 presenta la metodología y los abordajes para usar los datos sobre mortalidad en las investigaciones de salud pública y en la práctica clínica. En el capítulo 4 también se abordan otros aspectos relacionados con el impacto de la enfermedad, como la calidad de vida y las implicaciones futuras de la carga de la enfermedad.

Tras adquirir la base teórica para describir la morbilidad y la mortalidad en términos cuantitativos, se trata el tema de la valoración de la calidad de las pruebas diagnósticas y de cribado que se utilizan para determinar qué personas de la población sufren una enfermedad determinada (cap. 5). Después de identificar a las personas que sufren la enfermedad, necesitamos formas de describir la historia natural de la enfermedad en términos cuantitativos, lo que resulta fundamental para valorar la gravedad de una enfermedad y para evaluar los posibles efectos de nuevas intervenciones terapéuticas y preventivas en la supervivencia de las personas (cap. 6).

Por último, esta primera sección ofrece al estudiante una introducción a la nomenclatura de la epidemiología, el seguimiento y las ramificaciones de las enfermedades, con el objetivo de que sea capaz de definir los grupos de población en función de sus características de salud, así como de valorar las pruebas de detección y conocer la evolución natural de cada enfermedad.

CAPÍTULO 1

Introducción

Palabras clave

prevención: primaria, secundaria y terciaria
asociación
causalidad
observación
salud pública

Odio las definiciones.
—Benjamin Disraeli (1804-1881; primer ministro británico, 1868 y 1874-1880)

¿Qué es la epidemiología?

La epidemiología es el estudio de cómo se distribuyen las enfermedades en las poblaciones y los factores que determinan o influyen en esta distribución. ¿Por qué una enfermedad aparece en unas personas y no en otras? La premisa que subyace a la epidemiología es que las enfermedades, los trastornos, la mala salud y la salud excelente no se distribuyen de forma aleatoria en una población. En cambio, cada uno de nosotros tiene ciertas características que nos predisponen a diferentes enfermedades o nos protegen frente a ellas. Estas características pueden tener un origen sobre todo genético o deberse a la exposición a ciertos riesgos ambientales o comportamientos (apropiados e inadecuados) que desarrollamos. Quizás lo más frecuente es que en el desarrollo de la enfermedad nos enfrentemos con una interacción de factores genéticos y ambientales, así como de factores derivados de nuestro comportamiento y de factores sociales.

Una definición más extensa de la epidemiología que la ofrecida anteriormente ha sido ampliamente aceptada. Se define la epidemiología como «el estudio de la distribución y los determinantes de los estados o acontecimientos relacionados con la salud en poblaciones específicas y la aplicación de este estudio al control de los problemas sanitarios»[1]. Lo que resalta de esta definición es que comprende una descripción del contenido de esta disciplina y las razones por las que se llevan a cabo las investigaciones epidemiológicas.

Objetivos de la epidemiología

¿Cuáles son los objetivos específicos de la epidemiología? En primer lugar, identificar la *etiología* o la *causa* de una enfermedad y los factores de riesgo relevantes, es decir, los factores que aumentan el riesgo de que una persona padezca una enfermedad. Queremos conocer cómo se transmite la enfermedad de una persona a otra o de un reservorio no humano a una población humana, además de las razones por las que los comportamientos de riesgo causan enfermedades. Nuestro objetivo último es intervenir para reducir la morbilidad y la mortalidad de la enfermedad. Queremos crear una base racional para establecer programas de prevención. Si podemos identificar los factores causales o etiológicos de la enfermedad y reducir o eliminar la exposición a estos factores, podremos desarrollar la base de los programas de prevención. Además, podremos desarrollar vacunas y tratamientos apropiados, que pueden evitar la transmisión de la enfermedad de unas personas a otras.

El segundo objetivo de la epidemiología consiste en determinar la *extensión de la enfermedad* en la comunidad. ¿Cuál es el impacto de la enfermedad en la comunidad? Esta pregunta resulta fundamental para planear los servicios e instalaciones sanitarias y determinar el número de futuros profesionales sanitarios que deben ser formados.

El tercer objetivo es el de estudiar la *historia natural y el pronóstico de la enfermedad.* Claramente, ciertas enfermedades son más graves que otras; algunas pueden ser rápidamente mortales, mientras que otras se acompañan de periodos de supervivencia más prolongados. Muchas enfermedades no son mortales, aunque pueden influir negativamente en la calidad de vida o pueden causar cuadros de discapacidad. Queremos definir la historia natural basal de una enfermedad en términos cuantitativos, de modo que a medida que creemos nuevos métodos de intervención, ya sea a través de tratamientos o por nuevas formas de prevenir complicaciones, podamos comparar los resultados del uso de dichos métodos nuevos con los datos basales con el fin de determinar si nuestros nuevos abordajes han sido verdaderamente efectivos.

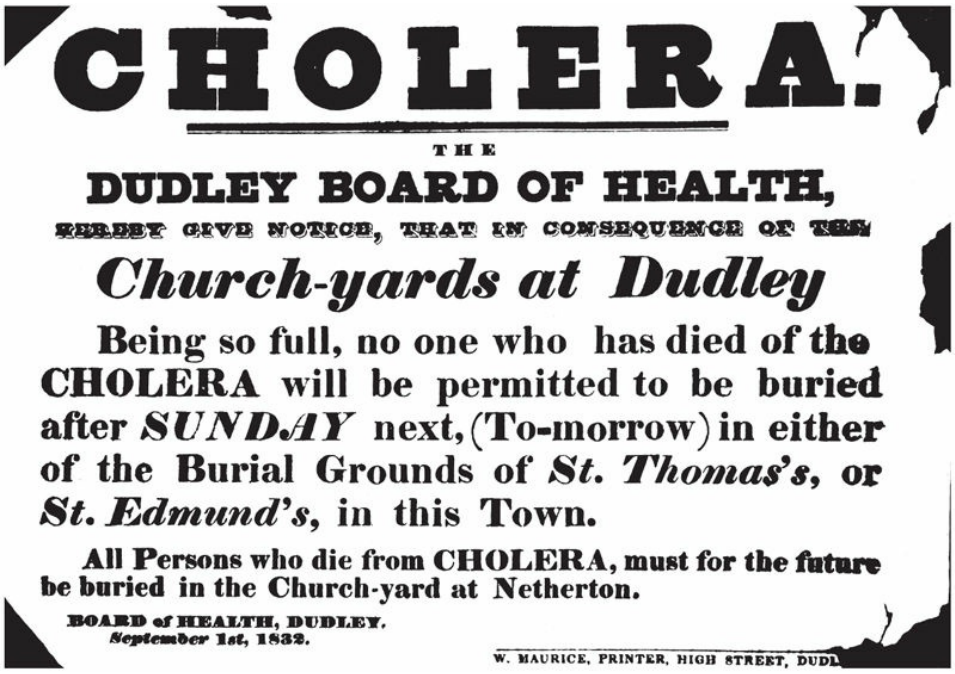
En cuarto lugar, utilizamos la epidemiología para evaluar *los modos de prestación de la asistencia sanitaria y las medidas terapéuticas y preventivas, tanto las existentes como las nuevas.* Por ejemplo, ¿el cribado del cáncer de próstata en los varones utilizando la prueba del antígeno prostático específico (PSA) aumenta la supervivencia en los pacientes con cáncer de próstata? ¿El crecimiento de la asistencia gestionada y otros nuevos sistemas de asistencia sanitaria y de seguros sanitarios ha tenido algún impacto sobre los resultados de salud de sus pacientes y sobre su calidad de vida? En caso afirmativo, ¿cuál ha sido la naturaleza de este impacto y cómo puede medirse?

Por último, la epidemiología proporciona la base para el *desarrollo de normativas públicas* relacionadas con problemas ambientales, aspectos

genéticos y otras consideraciones sociales y de nuestro comportamiento relacionadas con la prevención de las enfermedades y la promoción de la salud. Por ejemplo, ¿la radiación electromagnética emitida por los teléfonos móviles, las mantas eléctricas, los calentadores y otros electrodomésticos es un peligro para la salud humana?, ¿las concentraciones atmosféricas elevadas de ozono o de partículas en suspensión causan efectos adversos agudos o crónicos sobre la salud del ser humano?, ¿el ratón doméstico supone un riesgo importante para el ser humano?, ¿qué profesiones se asocian con un mayor riesgo de enfermedad en los trabajadores y qué normativas son necesarias para reducir dicho riesgo?

Cambio de los patrones de los problemas sanitarios en la comunidad

Un papel importante de la epidemiología es proporcionar información sobre los cambios que tienen lugar con el paso del tiempo en los problemas de salud que presenta una comunidad. En la figura 1.1 se muestra un cartel de un cementerio de Dudley, Inglaterra, de 1839. En esa época, el cólera era la principal causa de muerte en Inglaterra; el cementerio estaba tan lleno que no se permitían entierros de personas muertas por el cólera. El cartel da una idea de la importancia del cólera en la conciencia colectiva y en el espectro de problemas de salud pública a principios del siglo XIX. Claramente, el cólera ya no es un problema importante en la actualidad en Estados Unidos, pero en muchos países con escasos recursos económicos y azotados por las guerras sigue siendo una amenaza importante; se producen periódicamente brotes de cólera caracterizados por tasas de mortalidad elevadas, debidas a menudo a una asistencia médica inadecuada o inaccesible.


CHOLERA.

THE

DUDLEY BOARD OF HEALTH,

HEREBY GIVE NOTICE, THAT IN CONSEQUENCE OF THE

Church-yards at Dudley

Being so full, no one who has died of the CHOLERA will be permitted to be buried after SUNDAY next, (To-morrow) in either of the Burial Grounds of St. Thomas's, or St. Edmund's, in this Town.

All Persons who die from CHOLERA, must for the future be buried in the Church-yard at Netherton.

BOARD of HEALTH, DUDLEY.
September 1st, 1832.

W. MAURICE, PRINTER, HIGH STREET, DUDL

FIG. 1.1 Cartel en un cementerio de Dudley, Inglaterra, de 1839. (De Dudley Public Library, Dudley, England.)

Permítanos comparar las principales causas de muerte en Estados Unidos en 1900 y en 2014 (fig. 1.2). Las categorías de causas han sido codificadas mediante colores, como se describe en el pie de la figura. En 1900, las principales causas de muerte eran la neumonía y la gripe, seguidas por la tuberculosis, la diarrea y la enteritis. En 2014, las principales causas de muerte fueron las cardiopatías, el cáncer, la enfermedad respiratoria crónica de vías bajas y los traumatismos no deliberados. ¿Qué ha cambiado? Durante el siglo XX se ha producido un cambio importante de las causas de muerte en Estados Unidos. En 1900, las tres principales causas de muerte eran enfermedades infecciosas; sin embargo, en la actualidad se trata de enfermedades crónicas que en la mayoría de los casos no son transmisibles o de origen infeccioso. Por tanto, los tipos de investigaciones, intervenciones y servicios necesarios en la actualidad difieren de los que eran necesarios en Estados Unidos en 1900.

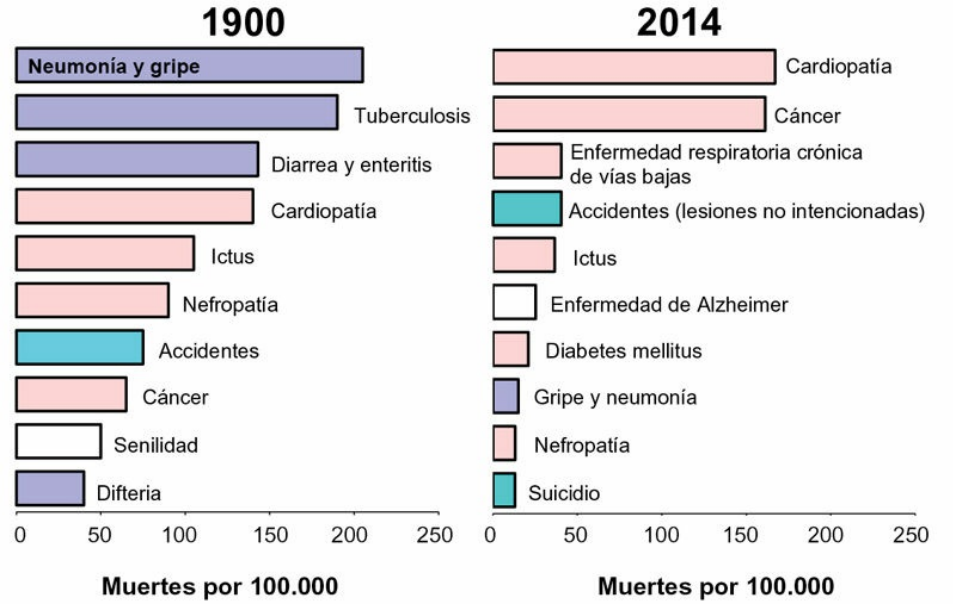


FIG. 1.2 Las diez principales causas de muerte en Estados Unidos en 1900 y en 2014. Aunque las definiciones de las enfermedades en esta figura no son comparables exactamente en 1900 y en 2014, las barras coloreadas de los gráficos representan enfermedades crónicas *(rosa)*, enfermedades infecciosas *(morado)*, lesiones *(turquesa)* y enfermedades relacionadas con el envejecimiento *(blanco)*. (Modificado de Grove RD, Hetzel AM. *Vital Statistics Rates of the United States, 1940–1960*. Washington, DC, US Government Printing Office, 1968; y Kochanek KD, Murphy SL, Xu JQ, Tejada-Vera B. Deaths: Final data for 2014. *Natl Vital Stat Rep*. 2016;65(4):1–122. [Hyattsville, MD: National Vital Statistics Report.])

El patrón de aparición de enfermedades en los países en vías de desarrollo observado en la actualidad es a menudo similar al observado en Estados Unidos en 1900: las enfermedades infecciosas siguen siendo las causas principales de muerte. Sin embargo, a medida que los países se industrializan, manifiestan con una frecuencia cada vez mayor los patrones de mortalidad observados en la actualidad en los países desarrollados, de modo que la principal causa de mortalidad corresponde a las enfermedades crónicas (un concepto denominado a menudo «transición epidemiológica»). No obstante, incluso en los países industrializados, debido a la aparición de la infección por el virus de la inmunodeficiencia humana (VIH) y al aumento de la incidencia de tuberculosis, las enfermedades infecciosas se están convirtiendo de nuevo en un problema importante de salud pública. En la tabla 1.1 se muestran las 10 principales causas de muerte en Estados Unidos en 2014. Las tres causas principales (las cardiopatías, el cáncer y las enfermedades crónicas del sistema respiratorio inferior) representan casi el 55% de todas las muertes, una observación que sugiere cuáles son los objetivos específicos para la prevención si se quiere lograr una reducción significativa de la mortalidad.

Tabla 1.1

Las 10 principales causas de muerte en Estados Unidos en 2014

y sus porcentajes en relación con las muertes totales

Puesto	Causa de muerte	Número de muertes	Porcentaje (%) del total de muertes	Tasa de mortalidad*
	Todas las causas	**2.626.418**	**100**	**724,6**
1	Cardiopatías	614.348	23,4	167,0
2	Neoplasias malignas (cáncer)	591.699	22,5	161,2
3	Enfermedades respiratorias crónicas de vías bajas	147.,101	5,6	40,5
4	Accidentes (lesiones no intencionadas)	136.153	5,2	40,5
5	Enfermedades cerebrovasculares	133.103	5,1	36,5
6	Enfermedad de Alzheimer	93.541	3,6	25,4
7	Diabetes mellitus	76.488	2,9	20,9
8	Gripe y neumonía	55.227	2,1	15,1
9	Nefritis, síndrome nefrótico y nefrosis	48.146	1,8	13,2
10	Autolesiones intencionadas (suicidio)	42.773	1,6	13,0
	Cualquier otra causa	687.939	26,2	

Nota: Los porcentajes pueden no sumar 100 debido al redondeo.

* Las tasas son por 100.000 habitantes y están ajustadas por edad para la población estándar de EE. UU. del año 2010.

Datos de los Centers for Disease Control and Prevention: Xu JQ, Murphy SL, Kochanek KD, Arias E. *Mortality in the United States, 2015*. NCHS data brief, no 267. Hyattsville, MD: National Center for Health Statistics; 2016. https://www.cdc.gov/nchs/data/databriefs/db267_table.pdf. Consultado el 17 de abril de 2017.

Otra prueba de los cambios que han tenido lugar a lo largo del tiempo se muestra en la figura 1.3, donde se exponen los años de esperanza de vida en Estados Unidos al nacer y a los 65 años de edad en los años 1900, 1950 y 2014 según raza y sexo.

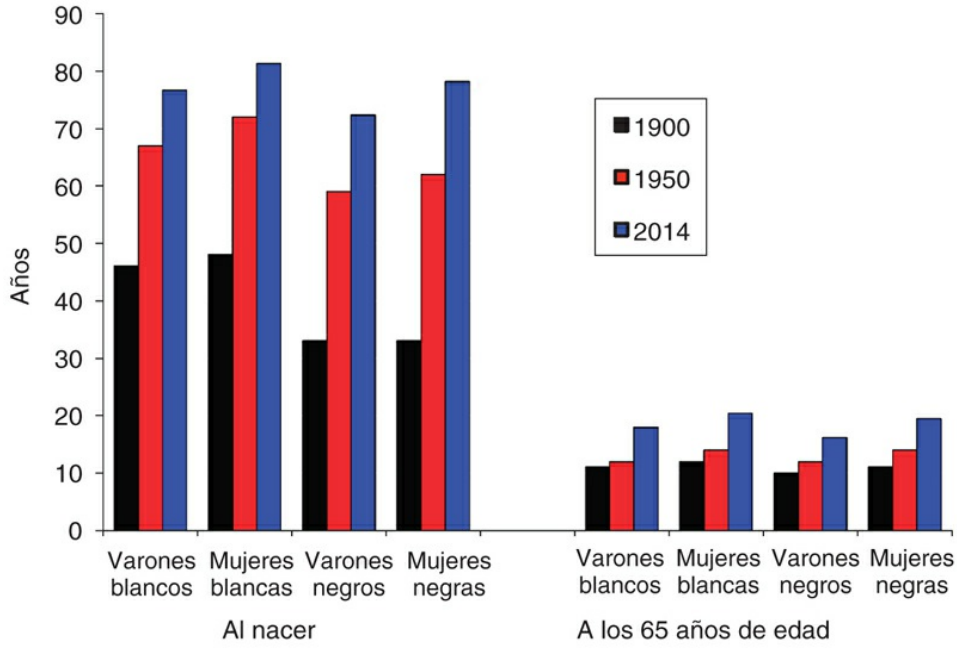


FIG. 1.3 Esperanza de vida al nacer y a los 65 años de edad por raza y sexo en Estados Unidos en 1900, 1950 y 2014. (Modificado de National Center for Health Statistics. *Health, United States, 1987 DHHS publication No. 88–1232.* Washington, DC, Public Health Service, March 1988; y National Center for Health Statistics. Health, United States, 2015: with special feature on racial and ethnic health disparities; 2016. https://www.cdc.gov/nchs/hus/contents2015.htm#015. Consultado el 2 de mayo de 2017.)

La esperanza de vida tras el nacimiento ha aumentado de manera importante en todos estos grupos. El aumento más significativo tuvo lugar de 1900 a 1950; a partir de 1950 ha sido mucho menor. Si observamos los años restantes de vida a partir de los 65 años, se ha producido muy poca mejoría de 1900 a 2014. Lo que justifica principalmente el aumento de la esperanza de vida al nacer es la disminución de la mortalidad en los lactantes y de la mortalidad por las enfermedades de la infancia. En relación con las enfermedades que afectan a los adultos, especialmente a los de 65 o más años de edad, hemos tenido mucho menos éxito ampliando la esperanza de vida, lo que sigue siendo un desafío importante.

Epidemiología y prevención

Un objetivo importante de los estudios epidemiológicos es la identificación de subgrupos de la población que posiblemente no hayan sido evaluados con anterioridad y que presentan un riesgo elevado de sufrir enfermedades. ¿Por qué debemos identificar estos grupos de alto riesgo? En primer lugar, si podemos identificar estos grupos de alto riesgo, podremos dirigir medidas preventivas, como programas de cribado para la detección precoz de enfermedades, a poblaciones que pueden beneficiarse con mayor probabilidad de intervenciones desarrolladas para dichas enfermedades. En África subsahariana, la orientación y la evaluación correspondientes al VIH y dirigidas hacia la población de hombres que desconocen su estado respecto a dicha infección pueden reducir eficazmente la epidemia, siempre en el contexto de la asistencia apropiada, el inicio del tratamiento antirretroviral y la continuidad asistencial.

En segundo lugar, si podemos identificar dichos grupos, podremos ser capaces de identificar los factores o las características específicas causantes de dicho riesgo elevado y podremos, por tanto, modificar dichos factores. Es importante tener en cuenta que dichos factores de riesgo pueden ser de dos tipos. Características como la edad, el sexo y la raza, por ejemplo, no son modificables, aunque pueden ayudarnos a identificar los grupos de alto riesgo. Por otro lado, características como la obesidad, el tabaquismo, la dieta, las prácticas sexuales y otros factores del estilo de vida pueden ser potencialmente modificables y pueden proporcionar, por tanto, una oportunidad para el desarrollo y la introducción de nuevos programas de prevención dirigidos a reducir o modificar exposiciones específicas o factores de riesgo.

Prevención primaria, secundaria y terciaria

Al estudiar la prevención, resulta útil distinguir entre prevención primaria, secundaria y terciaria (tabla 1.2).

Tabla 1.2

Los tres tipos de prevención

Tipos de prevención	Definición	Ejemplos
Primaria	Prevención del *desarrollo inicial* de una enfermedad	Vacunación, disminución de la exposición a un factor de riesgo
Secundaria	Detección precoz de una *enfermedad existente* para reducir la gravedad y las complicaciones	Cribado del cáncer
Terciaria	Reducción del *impacto de la enfermedad*	Rehabilitación del ictus

La *prevención primaria* se refiere a las acciones dirigidas a evitar el desarrollo de una enfermedad en una persona que se encuentra bien y que

(todavía) no presenta la enfermedad en cuestión. Por ejemplo, podemos vacunar a una persona frente a ciertas enfermedades para que la enfermedad no se presente nunca o, si la enfermedad es inducida por las condiciones ambientales, podemos prevenir la exposición de la población al factor ambiental involucrado y, por tanto, evitar el desarrollo de la enfermedad. La prevención primaria es nuestro principal objetivo. Por ejemplo, sabemos que la mayoría de los cánceres de pulmón son prevenibles. Si se consiguiera que la población dejara de fumar, sería posible eliminar el 80-90% de los cánceres de pulmón en el ser humano. Sin embargo, aunque nuestro objetivo es evitar que se produzcan enfermedades en la población, aún carecemos de datos biológicos, clínicos y epidemiológicos para muchas enfermedades (p. ej., la enfermedad de Alzheimer y el cáncer prostático) sobre los que basar programas efectivos de prevención primaria.

La *prevención secundaria* trata de identificar a pacientes en los que el proceso de la enfermedad ya ha comenzado pero que aún no presentan síntomas ni signos de la enfermedad. Este periodo de la historia natural de la enfermedad se denomina *fase preclínica* de la enfermedad y se estudia en el capítulo 18. Cuando una persona presenta síntomas o signos clínicos, por lo general se asume que en las condiciones ideales dicha persona buscará y obtendrá asistencia sanitaria. Nuestro objetivo con la prevención secundaria es detectar la enfermedad antes de lo que se hubiera detectado con la asistencia habitual. Al detectar la enfermedad en una etapa más temprana de su historia natural, a menudo a través de pruebas de cribado, es de esperar que el tratamiento sea más sencillo y/o más eficaz. Por ejemplo, la mayoría de los casos de cáncer de mama en las mujeres de edad avanzada pueden detectarse por medio de la mamografía. Varios estudios recientes indican que el estudio rutinario de las heces en busca de sangre oculta puede detectar el cáncer de colon tratable en fases tempranas de su historia natural, aunque la colonoscopia es una prueba más fiable y también bastante más cara e invasiva. La base de la prevención secundaria es que, si podemos identificar la enfermedad en etapas más tempranas de su historia natural de lo que normalmente ocurriría, las medidas de intervención podrán ser más eficaces, con el consiguiente incremento de la esperanza de vida. Quizás podemos prevenir la mortalidad o las complicaciones de la enfermedad y utilizar tratamientos menos invasivos o menos costosos para lograrlo. En el capítulo 18 se estudian y evalúan las pruebas de cribado de las enfermedades y el papel de dicha intervención en la estrategia de la prevención de la enfermedad.

La *prevención terciaria* consiste en la prevención de complicaciones cuando ya se han presentado signos y síntomas de una enfermedad y esta ya ha sido diagnosticada, es decir, en pacientes que se encuentran en la fase clínica de su enfermedad. Suele lograrse por medio del tratamiento precoz y adecuado de la enfermedad combinado con abordajes complementarios, como la fisioterapia, que son diseñados para prevenir complicaciones de las articulaciones.

Dos enfoques sobre la prevención: una visión diferente

Dos posibles enfoques preventivos son el enfoque sobre la población y el enfoque sobre la población de alto riesgo[2]. En el enfoque sobre la población, una medida preventiva se aplica ampliamente a una población entera. Por ejemplo, se pueden dar a toda una población consejos dietéticos prudentes para prevenir la enfermedad coronaria o consejos contra el tabaquismo, todo ello a través de los medios de comunicación y de diversas estrategias de educación para la salud. Un enfoque alternativo es el que dirige las medidas preventivas a un grupo de alto riesgo. Por ejemplo, las pruebas de cribado para el estudio del colesterol en los niños pueden limitarse a los niños de familias de alto riesgo. Claramente, las medidas aplicadas a toda una población deben ser relativamente poco costosas y no invasivas. Una medida que vaya a emplearse en un subgrupo de alto riesgo de la población puede ser más cara y también más invasiva o incómoda, pero también puede identificar correctamente las personas que padecen la enfermedad. En el capítulo 18 hay más información acerca de las pruebas de detección. Los enfoques sobre la población pueden considerarse acciones de salud pública, mientras que los enfoques sobre subgrupos de alto riesgo requieren más a menudo una acción clínica para identificar el subgrupo sobre el que van a dirigirse. Estos últimos enfoques (p. ej., la orientación respecto a la prevención) se limitan a menudo a breves encuentros con los médicos. La combinación de ambos enfoques resulta ideal en la mayoría de las situaciones. Estos enfoques se estudian con mayor profundidad en el capítulo 19.

Epidemiología y práctica clínica

La epidemiología es fundamental no solo para la salud pública sino también para la práctica clínica. La práctica de la medicina depende de datos poblacionales. Por ejemplo, si un médico escucha un soplo sistólico apical (un ruido cardiaco que se debe al paso de la sangre a través de las válvulas del corazón), ¿cómo sabe que indica una insuficiencia mitral? ¿Dónde se originó este conocimiento? El diagnóstico se basa en la correlación de los hallazgos clínicos (como los hallazgos de la auscultación: los sonidos oídos a través del estetoscopio) con los hallazgos encontrados en las autopsias o de la anatomía patológica quirúrgica y con los resultados de los estudios de ecocardiografía, resonancia magnética o cateterismo en un grupo extenso de pacientes. Por tanto, el proceso diagnóstico se basa en la población (v. cap. 5). Lo mismo puede aplicarse al pronóstico. Por ejemplo, un paciente pregunta a su médico: «¿Cuánto me queda de vida, doctor?» Y el médico le responde: «De seis meses a un año». ¿En qué basa el médico su pronóstico? El médico se basa en la experiencia con grupos grandes de pacientes que presentaban la misma enfermedad, que fueron estudiados en la misma etapa de la enfermedad y recibieron el mismo tratamiento. De nuevo, el pronóstico se basa en datos poblacionales (v. cap. 6). Por último, la selección del tratamiento apropiado también se basa en la población. Los ensayos clínicos aleatorizados que estudian los efectos de un tratamiento en grupos grandes de pacientes son el método ideal para identificar el tratamiento apropiado, lo que se denomina método de referencia (v. caps. 10 y 11). Por tanto, los conceptos y los datos basados en la población subyacen en los procesos críticos de la práctica clínica, como el diagnóstico, el pronóstico y la selección del tratamiento. En efecto, el médico aplica al paciente durante la exploración un modelo de probabilidad basado en la población.

La figura 1.4 presenta a un médico que demuestra que la práctica de la medicina clínica se basa de modo importante en conceptos poblacionales. Lo que se presenta con humor es un comentario verdadero de un aspecto de la práctica pediátrica; el pediatra a menudo emite un diagnóstico basándose en lo que los padres le dicen por teléfono y en lo que él sabe acerca de las enfermedades, como las infecciones víricas o bacterianas, que «están circulando» en la comunidad. Por tanto, los datos disponibles acerca de las enfermedades de la comunidad pueden ser de gran utilidad para conocer el diagnóstico, incluso aunque no sean concluyentes. Los datos referidos a la etiología de la faringitis en función de la edad del niño son especialmente importantes (fig. 1.5). Si la infección se produce en los primeros años de vida, probablemente será de origen vírico. Si ocurre entre los 4 y los 7 años, probablemente será de origen estreptocócico. En un niño mayor es más frecuente la infección por *Mycoplasma*. Aunque estos datos no dan el diagnóstico, proporcionan al médico o a otro profesional sanitario pistas valiosas sobre el microorganismo o microorganismos sospechosos.



FIG. 1.4 «Usted tiene lo que quiera que sea que está circulando por ahí.» (Al Ross/The New Yorker Collection/The Cartoon Bank.)

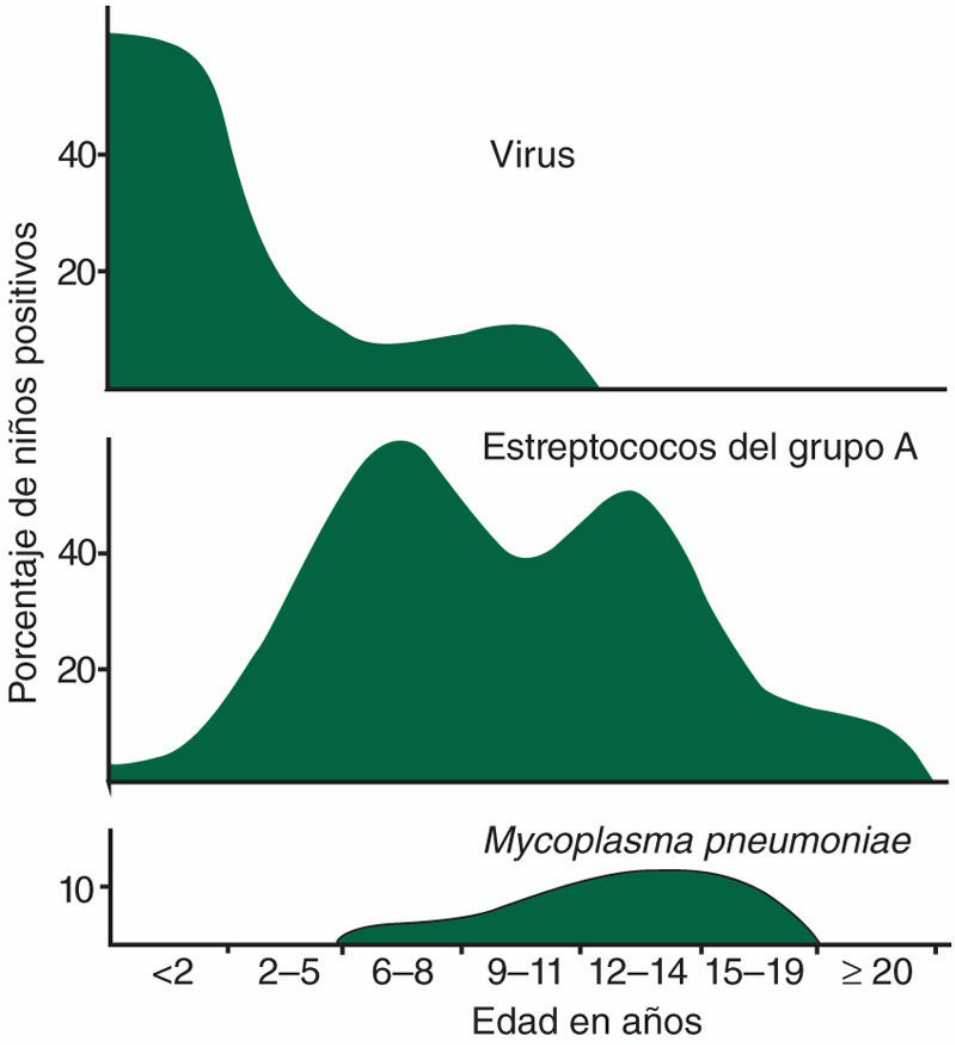


FIG. 1.5 Frecuencia de microorganismos según la edad en niños con faringitis, de 1964 a 1965. (De Denny FW. The replete pediatrician and the etiology of lower respiratory tract infections. *Pediatr Res.* 1969;3:464-470.)

Planteamiento epidemiológico

¿Cómo procede el epidemiólogo para identificar la causa de una enfermedad? El razonamiento epidemiológico es un proceso que tiene múltiples pasos. El primer paso es determinar si existe una asociación entre la exposición a un factor (p. ej., un agente ambiental) o una característica de una persona (p. ej., la concentración sérica de colesterol elevada) y la aparición de la enfermedad en cuestión. Esto se realiza estudiando las características de los grupos y de los individuos.

Si encontramos que de hecho existe una asociación entre una exposición y la enfermedad, ¿es necesariamente una relación causal? No, no todas las asociaciones son causales. Por tanto, el segundo paso consiste en intentar deducir inferencias apropiadas acerca de una posible relación causal a partir de los patrones de las asociaciones que hayan sido encontrados previamente. Estos pasos se analizan en detalle en el capítulo 14.

La epidemiología a menudo comienza con datos descriptivos. Por ejemplo, en la figura 1.6 se exponen las tasas de gonorrea en Estados Unidos en 2015 por estado. Claramente, existen importantes variaciones regionales en los casos declarados de gonorrea. La primera pregunta que debemos plantearnos cuando observamos dichas diferencias entre dos grupos o dos regiones o en dos momentos distintos es: «¿son reales estas diferencias?» En otras palabras: ¿son de calidad comparable los datos de cada área? Antes de intentar interpretar los datos, deberíamos asegurarnos de que son válidos. Si las diferencias son reales, a continuación debemos preguntarnos: «¿Por qué se han producidos estas diferencias?» ¿Existen diferencias en las exposiciones potenciales entre las áreas de alto riesgo y las de bajo riesgo o existen diferencias entre las poblaciones que viven en dichas áreas? Aquí es donde la epidemiología comienza sus estudios.

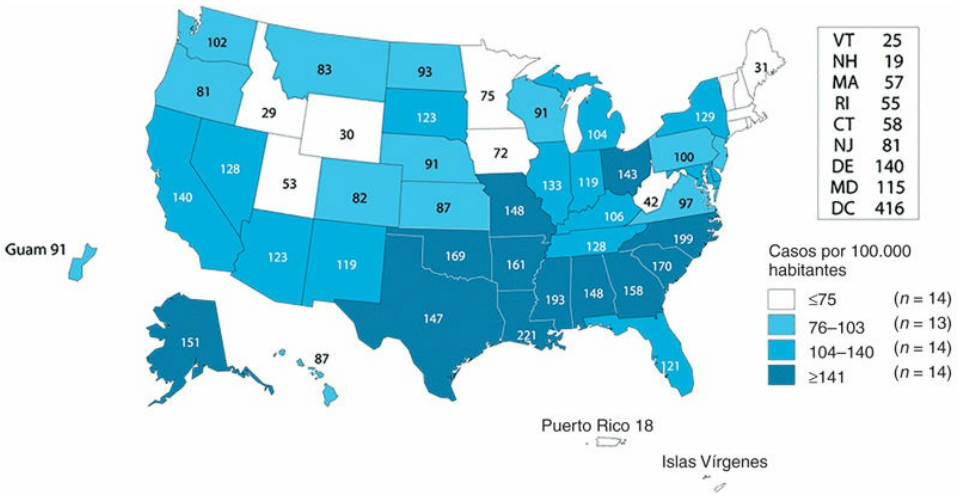


FIG. 1.6 Gonorrea: casos declarados por cada 100.000 habitantes, Estados Unidos y territorios, 2015. La tasa total de casos comunicados de gonorrea en

Estados Unidos continental y territorios exteriores (Guam, Puerto Rico e Islas Vírgenes) fue de 122,7 por cada 100.000 personas de la población general. (De Gonorrhea—rates by state, United States and outlying areas; 2015. https://www.cdc.gov/std/stats15/figures/15.htm. Consultado el 19 de abril de 2015.)

Hace muchos años se observó que en las comunidades en las que había diferencias en la concentración natural de flúor en el agua de bebida también existían diferencias en la frecuencia de caries dental en los dientes permanentes de los residentes. Las comunidades con concentraciones bajas de flúor natural presentaban una frecuencia elevada de caries y las comunidades con concentraciones más elevadas de flúor en el agua de bebida presentaban menor incidencia de caries (fig. 1.7). Esos hallazgos sugerían que el flúor podía ser una medida preventiva eficaz si se añadiese artificialmente al agua de bebida. Así pues, se efectuó un ensayo para comprobar la hipótesis. Aunque, idealmente, quisiéramos aleatorizar un grupo de personas para que recibiese flúor y otro para que no lo recibiese, esto no sería posible efectuarlo con el agua de bebida porque cada comunidad por lo general comparte un suministro de agua común. Por tanto, para realizar el ensayo se eligieron dos comunidades similares del norte del estado de Nueva York, Kingston y Newburgh. Se utilizó el índice CAO, que tiene en cuenta los dientes con caries, ausentes y obturados (empastados). Se recogieron datos basales en ambas ciudades. Al comienzo del estudio los índices CAO eran comparables en cada grupo de edad en las dos comunidades. Posteriormente se fluoró el agua en Newburgh y los niños fueron explorados de nuevo un decenio después. En la figura 1.8 se muestra que en cada grupo de edad el índice CAO se redujo en Newburgh de modo significativo 10 años después, mientras que no se observaron cambios en Kingston. El resultado sugiere con firmeza que el flúor estaba evitando las caries.

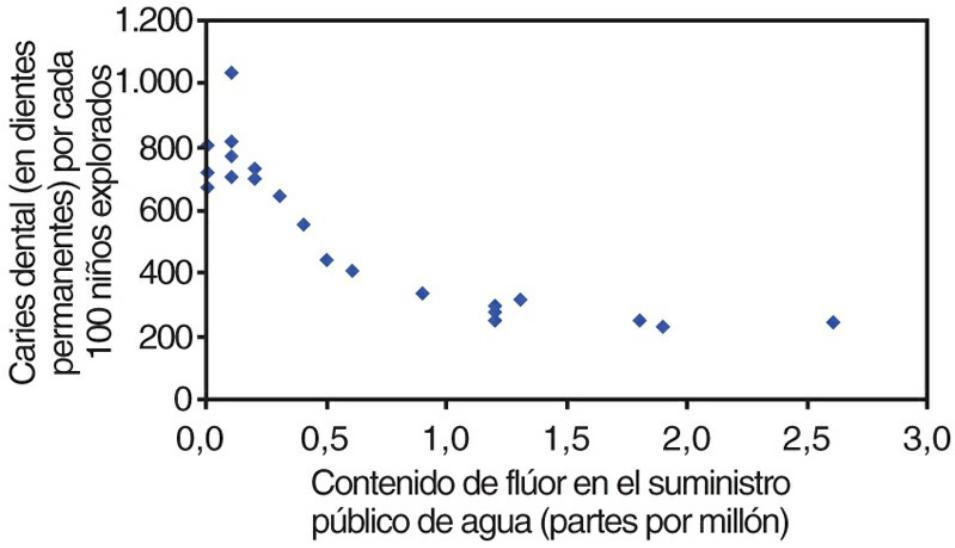


FIG. 1.7 Relación entre la frecuencia de caries dental en los dientes permanentes de los niños y el contenido de flúor en el suministro público de agua. (Modificado de Dean HT, Arnold FA Jr, Elvove E. Domestic water and dental

caries: V. Additional studies of the relation of fluoride in domestic waters to dental caries experience in 4,425 white children aged 12 to 14 years of 13 cities in 4 states. *Public Health Rep.* 1942;57:1155-1179.)

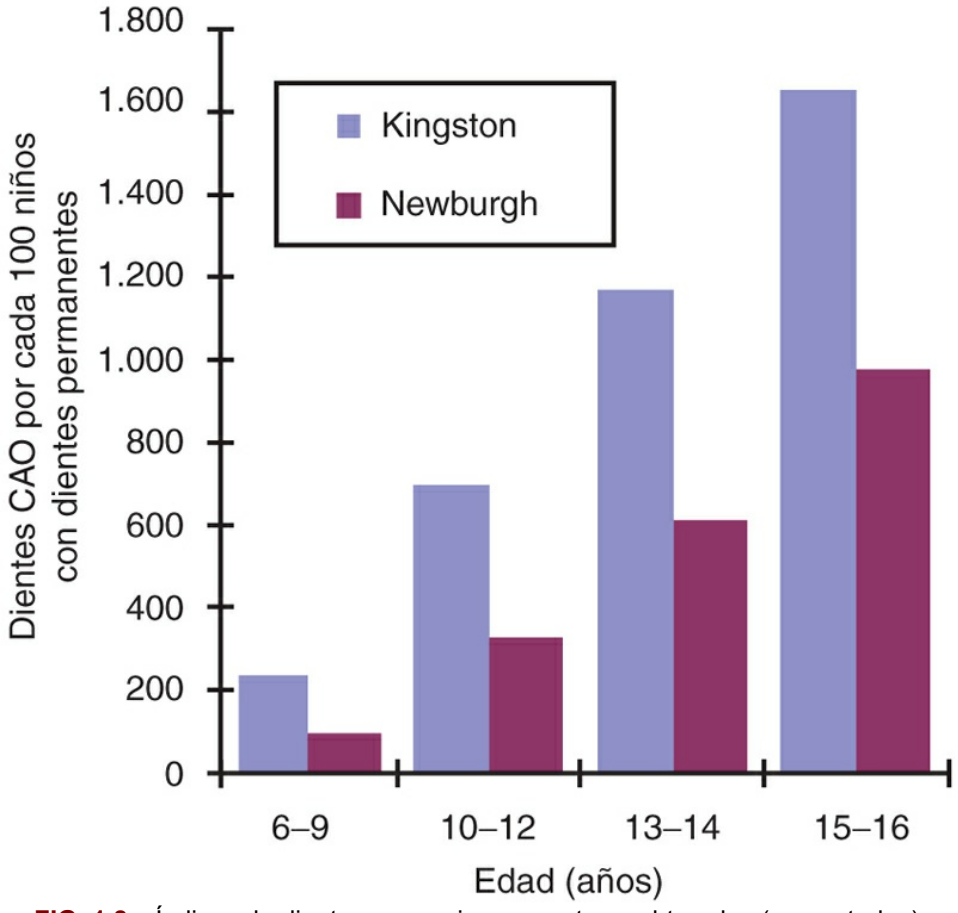


FIG. 1.8 Índices de dientes con caries, ausentes y obturados (empastados) *(CAO)* después de 10 años de fluoración, 1954-1955. *CAO,* dientes con caries, ausentes y obturados. (Modificado de Ast DB, Schlesinger ER. The conclusion of a 10-year study of water fluoridation. *Am J Public Health.* 1956;46:265-271. Copyright 1956 by the American Public Health Association. Adaptado con autorización.)

Fue posible ir más allá para intentar demostrar una relación causal entre la ingesta de flúor y la baja incidencia de caries. El tema de la fluoración de los suministros de agua ha sido muy controvertido y en algunas comunidades en las que el agua ha sido fluorada se han producido referéndums para interrumpir la fluoración. Así, fue posible estudiar el índice CAO en comunidades como Antigo, Wisconsin, en las que se fluoró el suministro de agua y, posteriormente, tras un referéndum, se interrumpió la fluoración. Como se observa en la figura 1.9, tras interrumpir la fluoración, el índice CAO se elevó. Esto proporcionó una prueba más de que el flúor servía para la prevención de la caries dental.

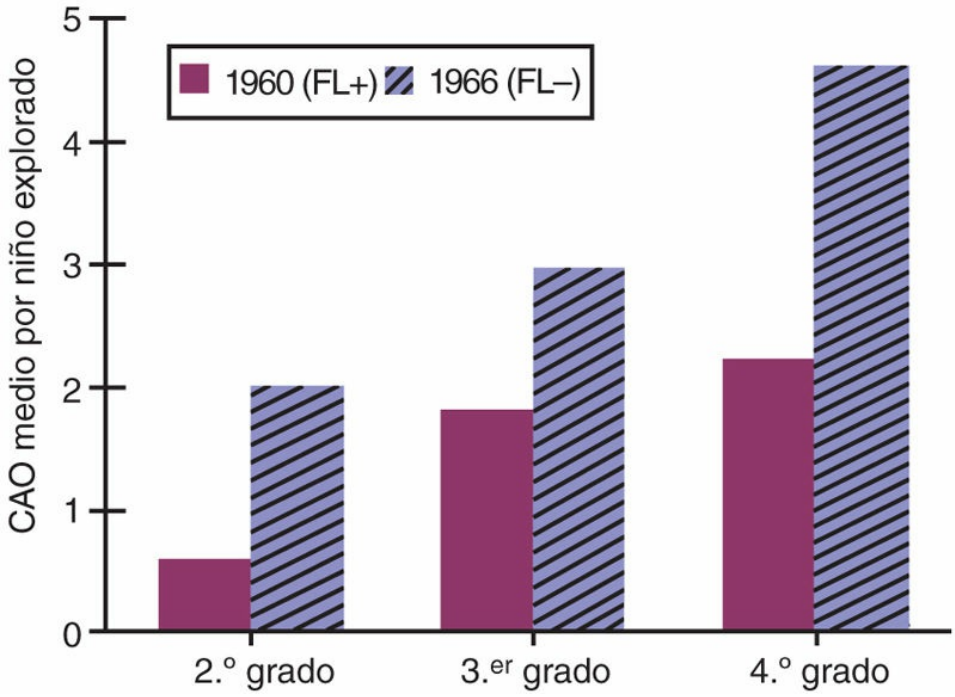


FIG. 1.9 Efecto de la interrupción de la fluoración en Antigo, Wisconsin; noviembre de 1960. *CAO,* dientes con caries, ausentes y obturados; *FL+,* durante la fluoración; *FL−*, tras interrumpir la fluoración. (Modificado de Lemke CW, Doherty JM, Arra MC. Controlled fluoridation: The dental effects of discontinuation in Antigo, Wisconsin. *J Am Dental Assoc.* 1970;80:782-786. Reproducido con autorización de ADA Publishing Co., Inc.)

De observaciones a acciones preventivas

En esta sección se analizan tres ejemplos históricos que demuestran cómo las observaciones epidemiológicas han conducido a medidas preventivas para las poblaciones humanas.

Ignáz Semmelweis y la fiebre puerperal

Ignáz Semmelweis (fig. 1.10) nació en 1818 y comenzó a estudiar Derecho hasta que abandonó sus estudios para comenzar la carrera de Medicina. Se especializó en obstetricia y se interesó en un problema de salud pública y clínico importante: la fiebre puerperal («puerperal» significa relacionado con el parto o con el periodo posterior al nacimiento).



FIG. 1.10 Retrato de Ignáz Philipp Semmelweis. (De la National Library of Medicine.)

A comienzos del siglo XIX, la fiebre puerperal era una causa importante de muerte en las mujeres poco tiempo después del parto, con tasas de mortalidad de incluso el 25%. En la época eran populares numerosas teorías acerca de su etiología, como toxinas atmosféricas, «constituciones epidémicas» de algunas mujeres, aire pútrido o influencias solares y magnéticas. Este periodo fue una época de interés creciente en la anatomía patológica. Como la causa de la fiebre puerperal era desconocida, surgió un gran interés en correlacionar los hallazgos encontrados en las autopsias de las mujeres que habían fallecido por la enfermedad con las manifestaciones clínicas que presentaban tras el parto.

Semmelweis fue nombrado responsable de la First Obstetrical Clinic del Allgemeine Krankenhaus (Hospital General) de Viena en julio de 1846. En esa época había dos clínicas obstétricas. Las mujeres embarazadas eran ingresadas para el parto en la primera o la segunda clínica basándose en un horario alternante de 24 horas. De la primera clínica se encargaban médicos y estudiantes de medicina y de la segunda, matronas. Los médicos y los estudiantes de medicina comenzaban su jornada realizando autopsias de las mujeres que habían fallecido de fiebre puerperal y posteriormente atendían a las mujeres hospitalizadas para dar a luz en la primera clínica. Las matronas encargadas de la segunda clínica no realizaban autopsias. Semmelweis estaba asombrado de las tasas de mortalidad de las dos clínicas en 1842 (fig. 1.11). La mortalidad de la primera clínica era más del doble que la de la segunda clínica (16% frente a 7%).

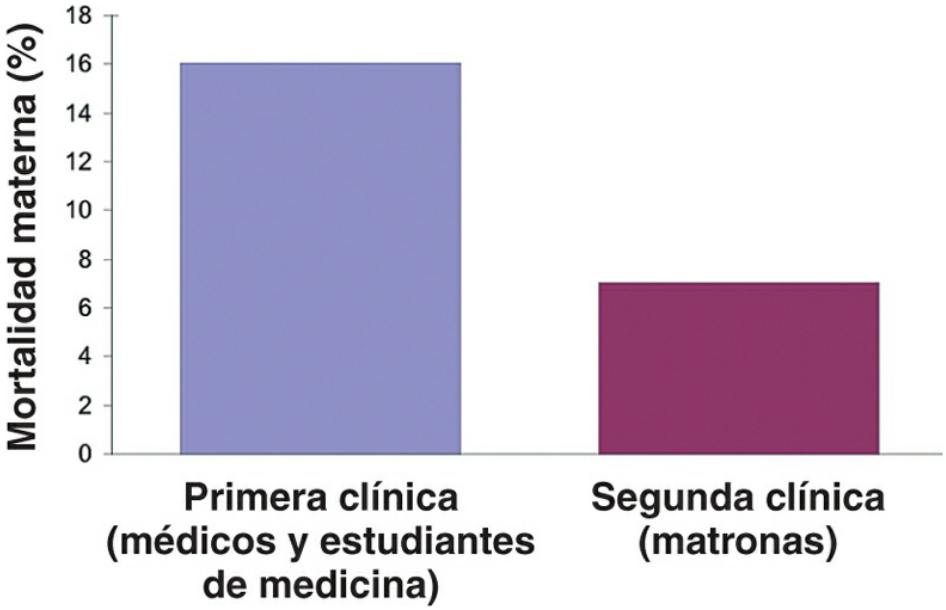


FIG. 1.11 Mortalidad materna debida a fiebre puerperal. Primera y segunda clínica, Hospital General de Viena, Austria, 1842. (Modificado de Centers for Disease Control and Prevention: Hand hygiene in health care settings—

Supplemental.
www.cdc.gov/handhygiene/download/hand_hygiene_supplement.ppt. Consultado el 11 de abril de 2013.)

Semmelweis llegó a la conclusión de que la mortalidad era mayor en la primera clínica porque los médicos y los estudiantes de medicina iban directamente desde las autopsias a tratar pacientes. Muchas de las mujeres que se encontraban de parto eran sometidas a múltiples exploraciones por parte de médicos y de estudiantes de medicina que estaban formándose en obstetricia. A menudo dichas exploraciones manuales producían traumatismos en los tejidos vaginales y uterinos. Semmelweis sugirió que las manos de los médicos y los estudiantes de medicina estaban transmitiendo partículas causantes de enfermedad desde los cadáveres de las mujeres a las mujeres que estaban a punto de dar a luz. Sus sospechas fueron confirmadas en 1847, cuando su amigo y colega Jakob Kolletschka murió de una infección que adquirió cuando se pinchó accidentalmente con el bisturí de un estudiante de medicina que estaba practicando una autopsia. La autopsia de Kolletschka demostró una anatomía patológica muy similar a la de las mujeres que fallecían de fiebre puerperal. Semmelweis concluyó que los médicos y los estudiantes de medicina transmitían la infección desde la sala de autopsias a las pacientes de la primera clínica y que esta era la causa de la alta tasa de mortalidad por fiebre puerperal en la primera clínica. La tasa de mortalidad en la segunda clínica seguía siendo baja porque las matronas que trabajaban en la segunda clínica no mantenían ningún contacto con la sala de autopsias.

Semmelweis desarrolló e implementó después una normativa para los médicos y los estudiantes de medicina de la primera clínica, unas normas diseñadas para evitar la fiebre puerperal. Exigió a los médicos y a los estudiantes de medicina de la primera clínica que se lavaran las manos y se cepillaran las uñas tras finalizar las autopsias y antes de explorar a las pacientes. Como se observa en la figura 1.12, la mortalidad en la primera clínica se redujo en 1848 del 12,2% al 2,4%, una tasa comparable a la observada en la segunda clínica. Cuando Semmelweis fue sustituido posteriormente por un obstetra que no comulgaba con sus teorías y que eliminó la norma que exigía el lavado de manos, la tasa de mortalidad por fiebre puerperal aumentó de nuevo en la primera clínica, prueba evidente que apoya la relación causal.

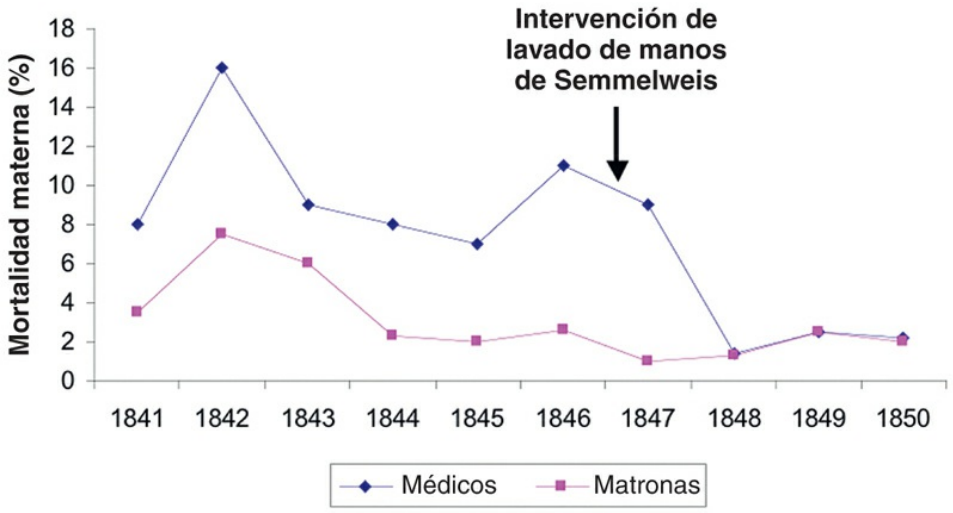


FIG. 1.12 Mortalidad materna debida a fiebre puerperal, por tipo de profesional sanitario, Hospital General de Viena, Austria, 1841-1850. (Modificado de Mayhall GC. *Hospital Epidemiology and Infection Control,* 2.ª ed. Filadelfia, Lippincott Williams & Wilkins; 1999.)

Desafortunadamente, durante muchos años Semmelweis se negó a presentar sus hallazgos en congresos importantes o a publicar sus estudios en revistas médicas. El hecho de no proporcionar pruebas científicas que apoyasen su teoría fue responsable al menos parcialmente de la falta de aceptación por parte de la comunidad médica de su hipótesis sobre la causa de la fiebre puerperal y la intervención propuesta más adelante de lavarse las manos antes de explorar a las pacientes. Entre otros factores que fomentaron la resistencia a su teoría se encontraba la reticencia de los médicos a aceptar la conclusión de que al transmitir el microorganismo responsable de la fiebre puerperal habían sido responsables involuntariamente de la muerte de un gran número de mujeres. Además, los médicos afirmaban que lavarse las manos antes de explorar a cada paciente sería una tarea que llevaría mucho tiempo. Otro factor importante es que Semmelweis era, cuando menos, poco diplomático y se había enemistado con muchos médicos de prestigio. Debido a todos estos factores, pasaron muchos años antes de que se adoptara ampliamente la norma de lavarse las manos. En 2003 se publicó una excelente biografía de Semmelweis, escrita por Sherwin Nuland[3].

Las lecciones de esta historia para el diseño de normas de eficacia probada siguen siendo importantes en la actualidad con respecto al desafío que supone aumentar la aceptación pública y por parte de los profesionales de recomendaciones de prevención basadas en la evidencia. Estas lecciones incluyen la necesidad de aportar pruebas científicas claras que apoyen la intervención propuesta, la necesidad de que la puesta en práctica de la intervención propuesta sea percibida como factible y económicamente viable, y la necesidad de preparar el trabajo preliminar necesario para implementar las recomendaciones, incluyendo la consecución de apoyo profesional, político y de la comunidad.

Años después se identificó a la infección por estreptococos como la causa principal de la fiebre puerperal. Las recomendaciones y los hallazgos

principales de Semmelweis finalmente surtieron efecto a nivel mundial en la práctica de la medicina. Sorprendentemente, sus observaciones y sus intervenciones propuestas precedieron al conocimiento de la teoría infecciosa y demostraron que es posible implementar una estrategia preventiva incluso cuando se desconoce la causa precisa de la enfermedad. Sin embargo, también resulta interesante que, aunque la necesidad del lavado de manos ha sido aceptada universalmente en la actualidad, estudios recientes han puesto de manifiesto que muchos médicos de hospitales de Estados Unidos y de otros países desarrollados todavía no siguen la recomendación de lavarse las manos (tabla 1.3).

Tabla 1.3

Cumplimiento del lavado de manos entre médicos, por especialidad, en los hospitales de la Universidad de Ginebra

Especialidad del médico	Número de médicos	Cumplimiento del lavado de manos (% de observaciones)
Medicina interna	32	87,3
Cirugía	25	36,4
Unidad de cuidados intensivos	22	62,6
Pediatría	21	82,6
Geriatría	10	71,2
Anestesiología	15	23,3
Medicina de urgencias	16	50,0
Otras	22	57,2

Datos de Pittet D. Hand hygiene among physicians: Performance, beliefs, and perceptions. *Ann Intern Med.* 2004;141:1-8.

Edward Jenner y la viruela

Edward Jenner (fig. 1.13) nació en 1749 y se interesó mucho por el problema de la viruela, que era un azote a nivel mundial. Por ejemplo, a finales del siglo XVIII, 400.000 personas morían de viruela al año y un tercio de los supervivientes sufrían ceguera como resultado de las infecciones corneales. Se sabía que los que sobrevivían a la viruela quedaban inmunizados frente a la enfermedad, por lo que una medida preventiva común era infectar a la población sana con la viruela administrándoles material obtenido de pacientes con viruela, una práctica denominada *variolización*. Sin embargo, este no era un método óptimo; algunos individuos variolizados fallecían de la viruela resultante, infectaban a otros de viruela o sufrían otras infecciones.



FIG. 1.13 Retrato de Edward Jenner. (Del Wellcome Historical Medical Museum and Library, Londres.)

Jenner estaba interesado en descubrir un método mejor y más seguro para prevenir la viruela. Observó, como otros lo habían hecho antes que él, que las ordeñadoras, las mujeres jóvenes encargadas de ordeñar las vacas, sufrían una enfermedad leve denominada viruela de las vacas. Posteriormente, durante los brotes de viruela, estas mujeres jóvenes parecían no contraer la enfermedad. En 1768, Jenner escuchó decir a una ordeñadora: «No puedo coger la viruela porque ya he pasado la viruela de las vacas». Estos datos eran observaciones y no se basaban en ningún estudio riguroso, pero Jenner estaba convencido de que la viruela de las vacas podía proteger frente a la viruela y decidió estudiar su hipótesis.

En la figura 1.14 se muestra un cuadro de Gaston Melingue en el que se observa a Jenner realizando la primera vacunación en 1796. (El término «vacunación» deriva del término latino *vacca.*) En este cuadro, una ordeñadora, Sarah Nelmes, se está vendando la mano de la que le acaban de extraer material de la viruela de las vacas. Jenner está administrando dicho material a James Phipps, un «voluntario» de 8 años. Jenner estaba tan convencido de que la viruela vacunal sería protectora que 6 semanas después, con el fin de comprobar su convicción, inoculó al niño material que acabada de ser obtenido de una pústula de viruela. El niño no contrajo la enfermedad. En este capítulo no queremos abordar las cuestiones éticas y las implicaciones de este experimento. (Evidentemente, Jenner no tuvo que justificar su estudio

ante un comité de ética.) En cualquier caso, los resultados de la primera vacunación y de las que siguieron fueron evitar a literalmente millones de seres humanos de todo el mundo la discapacidad y la muerte debidas al azote de la viruela. El punto importante es que Jenner no tenía ningún conocimiento sobre virus ni sobre la biología de la enfermedad; se basó únicamente en datos observacionales que le proporcionaron la base para una intervención preventiva.



FIG. 1.14 Una de las primeras vacunaciones de Edward Jenner, de Gaston Melingue. (Reproducido con autorización de la Bibliothèque de l'Académie Nationale de Médecine, París, 2007.)

En 1967, la Organización Mundial de la Salud (OMS) inició campañas internacionales para erradicar la viruela utilizando vacunas con el virus vacunal (de la viruela de las vacas). Se ha estimado que hasta ese momento la viruela afectaba a 15 millones de personas anualmente en todo el mundo, de los que 2 millones fallecían y millones quedaban ciegos o sufrían deformidades. En 1980, la OMS certificó que la viruela había sido erradicada. El programa de erradicación de la viruela[4], dirigido en la época por el Dr. D. A. Henderson (fig. 1.15), es uno de los mayores logros en la prevención de

enfermedades en la historia del ser humano. La OMS estimó que a lo largo de un periodo de 20 años se habían evitado 350 millones de nuevos casos. Sin embargo, tras los ataques terroristas al World Trade Center, en los que fallecieron casi 3.000 personas el 11 de septiembre de 2001, surgió una preocupación a nivel mundial por la posibilidad de un ataque terrorista biológico a raíz de los ataques con ántrax de 2001. Irónicamente, la posibilidad de que el virus de la viruela pudiese ser usado para dicho fin volvió a abrir debates relacionados con la viruela y la vacunación que muchos pensaban que habían quedado relegados permanentemente a la historia por los esfuerzos y éxitos obtenidos para erradicar la enfermedad. La magnitud de la amenaza de terrorismo biológico con el virus de la viruela, junto con aspectos relacionados con el riesgo de la vacuna (tanto para los vacunados como para los que pudieran entrar en contacto con vacunados, especialmente en el ámbito hospitalario), son algunos de los muchos que deben ser tenidos en cuenta. Sin embargo, a menudo solo disponemos de datos limitados o equívocos sobre estos temas que puedan guiar el desarrollo de una política pública preventiva relacionada con la amenaza del terrorismo biológico utilizando como arma la viruela.



FIG. 1.15 Fotografía del Dr. D. A. Henderson (1928–2016), que dirigió el programa de erradicación de la viruela de la Organización Mundial de la Salud.

John Snow y el cólera

Otro ejemplo del traslado de observaciones epidemiológicas a políticas públicas lo inmortalizó John Snow, cuyo retrato se muestra en la figura 1.16. Snow vivió en el siglo XIX y era conocido como el anestesista que administró cloroformo a la reina Victoria durante un parto. No obstante, la verdadera pasión de Snow era la epidemiología del cólera, una enfermedad que era un problema importante en Inglaterra a mediados del siglo XIX. En la primera semana de septiembre de 1854, cerca de 600 personas que vivían a unas manzanas del pozo de agua de Broad Street en Londres fallecieron de cólera. En esa época, el Registrador General era William Farr. Snow y Farr tenían un desacuerdo importante sobre la causa del cólera. Farr era partidario de la denominada teoría *miasmática* de la enfermedad. Según esta teoría, aceptada

comúnmente en dicha época, la enfermedad era transmitida por un miasma o nube que se mantenía próxima a la superficie de la tierra. Si esto fuese cierto, sería de esperar que las personas que viviesen a baja altitud presentasen un riesgo más elevado de contraer la enfermedad transmitida por esta nube que aquellas que viviesen en localizaciones más elevadas.



FIG. 1.16 Retrato de John Snow. (Retrato al óleo de Thomas Jones Baker, 1847, en Zuck D. Snow, Empson and the Barkers of Bath. *Anaesthesia.* 2001;56:227-230.)

Farr recopiló datos para apoyar su hipótesis (tabla 1.4). Los datos son bastante compatibles con su hipótesis: cuanto menor era la elevación, mayor era la mortalidad por el cólera. Snow no estaba de acuerdo y creía que el cólera se transmitía a través del agua contaminada (fig. 1.17). En Londres, en aquella época, la población obtenía agua suscribiendo un contrato con alguna de las compañías de suministro de agua. Los lugares de obtención de agua de

estas compañías se encontraban en una parte muy contaminada del río Támesis. En un momento, una de las compañías, la Lambeth Company, por motivos técnicos, no relacionados con la salud, cambió su zona de obtención de agua a una parte más alta del río Támesis, menos contaminada. Las otras compañías no modificaron su zona de obtención de agua. Snow razonó, entonces, que, según su hipótesis de que el agua contaminada era la causante del cólera, la mortalidad debida al cólera sería inferior en la población que obtenía el agua de la compañía Lambeth que en la que la obtenía de las otras compañías. Snow llevó a cabo lo que hoy en día denominamos «epidemiología de a pie»; fue de casa en casa contando todas las muertes por cólera en cada casa y determinando qué compañía suministraba agua a cada casa.

Tabla 1.4

Muertes por cólera por 10.000 habitantes en función de la elevación de la residencia por encima del nivel del mar, Londres, 1848-1849

Elevación por encima del nivel del mar (pies)	Número de muertes
<20	102
20-40	65
40-60	34
60-80	27
80-100	22
100-120	17
340-360	8

Datos de Farr W. *Vital Statistics: A Memorial Volume of Selections from the Reports and Writings of William Farr* (editado por el Sanitary Institute of Great Britain por Noel A. Humphreys). Londres: The Sanitary Institute; 1885.



A DROP OF LONDON WATER.

FIG. 1.17 Una gota de agua del Támesis, dibujada por *Punch* en 1850. (De The wonders of a London water drop. *Punch Magazine.* 11 de mayo, 1850;461:188.)

Los hallazgos de Snow se exponen en la tabla 1.5. La tabla muestra el número de casas, el número de muertes por cólera y las muertes por cada 10.000 casas. Aunque esta tasa no es el parámetro ideal, porque cada casa puede albergar diferentes números de habitantes, no es una mala aproximación. Observamos que en las casas en las que el suministro dependía de la compañía Southwark and Vauxhall, que obtenía el agua de una parte contaminada del río Támesis, la tasa de mortalidad fue de 315 fallecidos por cada 10.000 casas. En los hogares que obtenían el suministro de la compañía Lambeth, que había reubicado su lugar de captación aguas arriba, la tasa fue de solo 38 muertes por cada 10.000 casas. Sus datos fueron tan convincentes que hicieron que Farr, el Registrador General, exigiera que el registro de cada distrito del sur de Londres informase de qué compañía suministraba agua a cada casa en la que alguien hubiese fallecido por el cólera. Hay que recordar que en la época de Snow no se conocía la existencia del microorganismo enterotóxico *Vibrio cholerae*. Nada se conocía acerca de la biología de la enfermedad. La conclusión de Snow de que el agua contaminada se asociaba con el cólera se basaba exclusivamente en datos observacionales[5].

Tabla 1.5

Muertes debidas al cólera por cada 10.000 residencias. De Source of Water Supply, Londres, 1854

Suministro de agua	Número de residencias	Muertes debidas al cólera	Muertes por 10.000 residencias
Southwark y Vauxhall Co.	40.046	1.263	315
Lambeth Co.	26.107	98	38
Otros distritos de Londres	256.423	1.422	56

Datos modificados de Snow J. On the mode of communication of cholera. En: *Snow on Cholera: A Reprint of Two Papers by John Snow, M.D.* Nueva York: The Commonwealth Fund; 1936.

La cuestión es que, aunque es muy importante para nosotros maximizar nuestro conocimiento de la biología y la patogenia de la enfermedad, no siempre es necesario conocer todos los detalles del posible mecanismo patogénico para ser capaz de prevenir una enfermedad. Por ejemplo, sabemos que prácticamente todos los casos de fiebre reumática y de cardiopatía reumática eran secundarios a una infección estreptocócica. A pesar de que los estreptococos han sido estudiados y analizados exhaustivamente, todavía no conocemos cómo y por qué causan la fiebre reumática. Sabemos que tras una infección estreptocócica grave, como la que se observa en los reclutas militares, la fiebre reumática no se produce en 97 de cada 100 personas infectadas. En la población civil, como en escolares, en los que la infección es menos grave, la fiebre reumática se presenta en solo 3 de cada 1.000 escolares infectados, pero no en los 997 restantes[6]. ¿Por qué no se produce la enfermedad en esos 97 reclutas y en los 997 escolares si han estado expuestos al mismo microorganismo? No lo sabemos. Desconocemos si la enfermedad es el resultado de una diferencia no detectada en el microorganismo o si se debe a un cofactor que puede facilitar la adhesión de los estreptococos a las células epiteliales. Lo que sí sabemos es que, incluso sin comprender por completo la cadena patogénica desde la infección por *Streptococcus* hasta la fiebre reumática, podemos prevenir prácticamente todos los casos de fiebre reumática si podemos evitar o tratar con rapidez y de modo adecuado las infecciones estreptocócicas, tal como ha ocurrido en Estados Unidos. La ausencia de conocimiento biológico sobre la patogenia no debería ser un obstáculo o una excusa para no implementar servicios preventivos eficaces.

Consideremos el tabaquismo y el cáncer de pulmón. No sabemos qué carcinógenos específicos de los cigarrillos causan el cáncer, pero sabemos que más del 80% de los casos de cáncer de pulmón están causados por el tabaquismo. Esto no significa que no debamos realizar investigaciones de

laboratorio para comprender mejor cómo los cigarrillos causan cáncer. Pero de nuevo, en paralelo a dicha investigación, debemos crear programas comunitarios y de salud pública dirigidos hacia el abandono del tabaquismo que sean eficaces y que estén basados en los datos observacionales disponibles en la actualidad.

En la figura 1.18 se muestran los datos de mortalidad del cáncer de mama y de pulmón en las mujeres de Estados Unidos. Las tasas de mortalidad por cáncer de mama se mantuvieron relativamente constantes a lo largo de varias décadas, pero en los primeros años de siglo XXI se observaron signos de disminución. Sin embargo, la mortalidad debida al cáncer de pulmón en las mujeres ha estado aumentando constantemente, aunque parece haber comenzado a estabilizarse e incluso a disminuir levemente en los últimos años. Desde 1987, en Estados Unidos cada año han fallecido más mujeres de cáncer de pulmón que de cáncer de mama. Así pues, nos enfrentamos con el cuadro trágico de una forma de cáncer prevenible en una gran proporción, el cáncer de pulmón, que se debe a una adicción personal, el tabaquismo, como la principal causa de muerte por cáncer en la actualidad en las mujeres estadounidenses.

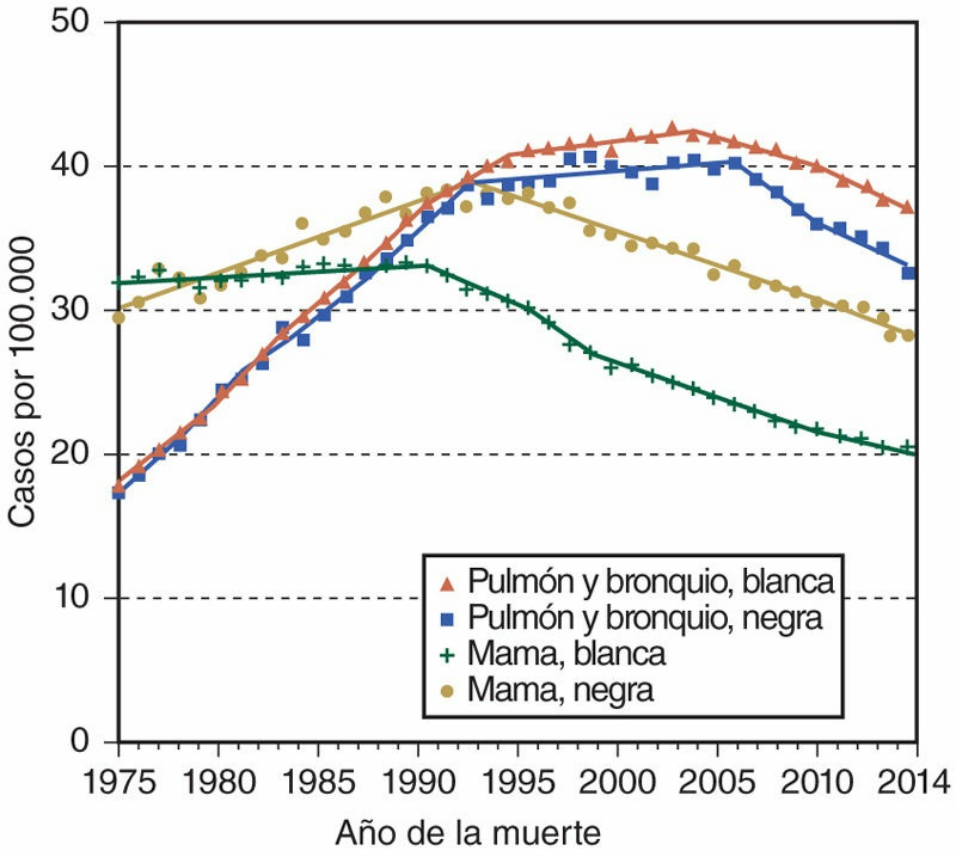


FIG. 1.18 Mortalidad por cáncer de mama frente a cáncer de pulmón: mujeres de raza blanca frente a mujeres de raza negra, Estados Unidos, 1975-2014, ajustada por edad al estándar de 2000. (De Howlader N, Noone AM, Krapcho M, et al, eds. *SEER Cancer Statistics Review, 1975–2014, National Cancer Institute.* Bethesda, MD, https://seer.cancer.gov/csr/1975_2014/, basado en datos de SEER de noviembre de 2016, publicados en la página web de SEER, abril de 2017.

https://seer.cancer.gov/csr/1975_2014/browse_csr.php; *figura 4.9.* Consultado el 14 de abril de 2017.)

Además, en 1993, el humo de tabaco en el ambiente (humo del tabaco de otras personas fumadoras) fue clasificado como un carcinógeno humano conocido por la Environmental Protection Agency, que atribuyó cerca de 3.000 muertes anuales por cáncer de pulmón en fumadores pasivos al humo de tabaco en el ambiente.

Cuando la frecuencia de una enfermedad disminuye, ¿de quién es el mérito?

A lo largo de los últimos cien años, la tasa de mortalidad de diversas enfermedades infecciosas comunes ha disminuido en Estados Unidos. Por ejemplo, las muertes por infecciones infantiles como la difteria, la tos ferina y la escarlatina (una infección estreptocócica) han disminuido de manera espectacular. Además, los fallecidos por tuberculosis en Estados Unidos han descendido de manera significativa.

Sería tentador relacionar estos descensos con mejoras en los tratamientos o en las vacunas disponibles para estas enfermedades durante este tiempo. Sin embargo, en 1971, Edward Kass publicó los gráficos que se exponen en la figura 1.19[7]. Estos gráficos demuestran que, para cada una de estas enfermedades, el principal descenso de su mortalidad se produjo muchos años antes de disponer de vacunas o tratamientos eficaces. En la figura 1.20 se muestra una presentación similar de la tendencia de la mortalidad a lo largo del tiempo de la fiebre reumática en el siglo XX[8]. Claramente, el principal descenso de la mortalidad por fiebre reumática se produjo bastante antes de que se dispusiera de penicilina y otros tratamientos antiestreptocócicos.

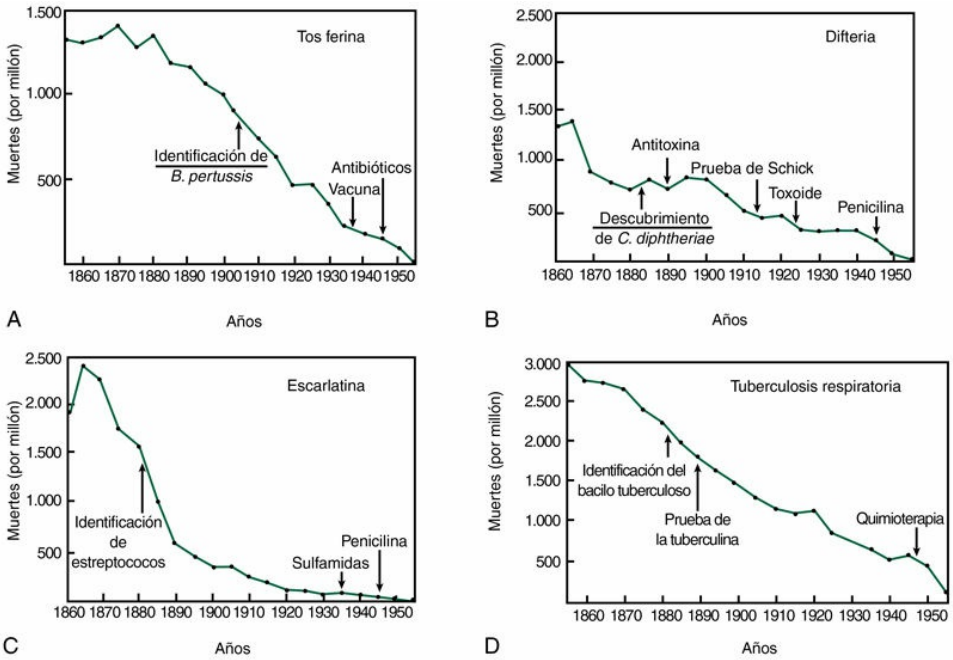


FIG. 1.19 Disminución de la tasa de mortalidad en Inglaterra y Gales por (A) tos ferina, (B) difteria, (C) escarlatina (en niños menores de 15 años) y (D) tuberculosis respiratoria. (De Kass EH. Infectious diseases and social change. *J Infect Dis.* 1971;123:110-114.)

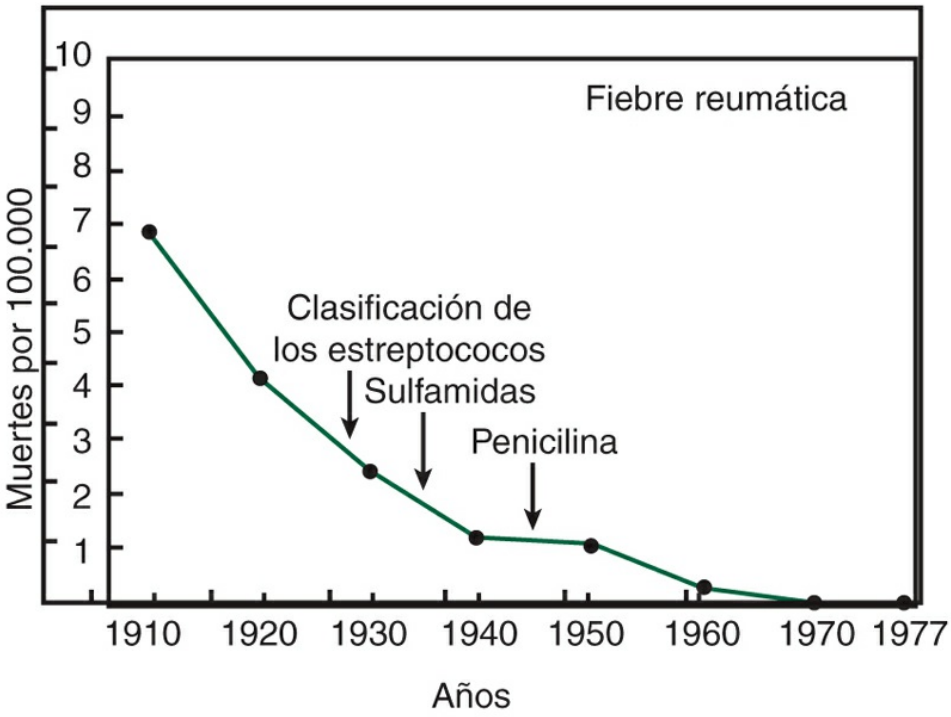


FIG. 1.20 Disminución de la tasa de mortalidad por fiebre reumática, Estados Unidos, 1910-1977. (De Gordis L. The virtual disappearance of rheumatic fever in the United States: lessons in the rise and fall of disease. T. Duckett Jones Memorial Lecture. *Circulation.* 1985;72:1155-1162.)

¿Cuál es la explicación de estos descensos espectaculares incluso antes de que se contase con vacunas o tratamientos? En teoría, es posible que cuando observamos un descenso en la mortalidad por una enfermedad infecciosa, la exposición humana al microorganismo causal puede haber disminuido o la virulencia del microorganismo puede ser menor. Sin embargo, una explicación más probable del descenso de la mortalidad en estos y otros ejemplos es que se debieron principalmente a mejoras en las condiciones sociales, a la mayor seguridad en el consumo de agua y a la reducción de la exposición a contaminantes, de manera que no se relacionaron con ninguna intervención médica. De hecho, Kass tituló su artículo de 1971, en el que aparecían los gráficos de la figura 1.19, «Enfermedades infecciosas y cambios sociales». Aunque los factores específicos que probablemente están implicados no siempre resultan claros, las mejoras en las viviendas, como las instalaciones sanitarias y la mejor nutrición, junto con cambios simultáneos del estilo de vida, son factores importantes que probablemente hayan contribuido de manera importante a dicho descenso.

A menudo atribuimos con entusiasmo los descensos temporales de mortalidad a intervenciones médicas. Sin embargo, la lección ilustrada por los ejemplos de estos gráficos es que debemos ser prudentes antes de concluir que el descenso de mortalidad observado es el resultado de una intervención médica. En vista de la dificultad para extraer conclusiones acerca de la eficacia de intervenciones médicas únicamente en descensos importantes de mortalidad en la población, resultan claramente esenciales estudios epidemiológicos rigurosos para valorar la eficacia de diferentes

intervenciones médicas. En el capítulo 17 se exponen algunos de los abordajes empleados y el diseño de dichos estudios para evaluar los servicios de salud.

Integración de la prevención y el tratamiento

La prevención y el tratamiento son consideradas con demasiada frecuencia actividades mutuamente excluyentes, como se expone en la figura 1.21. Sin embargo, resulta claro que la prevención no solo es fundamental para la salud pública, sino que también es una parte integral de la práctica clínica. El papel del médico es mantener la salud, así como tratar la enfermedad. No obstante, incluso el tratamiento de la enfermedad incluye un componente importante de prevención. Cuando tratamos una enfermedad, estamos previniendo la muerte, las complicaciones en el paciente y el impacto en la familia del paciente. Por tanto, gran parte de la dicotomía entre tratamiento y prevención es una ilusión. El tratamiento implica la prevención secundaria y la terciaria, y esta última consiste en la prevención de complicaciones, como la discapacidad. En ocasiones también implica la prevención primaria. Por tanto, todo el espectro de la prevención debe considerarse esencial para la salud pública y la práctica clínica.



FIG. 1.21 «El doctor es especialista en medicina preventiva…Así que si usted está enfermo, ya no le interesa.» La prevención y el tratamiento vistos como actividades mutuamente excluyentes. (ZIGGY© 1986 ZIGGY AND FRIENDS, INC. Reimpreso con autorización de ANDREWS MCMEEL SYNDICATION. Todos los derechos reservados.)

Dos decisiones muy diferentes adoptadas en 2012 hacen hincapié aún más en la relación entre prevención y tratamiento. En julio de 2012, la Food and Drug Administration (FDA) estadounidense aprobó el uso de un fármaco, Truvada (la combinación de los antivirales tenofovir y emtricitabina, Gilead Sciences, Foster City, CA, Estados Unidos), para la prevención de la infección por VIH en personas con alto riesgo de infectarse por dicho virus (la denominada profilaxis preexposición [PPrE]). Desde 2004, el fármaco se había comercializado únicamente como tratamiento de pacientes ya infectados por VIH, tanto de los que presentaban infección crónica como de los expuestos a pinchazos con agujas u otros riesgos de tipo traumático (la denominada profilaxis postexposición [PPE]).

La segunda decisión, anunciada en mayo de 2012, fue que los National Institutes of Health iban a realizar un ensayo clínico de 5 años de duración con el fin de prevenir una forma de enfermedad de Alzheimer genéticamente determinada. Los investigadores estudiarán a 300 personas con una capacidad cognitiva normal, pero con un riesgo elevado de desarrollar la

enfermedad de Alzheimer. El estudio se inició en 2013 y se espera su finalización en 2020. La mayoría de los participantes en el estudio pertenecen a una familia numerosa de Medellín, Colombia, que presenta un riesgo alto de padecer una forma de enfermedad de Alzheimer genéticamente determinada, caracterizada por la aparición temprana de alteraciones cognitivas seguida de demencia florida a la edad de 53 años. El fármaco que está siendo estudiado, crenezumab (anticuerpos frente a dos tipos de amiloide beta humano; Genentech, South San Francisco, CA, Estados Unidos), está siendo evaluado en la actualidad en otros dos ensayos clínicos en pacientes que ya presentan una demencia de leve a moderada para determinar si puede enlentecerse la formación de acumulación de amiloide o el deterioro cognitivo. Así, tanto en el estudio sobre el VIH expuesto en el párrafo anterior como en este estudio sobre la enfermedad de Alzheimer, fármacos que han sido utilizados en pacientes con diagnósticos claros de las enfermedades en cuestión están siendo evaluados en la actualidad como fármacos que podrían prevenir estas enfermedades en pacientes de alto riesgo. Ambos estudios hacen hincapié en la necesidad de aunar también el tratamiento y la prevención en nuestra consideración cambiante de otras enfermedades.

Conclusión

La epidemiología es una herramienta valiosa que proporciona una base racional sobre la que pueden planificarse e implementarse programas de prevención eficaces. La epidemiología también es valiosa para llevar a cabo investigaciones clínicas para evaluar nuevos tratamientos y terapias utilizadas durante algún tiempo, así como intervenciones novedosas desarrolladas para la prevención de enfermedades. El objetivo último es mejorar el control de la enfermedad por medio de la prevención y el tratamiento para evitar muertes debidas a la enfermedad y mejorar la calidad de vida de los que padecen una enfermedad grave. Los diseños de estudios utilizados en epidemiología se analizan en capítulos posteriores.

Bibliografía

1. Porta M. A Dictionary of Epidemiology. 5th ed. New York: Oxford University Press; 2008.
2. Rose G. Sick individuals and sick populations. *Int J Epidemiol*. 1985;14:32–38.
3. Nuland SB. The Doctors' Plague: Germs, Childbed Fever and the Strange Story of Ignáz Semmelweis. New York: WW Norton/Atlas Books; 2003.
4. Fenner F, Henderson DA, Arita I, et al. Smallpox and Its Eradication. Geneva, Switzerland: World Health Organization; 1988.
5. Johnson S. The Ghost Map: The Story of London's Most Terrifying Epidemic—and How It Changed Science, Cities, and the Modern World. New York: Riverhead Books; 2006.
6. Markowitz M, Gordis L. Rheumatic Fever. 2nd ed. Philadelphia: WB Saunders; 1972.
7. Kass EH. Infectious diseases and social change. *J Infect Dis*. 1971;123:110–114.
8. Gordis L. The virtual disappearance of rheumatic fever in the United States: lessons in the rise and fall of disease. *Circulation*. 1985;72:1155–1162.

CAPÍTULO 2

Dinámica de la transmisión de las enfermedades

Palabras clave

tríada epidemiológica
transmisión directa e indirecta
infección subclínica
endémica
epidémica y pandémica
susceptibilidad
inmunidad grupal
tasa de ataque
periodo de incubación

Conservo seis sirvientes honestos

(Ellos me enseñaron todo lo que sé);

Sus nombres son qué, por qué, cuándo, cómo, dónde y quién.
—Rudyard Kipling[1] (1865-1936)

Objetivos de aprendizaje

- Introducir conceptos relacionados con la transmisión de la enfermedad empleando como modelo el enfoque epidemiológico de las enfermedades contagiosas.
- Definir términos importantes relacionados con la aparición de enfermedades en la población.
- Calcular la tasa de ataque e ilustrar cómo puede emplearse para medir la transmisión de una enfermedad de una persona a otra.

- Describir los pasos en la investigación de un brote y presentar cómo puede emplearse la tabulación cruzada para identificar la fuente.

Las enfermedades humanas no se originan de la nada. Se deben a la interacción entre el hospedador (una persona), el agente (p. ej., una bacteria) y el ambiente (p. ej., aire contaminado). Aunque algunas enfermedades son de origen principalmente genético, prácticamente todas las enfermedades se deben a la interacción de factores genéticos, conductuales y ambientales, de manera que las proporciones difieren en las diferentes enfermedades. Muchos de los principios que subyacen en la transmisión de las enfermedades se demuestran más claramente utilizando como modelo las enfermedades infecciosas. Por tanto, este capítulo utiliza como ejemplo principalmente a dichas enfermedades para revisar estos principios. No obstante, los conceptos analizados también son aplicables a las enfermedades que no son de origen infeccioso (p. ej., el cáncer de los fumadores pasivos).

La enfermedad se ha descrito clásicamente como el resultado de la tríada epidemiológica que se muestra en la figura 2.1. Según este diagrama, es el producto de la interacción entre el hospedador humano, el agente infeccioso o de otro tipo y el ambiente que favorece la exposición. En ocasiones participa un vector, como el mosquito o la garrapata del ciervo. Para que dicha interacción tenga lugar, el hospedador debe ser susceptible. La susceptibilidad humana está determinada por diversos factores como la carga genética y características propias de nuestro comportamiento, nutricionales e inmunológicas. El estado inmunitario de un individuo está determinado por numerosos factores como el contacto previo bien con la infección natural o con la vacunación.

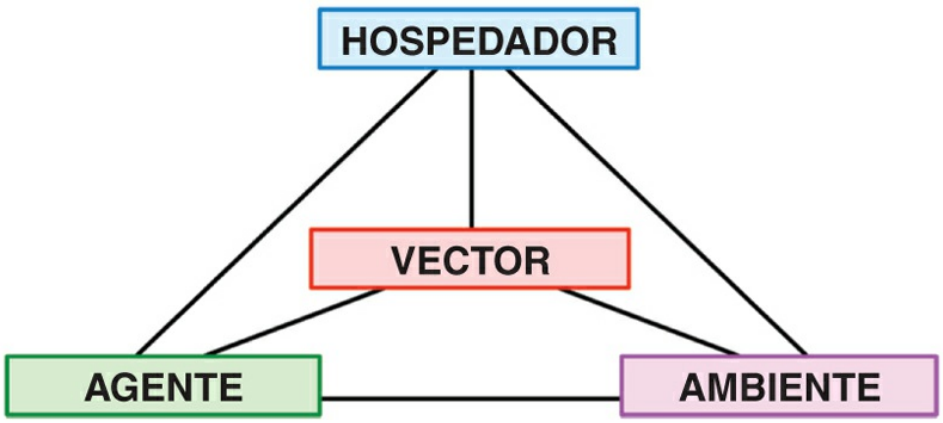


FIG. 2.1 La tríada epidemiológica de una enfermedad.

Los factores que pueden causar enfermedad humana son factores biológicos, físicos y químicos, así como otros tipos, como el estrés o riesgos de nuestra conducta, que pueden ser más difíciles de clasificar (tabla 2.1).

Tabla 2.1

Factores que pueden asociarse a un mayor riesgo de enfermedad en el ser humano

Características del hospedador	Tipos de agentes y ejemplos	Factores ambientales
Edad	Biológicos	Temperatura
Sexo	Bacterias, virus	Humedad
Raza	Químicos	Altitud
Religión	Metales pesados, alcohol, humo	Hacinamiento
Costumbres	Físicos	Hogar
Ocupación laboral	Traumatismo, radiación, fuego	Vecindario
Perfil genético	Nutricionales	Agua
Estado civil	Carencia, exceso	Leche
Antecedentes familiares		Alimentos
Enfermedades previas		Radiación
Estado inmunitario		Contaminación del aire
		Ruido

Modos de transmisión

Las enfermedades pueden transmitirse *directamente* o *indirectamente*. Por ejemplo, una enfermedad puede transmitirse entre personas (transmisión directa) a través del contacto directo (como en el caso de las infecciones de transmisión sexual). La transmisión indirecta puede producirse a través de un vehículo común, como el aire o los suministros de agua contaminados, o por un vector como el mosquito. En el cuadro 2.1 se exponen algunos de los modos de transmisión.

Cuadro 2.1 Modos de transmisión de la enfermedad

1. Directa:
    a. Contacto de persona a persona.
2. Indirecta:
    a. Vehículo común:
        1) Exposición única.
        2) Exposiciones múltiples.
        3) Exposición continua.
    b. Vector.

La figura 2.2 es una fotografía clásica en la que se observa la dispersión de gotículas después de un estornudo. Demuestra claramente el potencial que tiene un paciente de infectar a un elevado número de personas en un breve periodo de tiempo. Como Mims ha señalado:

Un paciente infectado puede transmitir la gripe o el catarro común a muchos otros en el transcurso de una única hora en una habitación abarrotada. Una infección venérea también debe propagarse progresivamente de una persona a otra para persistir en la naturaleza, pero sería una tarea formidable transmitir una infección venérea a dicha escala[2].

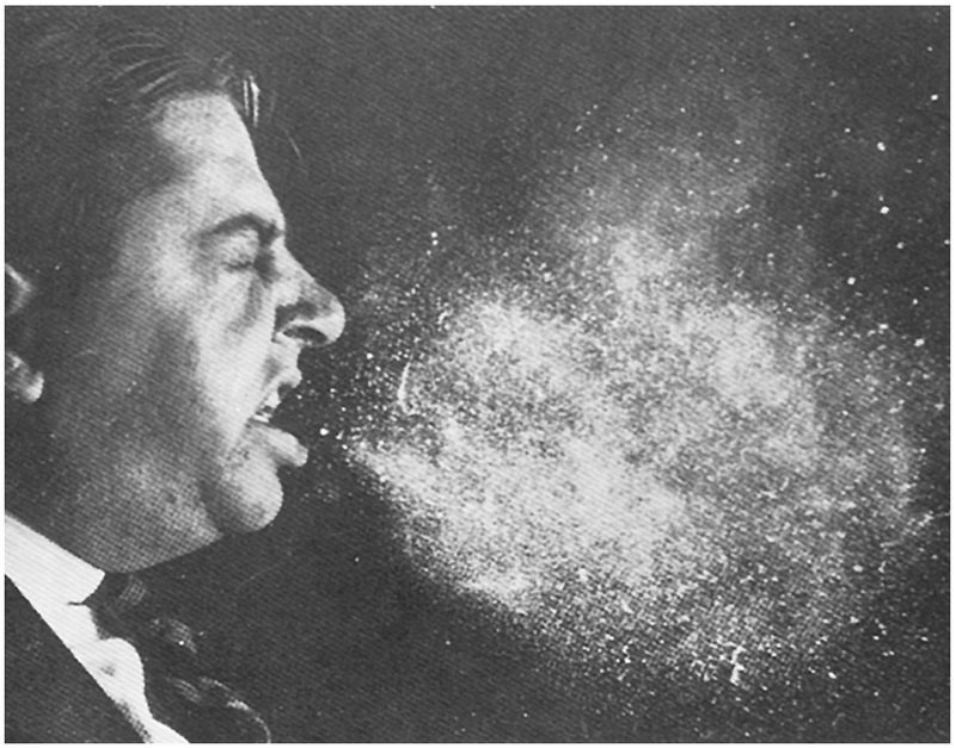


FIG. 2.2 Dispersión de gotículas tras un estornudo violento. (Reproducido con autorización de Jennison MW. *Aerobiology* 17:102, 1947. Copyright 1947 American Association for the Advancement of Science.)

Así, los diferentes microorganismos se propagan de diferentes maneras, y el potencial de propagación y de causar brotes de un microorganismo concreto depende de las características del microorganismo, como su tasa de crecimiento, la vía por la que es transmitido de una persona a otra y la cantidad de individuos susceptibles en la comunidad.

La figura 2.3 es un diagrama esquemático de las superficies del cuerpo humano como lugares de infección y diseminación microbiana. El tracto alimentario puede considerarse como un tubo abierto que atraviesa el cuerpo y los aparatos respiratorio y urogenital se muestran como invaginaciones ciegas. Cada uno ofrece oportunidades para la infección. La piel es otra puerta de entrada importante para los agentes infecciosos, principalmente a través de arañazos, picaduras o lesiones. Entre los microorganismos que a menudo penetran a través de la piel se encuentran los estreptococos, los estafilococos y los hongos, como la tiña (circinada). A este respecto hay que tener en cuenta dos puntos: en primer lugar, la piel no es la puerta de entrada exclusiva de muchos de estos microorganismos y, en segundo lugar, las infecciones pueden adquirirse a través de más de una vía. Las mismas vías también sirven de puntos de entrada de agentes no infecciosos causantes de enfermedad. Por ejemplo, las toxinas ambientales pueden ser ingeridas, inspiradas con la respiración o absorberse directamente a través de la piel. Las características clínicas y epidemiológicas de muchas enfermedades infecciosas y no infecciosas a menudo se relacionan con el punto de exposición al microorganismo o a la sustancia ambiental y con la puerta de entrada en el organismo.

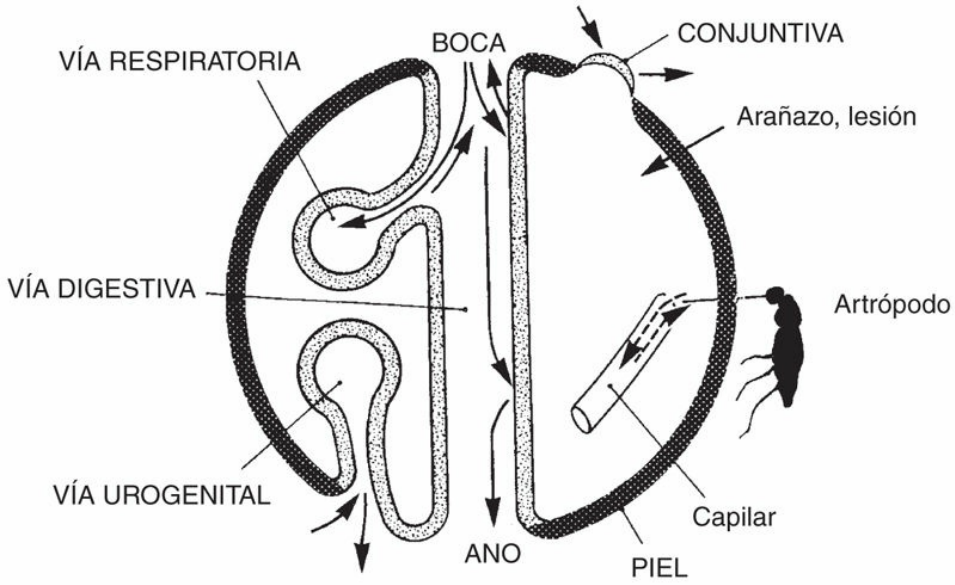


FIG. 2.3 Superficies corporales como sitios de infección y diseminación microbiana. (De Mims CA, Nash A, Stephen J. *Mims' Pathogenesis of Infectious Disease,* 5.ª ed. Londres: Academic Press; 2001.)

Enfermedad clínica y subclínica

Es importante tener en cuenta el amplio espectro de gravedad de la enfermedad. En la figura 2.4 se muestra el concepto «iceberg» de la enfermedad. Al igual que la mayor parte de un «iceberg» se encuentra por debajo del agua y oculto a la vista excepto por su punta visible, lo mismo ocurre con la enfermedad: tan solo la enfermedad clínica es fácilmente visible (como se expone en la *Respuesta del hospedador,* parte derecha de la fig. 2.4). Sin embargo, las infecciones sin enfermedad clínica aparente son importantes, en especial en lo referente a la transmisión de la enfermedad, aunque no sean clínicamente visibles. En la parte izquierda de la figura 2.4 se observan las etapas biológicas correspondientes de la patogenia (mecanismos biológicos) y la enfermedad a nivel celular. El concepto «iceberg» es importante porque no es suficiente tener en cuenta únicamente los casos clínicamente visibles que observamos; por ejemplo, la mayoría de los casos de poliomielitis en la época anterior a la vacunación eran subclínicos, es decir, muchas personas infectadas por la poliomielitis no presentaban enfermedad aparente. No obstante, eran capaces de propagar el virus a otros. Como resultado, no podemos comprender y explicar la diseminación de la poliomielitis a no ser que reconozcamos el grupo de casos inaparentes. Desde el punto de vista de la enfermedad inaparente, esta situación no es diferente en muchas enfermedades no transmisibles. Aunque estas enfermedades no se transmiten de persona a persona, muchos pacientes, por ejemplo, pueden vivir mucho tiempo con una enfermedad renal crónica inaparente y solo cuando experimentan una complicación clínica se hace el diagnóstico de la enfermedad renal crónica.

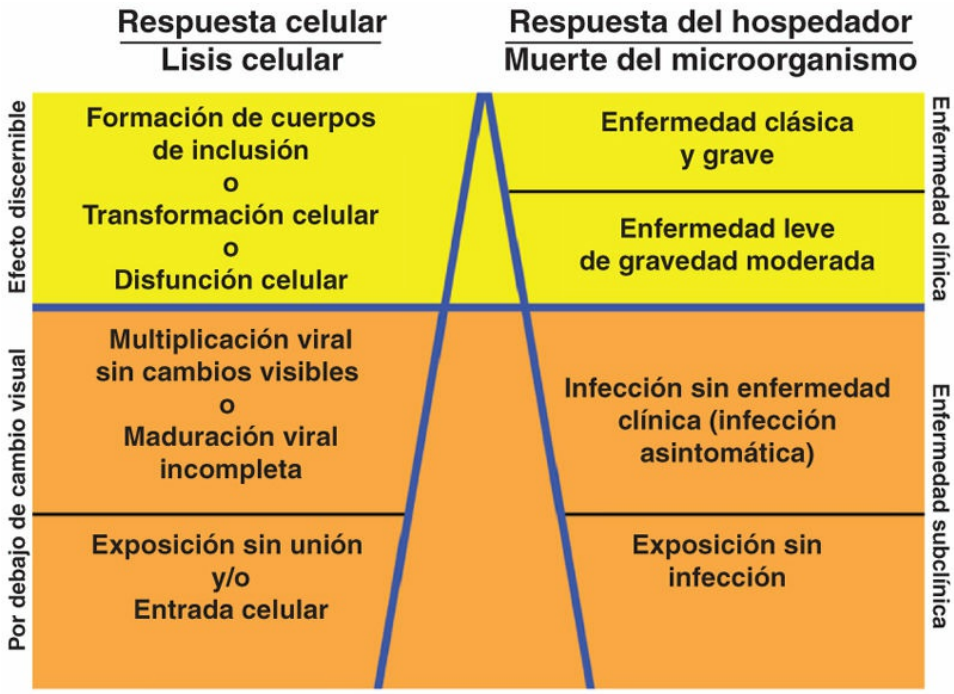


FIG. 2.4 El concepto «iceberg» de las enfermedades infecciosas en la célula y en el hospedador. (Modificado de Evans AS, Kaslow RA, eds. *Viral Infections of Humans: Epidemiology and Control.* 4.ª ed. Nueva York: Plenum; 1997.)

En la figura 2.5 se muestra el espectro de gravedad de varias enfermedades. La mayoría de los casos de tuberculosis, por ejemplo, son inaparentes. Sin embargo, como los casos inaparentes pueden transmitir la enfermedad, dichos casos deben ser identificados y tratados para controlar la propagación adicional de la enfermedad. Respecto al sarampión, muchos casos son de gravedad moderada y solo unos pocos son inaparentes. En el otro extremo, sin intervención, la rabia no cursa con casos inaparentes y la mayoría de los casos no tratados son mortales. Por tanto, observamos un espectro de patrones de gravedad que varía con la enfermedad. La gravedad parece estar relacionada con la virulencia del microorganismo (de lo eficaz que sea el microorganismo produciendo la enfermedad) y con la localización corporal en la que el microorganismo se multiplica. Todos estos factores, así como características del hospedador como la respuesta inmunitaria, deben ser tenidos en cuenta para comprender cómo se propaga la enfermedad de un individuo a otro.

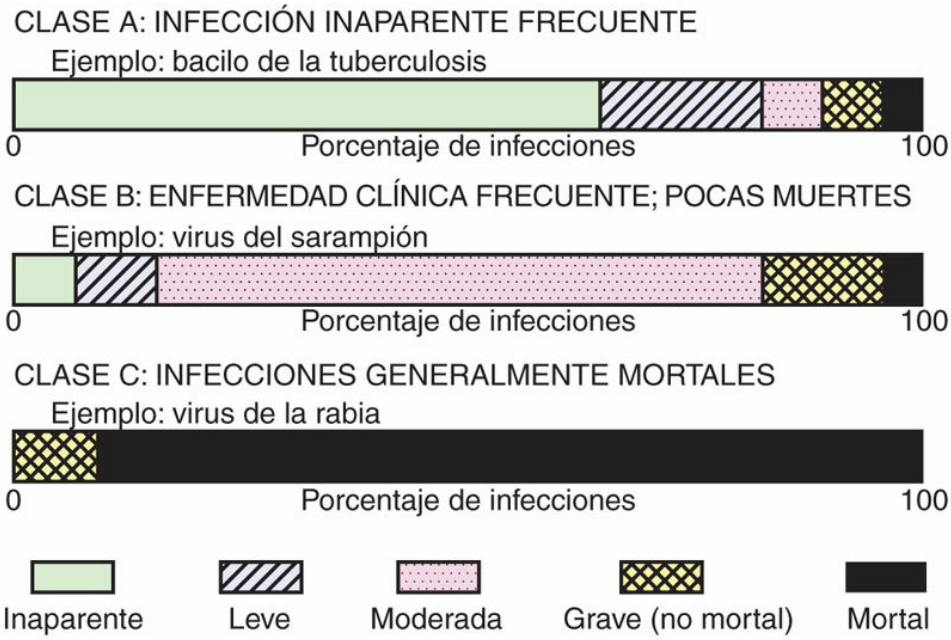


FIG. 2.5 Distribución de la gravedad clínica de tres clases de infecciones (no dibujadas a escala). (Modificado de Mausner JS, Kramer S. *Epidemiology: An Introductory Text.* Filadelfia: WB Saunders; 1985: 265.)

Al igual que con el paso de los años ha aumentado el conocimiento de los aspectos clínicos y biológicos, del mismo modo lo ha hecho nuestra capacidad para distinguir las diferentes etapas de la enfermedad. Estas consisten en las enfermedades clínicas y no clínicas.

Enfermedad clínica

La enfermedad clínica se caracteriza por signos y síntomas.

Enfermedad no clínica (inaparente)

La enfermedad no clínica puede incluir las siguientes:

1. *Enfermedad preclínica.* Enfermedad que todavía no es clínicamente aparente pero que está destinada a progresar a enfermedad clínica.
2. *Enfermedad subclínica.* Enfermedad asintomática sin signos y síntomas visibles y que no está destinada a volverse clínicamente aparente. Este tipo de enfermedad a menudo se diagnostica por su respuesta serológica (anticuerpos) o por el cultivo del microorganismo.
3. *Enfermedad persistente (crónica).* Enfermedad de larga duración y de progresión lenta que puede persistir durante años, a veces de por vida. En los últimos años se ha observado un fenómeno interesante: la manifestación de síntomas muchos años después de que una infección se hubiese considerado resuelta. Algunos adultos que se recuperaron de una poliomielitis en la infancia presentan cansancio y debilidad crónica grave; esto se ha denominado síndrome

pospoliomielitis en la vida adulta. Estos se han convertido así en casos de enfermedad clínica, aunque algo diferentes de la enfermedad inicial.

4. *Enfermedad latente*. Una infección sin multiplicación activa del microorganismo, como cuando el ácido nucleico viral es incorporado al núcleo de una célula como un provirus. A diferencia de la infección persistente, solo el mensaje genético está presente en el hospedador, no el microorganismo viable.

Estado de portador

Un portador es un individuo que alberga el microorganismo pero que no está infectado según el resultado de estudios serológicos (no existe respuesta de anticuerpos) ni existen signos o síntomas de enfermedad clínica. Esta persona puede infectar a otras, aunque la infectividad por lo general es menor que con otras infecciones. El estado de portador puede ser de duración limitada o puede ser crónico; puede durar meses o años. Uno de los ejemplos más conocidos de portador crónico fue Mary Mallon, más conocida como María Tifoidea, portadora de *Salmonella typhi,* que murió en 1938. Durante un periodo de muchos años trabajó como cocinera en el área urbana de Nueva York; se trasladaba de una casa a otra bajo nombres diferentes. Se considera que fue la causante de al menos 10 brotes de fiebre tifoidea que produjeron 51 casos y 3 muertes.

Endémico, epidémico y pandémico

Se deben definir otros tres términos: *endémico, epidémico* y *pandémico. Endémico* se define como la presencia habitual de una enfermedad en una zona geográfica determinada. También puede hacer referencia a la aparición habitual de una enfermedad determinada en dicha zona (a veces referido como la «tasa de fondo de la enfermedad»). *Epidémico* se define como la aparición en una región o en una comunidad de un grupo de enfermedades de naturaleza similar, con una frecuencia claramente superior a la normal y originada a partir de una fuente común o propagada (fig. 2.6). *Pandémico* hace referencia a una epidemia mundial.

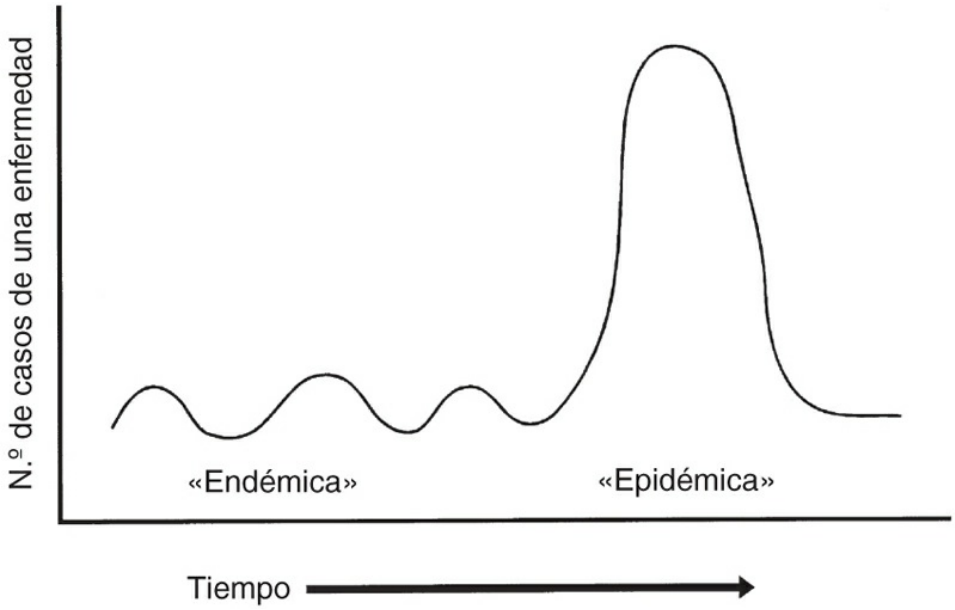


FIG. 2.6 Enfermedad endémica frente a epidémica.

¿Cómo sabemos que la frecuencia es superior a la esperada? De hecho, ¿cómo sabemos cuánto es lo esperado? No existe una respuesta precisa a ninguna de las dos preguntas. A través de la vigilancia continua podemos determinar cuál es la frecuencia habitual o esperada. Con respecto al exceso, en ocasiones una «prueba interocular» puede ser convincente: la diferencia es tan clara que salta a la vista.

Dos ejemplos nos demuestran cómo las pandemias y el miedo a las pandemias se relacionan con la creación de políticas públicas. Los pacientes con enfermedad renal crónica a menudo sufren anemia, que comúnmente se corrige mediante la inyección de agentes estimulantes de la eritropoyesis (AEE); estos son formas genéticamente modificadas de la hormona eritropoyetina humana. Los fabricantes de medicamentos pagan a los médicos millones de dólares cada año a cambio de recetar este medicamento para la anemia, lo que dio lugar a un uso frecuente para indicaciones no

autorizadas y a la sobreutilización de los AEE en Estados Unidos (fig. 2.7). En 2006, se publicaron dos ensayos clínicos en el *New England Journal of Medicine* que plantearon dudas sobre la seguridad del uso de los AEE para la corrección de la anemia a niveles óptimos en pacientes con enfermedad renal crónica, ya que ninguno de los estudios anticipó estos resultados. El primer ensayo clínico, Reducción del Riesgo Cardiovascular mediante el Tratamiento Temprano de la Anemia con Epoetina Beta (CREATE, *Cardiovascular Risk Reduction by Early Anemia Treatment with Epoetin Beta*)[3], demostró que la corrección temprana y completa de la anemia (hasta una concentración de hemoglobina objetivo en el rango normal) no logró reducir la incidencia de complicaciones cardiovasculares en comparación con la corrección parcial de esta. El segundo ensayo, Corrección de la Hemoglobina y Resultados en la Insuficiencia Renal (CHOIR, *Correction of Hemoglobin and Outcomes in Renal Insufficiency*)[4], mostró que un valor de hemoglobina objetivo más elevado de 13,5 g/dl se asoció con un mayor riesgo de muerte, infarto de miocardio, hospitalización por insuficiencia cardiaca congestiva y accidente cerebrovascular, todo sin una mejora en la calidad de vida en comparación con una cifra objetivo de hemoglobina inferior de 11,3 g/dl (fig. 2.8). Como resultado, en 2007, la agencia estadounidense para la administración de medicamentos y alimentos emitió una advertencia de recuadro negro que incluía restricciones significativas al uso de AEE. La advertencia de recuadro negro incluye lo siguiente: 1) los médicos deben usar la dosis más baja de AEE que aumente gradualmente la concentración de hemoglobina al nivel más bajo suficiente para evitar la necesidad de transfusión de hematíes, y 2) los AEE aumentan el riesgo de muerte y de complicaciones cardiovasculares graves cuando se administran para lograr una hemoglobina mayor de 12 g/dl.

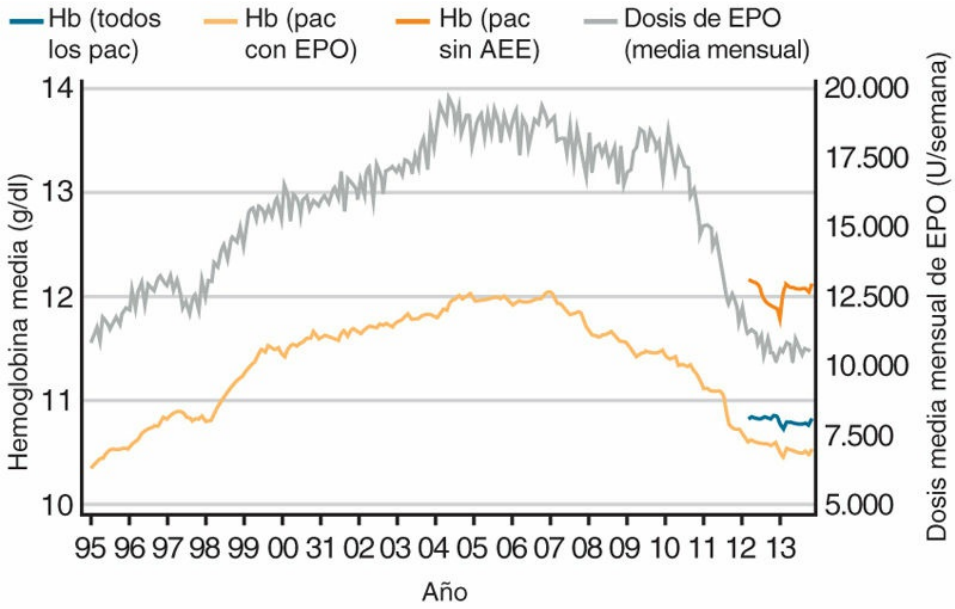


FIG. 2.7 Concentración media mensual de hemoglobina *(Hb)* y dosis media mensual de epoetina *(EPO)* (expresada como unidades por semana) en pacientes adultos hemodializados con diálisis ≥90 días, solicitudes de reembolso de

Medicare, 1995–2013. (De US Renal Data System. 2015 USRDS annual data report volume 2: ESRD in the United States. https://www.usrds.org/2015/download/vol2_USRDS_ESRD_15.pdf. Consultado el 5 de junio de 2017.)

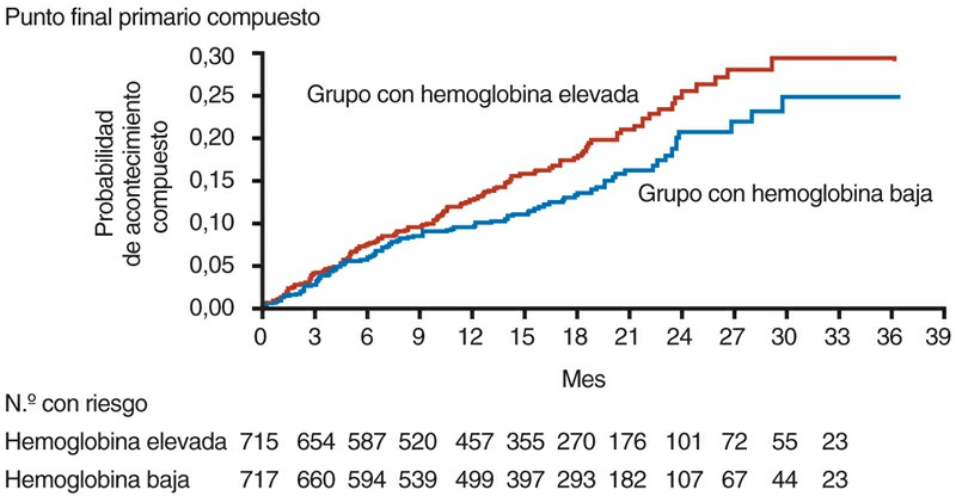


FIG. 2.8 Estimaciones de Kaplan-Meier de la probabilidad del punto final primario compuesto (hospitalización por insuficiencia cardiaca congestiva sin terapia de reemplazo renal, infarto de miocardio, accidente cerebrovascular y muerte). (De Singh AK, Szczech L, Tang KL, et al. Correction of anemia with epoetin alfa in chronic kidney disease. *N Engl J Med.* 2006;355:2085–2098.)

El segundo ejemplo aborda un asunto que surgió en 2011 relacionado con la investigación de laboratorio del virus H5N1 o de la «gripe aviar» (fig. 2.9). Aunque la transmisión del virus H5N1 de aparición natural se ha limitado principalmente a sujetos en contacto directo con animales infectados, en los casos raros en los que las personas adquieren la infección a partir de animales, la enfermedad a menudo es muy grave y con una elevada mortalidad. Por tanto, existe una preocupación importante acerca de que ciertas mutaciones del virus puedan aumentar la capacidad de transmisión del virus a humanos, lo que podría causar una pandemia humana. Con el fin de conocer por completo la posibilidad de dicha mutación y la posibilidad de evitarla, dos laboratorios financiados por el gobierno, uno en el Erasmus Medical Center de Holanda y otro en la Universidad de Wisconsin-Madison de Estados Unidos, crearon cepas H5N1 genéticamente alteradas que podían transmitirse entre mamíferos (hurones) a través del aire.

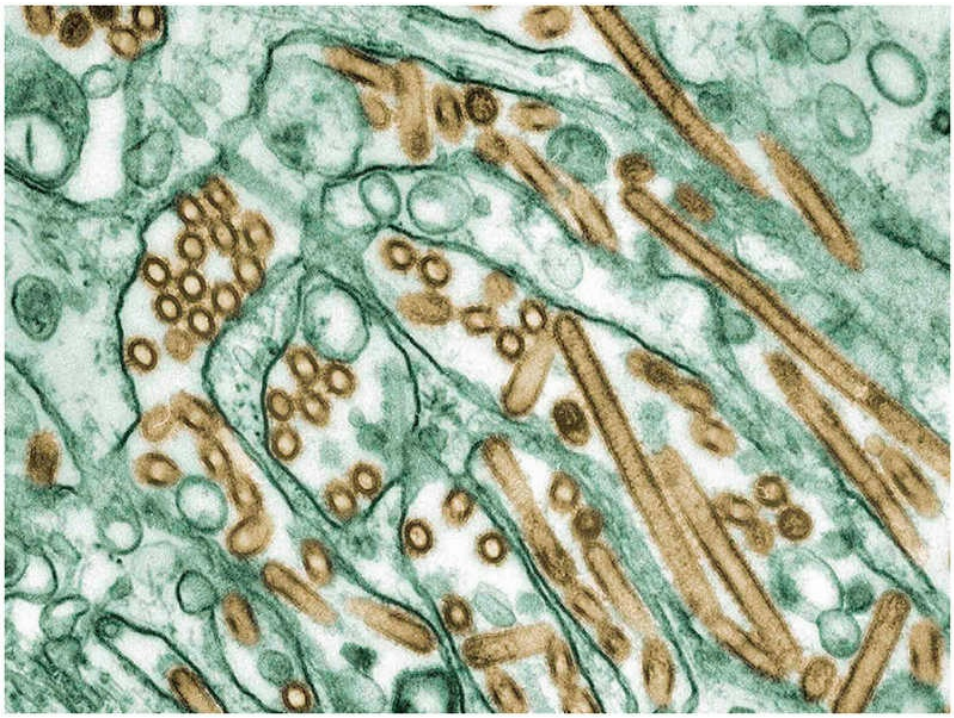


FIG. 2.9 Imagen coloreada de microscopia electrónica de transmisión del virus de la gripe aviar A H5N1 (en *dorado*) cultivado en células MDCK (en *verde*). (De Centers for Disease Control and Prevention, por cortesía de Cynthia Goldsmith, Jacqueline Katz y Sherif R. Zaki.)

Tras analizar los dos estudios, por primera vez en su historia, el National Science Advisory Board for Biosecurity de Estados Unidos recomendó no publicar los detalles de las metodologías empleadas en estos estudios. El comité citó la posibilidad de un mal uso por parte de «los que buscan hacer daño» por medio de actividades de terrorismo biológico. Sin embargo, otros científicos, incluyendo miembros de un comité de expertos reunidos por la Organización Mundial de la Salud (OMS), discreparon, afirmando que el trabajo era importante para los esfuerzos de salud pública para evitar una posible pandemia en humanos. En enero de 2012, los investigadores se autoimpusieron una moratoria en algunos tipos de investigación sobre el virus H5N1 para dejar tiempo para que estos asuntos fueran debatidos por expertos y por el público. Los resultados de los dos estudios fueron finalmente publicados en mayo y en junio de 2012[4,6,7].

El principal aspecto no resuelto fue si los potenciales beneficios para la sociedad de estos tipos de estudios superan los riesgos de la propagación incontrolada de virus mutados, debido o a fallos en la bioseguridad de los laboratorios (escape accidental del virus) o a las actividades del terrorismo biológico (liberación intencionada del virus). Los científicos y los legisladores están obligados a desarrollar métodos para valorar los riesgos y los beneficios de realizar diferentes tipos de investigaciones experimentales. Además, estos hechos ilustran cómo la censura y la libertad académica en la ciencia siguen siendo temas de gran importancia en la actualidad.

Brotes de enfermedad

Supongamos que un alimento se contamina por un microorganismo. Si se produce un brote en el grupo de personas que han consumido dicho alimento nos encontramos ante una *exposición a un vehículo común,* porque todos los casos que se han producido han sido en personas expuestas al alimento presuntamente contaminado. El alimento puede servirse solamente una vez, por ejemplo, en un almuerzo servido por una empresa de servicio de banquetes a domicilio; esto es una *exposición única* de las personas que lo consumieron. O el alimento puede servirse más de una vez, y por tanto son *exposiciones múltiples* de las personas que lo consumieron más de una vez. Cuando un suministro de agua es contaminado con aguas residuales debido a fugas en las tuberías, la contaminación puede ser *periódica* (provoca exposiciones múltiples como resultado de cambios de presión en el sistema de suministro de agua, lo que puede causar una contaminación intermitente) o *continua* (la fuga constante causa una contaminación persistente). El cuadro epidemiológico resultante depende de si la exposición es única, múltiple o continua.

En esta exposición nos centraremos en el *brote con vehículo común y exposición única,* ya que los temas que vamos a exponer se observan con mayor claridad en este tipo de brote. ¿Cuáles son las características de este tipo de brote? En primer lugar, estos brotes suelen ser explosivos, es decir, existe un aumento repentino y rápido del número de casos de una enfermedad en una población. (Curiosamente, las epidemias de exposición única a vehículos comunes de enfermedades no transmisibles, como la epidemia de leucemia tras la explosión de una bomba atómica en Hiroshima y Nagasaki, también parecen seguir el mismo patrón.) En segundo lugar, los casos se limitan a las personas que compartieron la exposición común. Esto resulta evidente porque en la primera oleada de casos no esperaríamos que la enfermedad se presentase en personas no expuestas a no ser que existiera otra fuente independiente de la enfermedad en la comunidad. En tercer lugar, en un brote de origen alimentario, los casos raramente se presentan en personas que no consumieron el alimento, es decir, en individuos que adquieren la enfermedad a partir de un caso primario que ingirió el alimento. El motivo de la rareza relativa de dichos casos secundarios en este tipo de brote no es bien conocido.

En Estados Unidos, la causa principal de enfermedades relacionadas con la comida es la contaminación por norovirus (de la familia de virus Norwalk). A nivel mundial, las infecciones por norovirus suponen un total de 4,2 miles de millones de dólares en costos directos al sistema sanitario y 60,3 miles de millones de dólares en costos sociales por año[8].

En las últimas décadas se ha producido un número creciente de brotes de gastroenteritis aguda (GEA) a bordo de barcos de crucero. Los Centros para el Control y la Prevención de Enfermedades (CDC, Centers for Disease

Control and Prevention) informaron que las tasas de GEA entre los pasajeros de cruceros han disminuido de 27,2 casos por 100.000 días de viaje en 2008 a 22,3 en 2014, mientras que la tasa entre los miembros de la tripulación se mantuvo prácticamente sin cambios[9]. Esto podría atribuirse potencialmente a la creación de manuales operativos o pautas específicas que brindan estándares de higiene, lo que aumenta la conciencia entre los pasajeros y los miembros de la tripulación a los programas de vigilancia de enfermedades transmisibles y a las intervenciones preventivas, además de la aplicación de regulaciones e inspecciones estrictas a través del Programa de Saneamiento de Buques (VSP, *Vessel Sanitation Program*) de los CDC, que controla los brotes en los cruceros y sirve para prevenir y controlar la transmisión de enfermedades a bordo de estos barcos. (Los datos de cada brote están disponibles en la página web, http://www.cdc.gov/nceh/vsp/). En áreas con una alta prevalencia de norovirus, particularmente el tipo GII.2 recombinante, como en las provincias de Guangdong y Jiangsu en China, los brotes de GEA continúan ocurriendo con frecuencia[10]. Por ejemplo, el 14 de diciembre de 2014, un estudiante de tercer grado vomitó en el aula y el baño varias veces y fue considerado el primer caso sospechoso de norovirus. Durante los siguientes 3 días se notificaron 27 casos más, que se ubicaron principalmente en la cuarta planta (12 casos) y en la tercera planta (9 casos) del edificio de la escuela. La figura 2.10 muestra la curva epidémica con el número de casos cada día. El primer pico del brote tuvo lugar el 17 de diciembre y comenzó a disminuir cuando se pusieron en práctica medidas de control como la cuarentena y la desinfección. Sin embargo, unos días después, el 25 de diciembre, la tasa de ataque repuntó nuevamente, con casos que se presentaron principalmente en la segunda planta (12 casos) y en la tercera (5 casos). Para contener el brote enérgicamente, la escuela se cerró con carácter temporal y el brote finalizó el 31 de diciembre.

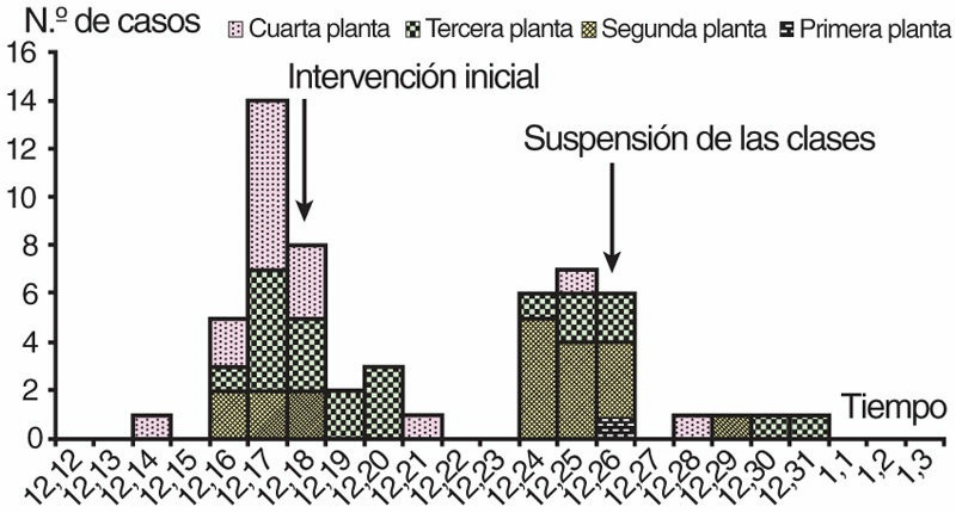


FIG. 2.10 Distribución de casos agrupados por planta y fecha de inicio del brote en una escuela en la provincia de Jiangsu, China, en 2014. (De Shi C, Feng W-H, Shi P, et al. An acute gastroenteritis outbreak caused by GII.17 norovirus in

Jiangsu Province, China. *Int J Infect Dis.* 2016;49:30–32.)

Inmunidad y susceptibilidad

El grado de enfermedad en una población depende del equilibrio entre el número de personas en dicha población que son susceptibles y, por tanto, con riesgo de presentar la enfermedad, y el número de personas no susceptibles o inmunes y, por tanto, sin riesgo. Pueden ser inmunes porque han pasado previamente la enfermedad (y poseen anticuerpos) o porque han sido vacunados. También pueden no ser susceptibles por motivos genéticos. Evidentemente, si toda la población fuese inmune, no se desarrollarían epidemias. Pero el equilibrio suele verse alterado en algún punto entre inmunidad y susceptibilidad, y cuando se desplaza hacia la susceptibilidad, la probabilidad de que se produzca un brote aumenta. Esto ha sido observado particularmente en poblaciones previamente aisladas que fueron expuestas posteriormente a la enfermedad. Por ejemplo, en el siglo XIX, Panum observó que el sarampión se presentaba en las islas Feroe en forma epidémica cuando los individuos infectados entraban en contacto con la población aislada y susceptible[11]. En otro ejemplo, se produjeron brotes graves de faringitis estreptocócica cuando llegaron nuevos reclutas susceptibles a la Base Naval de los Grandes Lagos[12].

Inmunidad grupal

La *inmunidad grupal* se define como la resistencia de un grupo de personas al ataque de una enfermedad a la que una gran proporción de miembros del grupo son inmunes. Cuando un gran porcentaje de la población es inmune, es probable que toda la población se encuentre protegida, no solo aquellos que son inmunes. ¿Por qué se produce la inmunidad grupal? Se produce porque la enfermedad se propaga de una persona a otra en cualquier comunidad. Una vez que cierta proporción de personas en la comunidad son inmunes, disminuye la probabilidad de que una persona infectada encuentre una persona susceptible a la que transmitir la infección; la mayoría de sus encuentros serán con personas que son inmunes. La presencia de una gran proporción de personas inmunes en la población disminuye la probabilidad de que una persona con la enfermedad entre en contacto con un sujeto susceptible.

¿Por qué es tan importante el concepto de inmunidad grupal? Cuando se realizan programas de inmunización puede no ser necesario lograr tasas de inmunización del 100% para inmunizar a toda la población de modo eficaz. Podemos lograr una protección muy efectiva inmunizando a una gran parte de la población; la parte restante se encontrará protegida gracias a la inmunidad grupal.

Para que exista inmunidad grupal, se deben cumplir ciertas condiciones. El agente causante de la enfermedad debe limitarse a una sola especie de hospedador en la que se produce la transmisión y dicha transmisión debe ser relativamente directa de un miembro de la especie hospedadora a otro. Si existe un reservorio en el que el microorganismo puede existir fuera del hospedador humano, la inmunidad grupal no se producirá porque existen otros medios de transmisión. Además, las infecciones deben inducir una inmunidad total. Si la inmunidad es únicamente parcial, no se creará una gran proporción de personas inmunes en la comunidad.

¿Qué significa esto? La inmunidad grupal funciona cuando la probabilidad de que una persona infectada se encuentre con cualquier *otro individuo* de la población («mezcla aleatoria») es la misma. Pero si una persona se infecta y todas sus interacciones son con personas susceptibles (es decir, no se produce una mezcla aleatoria de la población), es probable que transmita la enfermedad a otras personas susceptibles. La inmunidad grupal funciona de manera óptima cuando las poblaciones se mezclan constantemente entre sí. Este es un concepto teórico, porque, evidentemente, las poblaciones nunca se mezclan aleatoriamente por completo. Todos nosotros nos relacionamos con familiares y amigos, por ejemplo, más que con extraños. Sin embargo, el grado de inmunidad grupal que se consigue depende de la extensión de la mezcla aleatoria que se produzca en la población. Por tanto, podemos interrumpir la transmisión de una enfermedad incluso aunque no todos los miembros de la población sean inmunes siempre que un porcentaje

importante de la población lo sea.

¿Qué porcentaje de una población debe ser inmune para que opere la inmunidad grupal? Este porcentaje varía de una enfermedad a otra. Por ejemplo, en el caso del sarampión, que es muy contagioso, se ha estimado que el 94% de la población debe ser inmune para que se interrumpa la cadena de transmisión. Con la disminución de las tasas de inmunización infantil en Estados Unidos, asociada con las preocupaciones de los padres con respecto al riesgo de que produzcan trastornos del espectro autista, los brotes de sarampión son cada vez más comunes. Un total de 125 casos de sarampión con erupción ocurrió en un periodo de 6 semanas; entre estos casos, 110 eran residentes de California (45% no vacunados), de los cuales el 35% había visitado uno o ambos parques temáticos de Disney entre el 17 y el 20 de diciembre de 2014, la fuente sospechosa de exposición. De los casos secundarios, la mayoría (26/34) fueron contactos cercanos. Se notificaron 15 casos adicionales relacionados con los parques temáticos de Disney en siete estados adicionales[13].

Consideremos la vacunación de la poliomielitis y la inmunidad grupal. De 1951 a 1954 se produjo una media de 24.220 casos de poliomielitis paralítica al año en Estados Unidos. Se dispone de dos tipos de vacunas. La vacuna de la poliomielitis oral (VPO) no solo protege a los vacunados, sino que también protege a otros en la comunidad a través de la inmunización secundaria que se produce cuando el sujeto vacunado propaga el virus de la vacuna activa a los contactos. En efecto, los contactos son inmunizados por la propagación del virus a partir de la persona vacunada. Si una proporción suficiente de personas en la comunidad es protegida de este modo, la cadena de transmisión se interrumpe. Sin embargo, incluso la vacuna de poliovirus inactivados (VPI), que no produce inmunidad secundaria (no propaga el virus a individuos susceptibles), puede producir inmunidad grupal si una proporción suficiente de la población es inmunizada. Incluso los no inmunizados estarán protegidos porque la cadena de transmisión en la comunidad se ha interrumpido.

De 1958 a 1961 en Estados Unidos solo se disponía de la VPI. En la figura 2.11A se muestra el número esperado de casos anuales si la vacuna hubiese protegido únicamente a los que recibieron la vacuna. En la figura 2.11B se muestra el número de casos de poliomielitis observados realmente. Resulta claro que el número de casos reales fue muy inferior al que cabría esperar solo por los efectos directos de la vacuna. La diferencia entre las dos curvas representa el efecto de la inmunidad grupal de la vacuna. Así, los sujetos no inmunizados pueden obtener cierta protección a partir de la VPO o de la VPI.

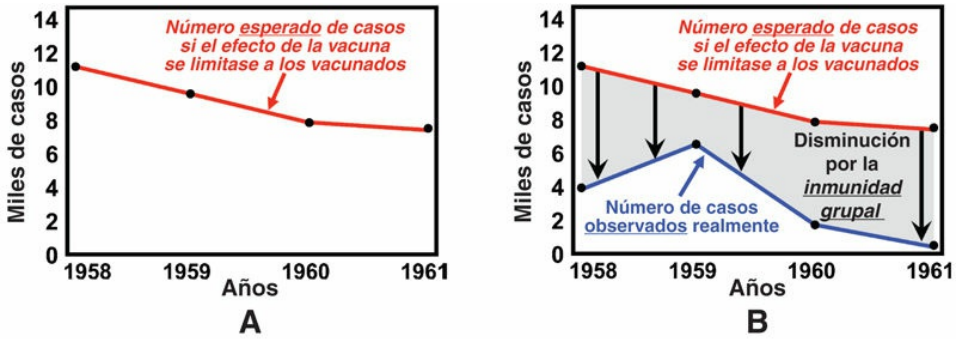


FIG. 2.11 Efecto de la inmunidad grupal, Estados Unidos, 1958-1961. (A) Número esperado de casos de poliomielitis paralítica si el efecto de la vacuna se limitase a la población vacunada. (B) Número de casos observados como resultado de la inmunidad grupal. (Modificado de American Academy of Pediatrics News. Copyright 1998. De Stickle G. Observed and expected poliomyelitis in the United States, 1958–1961. *Am J Public Health.* 1964;54:1222–1229.)

Periodo de incubación

El periodo de incubación se define como el *intervalo desde la recepción de la infección hasta el momento de aparición de la infección clínica* (la aparición de síntomas reconocibles). Si usted se infecta hoy, la enfermedad con la que se ha infectado puede que no se desarrolle durante varios días o semanas. Durante este tiempo, *periodo de incubación,* usted se encuentra completamente bien y no presenta signos de la enfermedad.

¿Por qué no se desarrolla la enfermedad inmediatamente después de la infección? ¿Qué quiere decir periodo de incubación? Puede reflejar el tiempo necesario para que el microorganismo se replique lo suficiente hasta que alcanza la masa crítica necesaria para que se produzca la enfermedad clínica. Probablemente también se relacione con la localización corporal en la que se replica el microorganismo, bien sea superficialmente, cerca de la superficie de la piel o en localizaciones más profundas (p. ej., en el intestino). La dosis del microorganismo infeccioso recibida en el momento de la infección también puede influir en la duración del periodo de incubación. Cuando la dosis es grande, el periodo de incubación puede ser más corto.

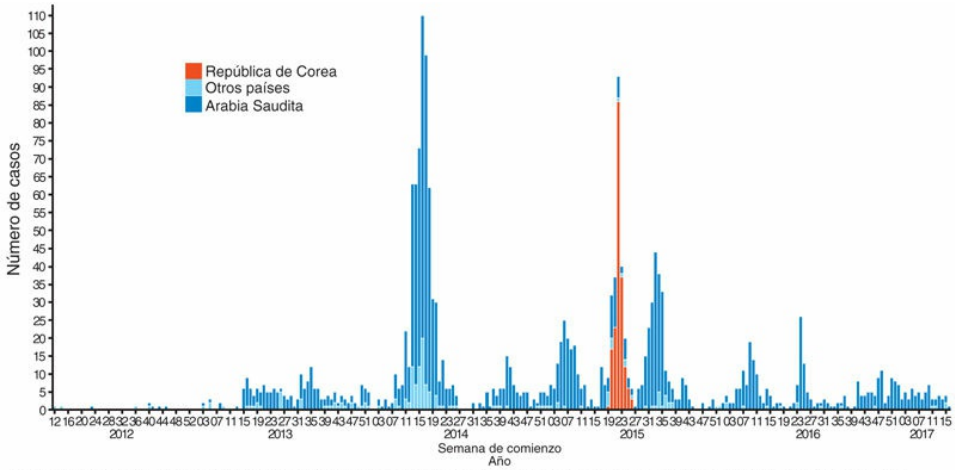
El periodo de incubación también tiene interés histórico porque se relaciona con lo que puede haber sido el único avance médico asociado con la peste negra en Europa. En 1374, cuando la población estaba aterrorizada por la peste negra, la república veneciana nombró a tres oficiales para que se responsabilizaran de inspeccionar todos los barcos que atracasen en el puerto e impidieran la entrada a los barcos con enfermos a bordo. Se esperaba que esta intervención protegiese a la comunidad. En 1377, en el puerto italiano de Ragusa, los viajeros eran detenidos en un área aislada durante treinta días *(trentini giorni)* tras la llegada para ver si desarrollaban alguna infección. Se observó que este periodo era insuficiente, por lo que el periodo de detención se alargó a 40 días *(quarante giorni).* Este fue el origen de la palabra *cuarentena*.

¿Durante cuánto tiempo querríamos mantener aislada a una persona? Quisiéramos aislar a una persona hasta que ya no sea infecciosa para el resto (tras concluir el periodo de incubación sospechoso). Cuando una persona se encuentra enferma desde el punto de vista clínico, tenemos por lo general un signo claro de infecciosidad. Un problema importante surge *antes* de que la persona presente la enfermedad, es decir, durante el periodo de incubación. Si supiéramos cuándo se infectó y también pudiéramos saber la duración general del periodo de incubación de la enfermedad, podríamos aislar a la persona infectada durante este periodo (y quizás unos días más para ser especialmente prudentes) para evitar la transmisión de la enfermedad a otros. En la mayoría de las situaciones, sin embargo, desconocemos que la persona ha sido infectada y puede que no lo sepamos hasta que los signos clínicos de la enfermedad se pongan de manifiesto. Además, es posible que no sepamos la duración del periodo de incubación.

Esto lleva a una pregunta importante: ¿merece la pena poner en cuarentena

(aislar) a un paciente, como, por ejemplo, un niño con varicela? El problema es que, durante al menos parte del periodo de incubación, cuando una persona aún no presenta una enfermedad clínica, es capaz de transmitir la enfermedad a otros. Por tanto, existen personas que no presentan (todavía) la enfermedad, pero que han sido infectadas, desconocen su estado infeccioso y pueden transmitir la enfermedad. Para muchas enfermedades comunes de la infancia, cuando la enfermedad se presenta en el niño, puede que ya haya contagiado la enfermedad a otros. Por tanto, el aislamiento de un paciente cuando ya presenta la enfermedad no será necesariamente eficaz. Por otra parte, el aislamiento puede ser muy valioso. En septiembre de 2012, las autoridades sanitarias de Arabia Saudita notificaron por primera vez una enfermedad respiratoria aguda grave con síntomas de fiebre, tos y dificultad respiratoria. Se demostró que el organismo causante es el coronavirus del síndrome respiratorio de oriente medio (MERS-CoV, *Middle East Respiratory Syndrome*), que tiene un periodo de incubación de aproximadamente 5 o 6 días. El MERS-CoV probablemente provino de camellos infectados en la península arábiga y se propagó a través del contacto cercano de persona a persona, con los profesionales sanitarios en mayor riesgo de infección si no se siguieron las precauciones universales. Todos los casos de MERS-CoV que se han identificado tienen una historia positiva de alguien que vive o viaja a países de la península arábiga o limítrofes. Otro brote de MERS-CoV ocurrió en la República de Corea en 2015 y también estuvo vinculado a un viajero que regresó de la península arábiga. Hasta mayo de 2017, la OMS informó que hubo 1.952 casos confirmados por laboratorio de infección por MERS-CoV en 27 países, de los cuales 693 (36%) fueron mortales.

La figura 2.12 muestra la curva epidémica de los casos globales confirmados de MERS-CoV notificados a la OMS hasta el 5 de mayo de 2017. (Tenga en cuenta que, a diferencia de la curva epidémica de epidemias por vehículos comunes, la curva para la propagación de persona a persona es multimodal.) Un brote de MERS-CoV observado en la República de Corea en 2015 fue bastante contenido, mientras que la epidemia permanece activa en Arabia Saudita. Un importante contribuyente al control de la epidemia coreana fueron probablemente las fuertes medidas de control de infecciones implementadas desde el principio para diagnosticar y aislar los casos probables de MERS-CoV y para reducir los contactos interpersonales de los viajeros con antecedentes de viajes a áreas altamente afectadas.



Otros países: Argelia, Austria, Bahréin, China, Egipto, Francia, Alemania, Grecia, Irán, Italia, Jordania, Kuwait, Líbano, Malasia, Países Bajos, Omán, Filipinas, Qatar, Tailandia, Túnez, Turquía, Emiratos Árabes Unidos, Reino Unido, Estados Unidos, Yemen.
Tenga en cuenta que los datos subyacentes están sujetos a cambios ya que las investigaciones sobre los casos están en curso. Fecha de inicio estimada si no está disponible.

FIG. 2.12 Curva epidémica de todos los casos globales confirmados de síndrome respiratorio de Oriente Medio por coronavirus (MERS-CoV) desde 2012 hasta abril de 2017. (World Health Organization. http://www.who.int/emergencies/mers-cov/mers-epi-5-may-2017.png?ua=1. Consultado el 14 de mayo de 2017.)

Las diferentes enfermedades poseen diferentes periodos de incubación. No existe un periodo de incubación preciso para una enfermedad dada; en cambio, es característico un rango de periodos de incubación para dicha enfermedad. En la figura 2.13 se muestra el rango de periodos de incubación de varias enfermedades. Por lo general, la duración del periodo de incubación es característica del microorganismo infeccioso.

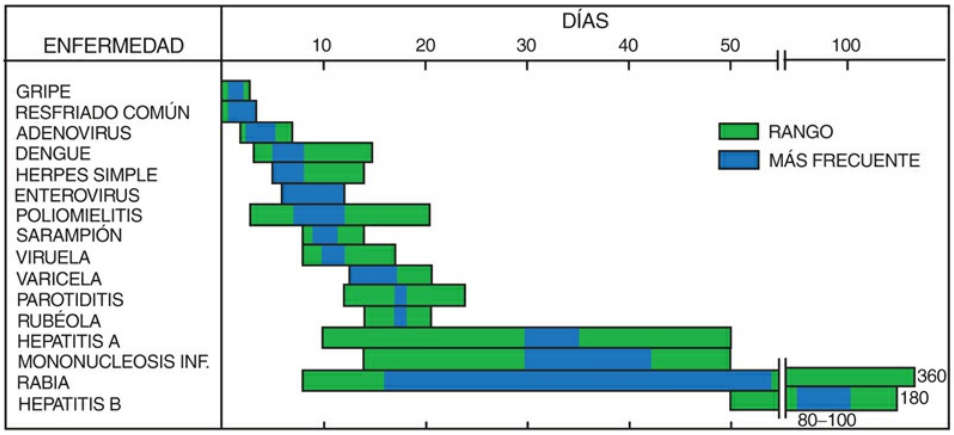


FIG. 2.13 Periodos de incubación de las enfermedades víricas. *INF,* infecciosa. (De Evans AS, Kaslow RA, eds. *Viral Infections of Humans: Epidemiology and Control,* 4.ª ed. Nueva York: Plenum; 1997.)

El periodo de incubación de las enfermedades infecciosas tiene su análogo en las enfermedades no infecciosas. Por tanto, incluso cuando un individuo se expone a un carcinógeno o a otra toxina ambiental, la enfermedad a menudo se manifiesta solo meses o incluso años después. Por ejemplo, los mesoteliomas debidos a la exposición al asbesto pueden producirse 20-30

años después de la exposición. El periodo de incubación de las enfermedades no infecciosas a menudo se conoce como periodo de latencia.

En la figura 2.14 se expone una representación gráfica de un brote de *Salmonella typhimurium* en un congreso médico en Gales en 1986. Cada barra representa el número de casos de la enfermedad que se produjo en un punto concreto en el tiempo tras la exposición; el número de horas desde la exposición se muestra en el eje horizontal. Si trazamos una línea que conecte la parte superior de las barras, obtenemos la denominada *curva epidémica*, que se define como la distribución de los tiempos de comienzo de la enfermedad. En una *epidemia con vehículo común y exposición única*, la curva epidémica representa la distribución de los periodos de incubación. Esto debería poder verse de forma intuitiva: si la infección se produjo en un momento en el tiempo, el intervalo entre dicho punto y el comienzo de cada caso es el periodo de incubación en dicha persona.

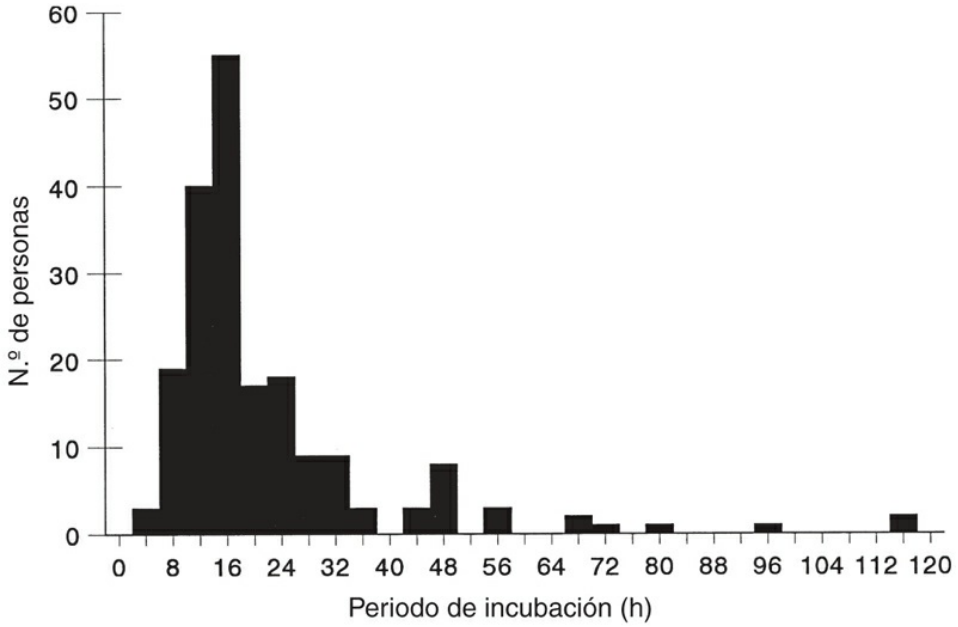


FIG. 2.14 Periodos de incubación en 191 delegados afectados por un brote de *Salmonella typhimurium* en un congreso médico en Gales, 1986. (Modificado de Gwynn JR, Palmer SR. Incubation period, severity of disease, and infecting dose: evidence from a *Salmonella* outbreak. *Am J Epidemiol.* 1992;136:1369–1377.)

Como se observa en la figura 2.14, se produjo un aumento rápido, explosivo, en el número de casos de infección por *Salmonella typhimurium* en las primeras 16 horas, lo que sugiere una epidemia con vehículo común y exposición única. De hecho, este patrón es la curva clásica de un brote epidémico con vehículo común y exposición única (fig. 2.15, *izquierda*). El motivo de esta configuración es desconocido, pero posee una propiedad interesante: si la curva se traza teniendo en cuenta el logaritmo del tiempo en vez del tiempo, se obtiene una curva normal, que posee propiedades estadísticas útiles (v. fig. 2.15, *derecha*). Si se dibuja sobre papel gráfico logarítmico normal, obtenemos una línea recta, lo que facilita la estimación

del periodo de incubación medio. Armenian y Lilienfeld[14] demostraron que una curva logarítmica normal también es típica de epidemias de enfermedades no infecciosas con vehículo común y exposición única.

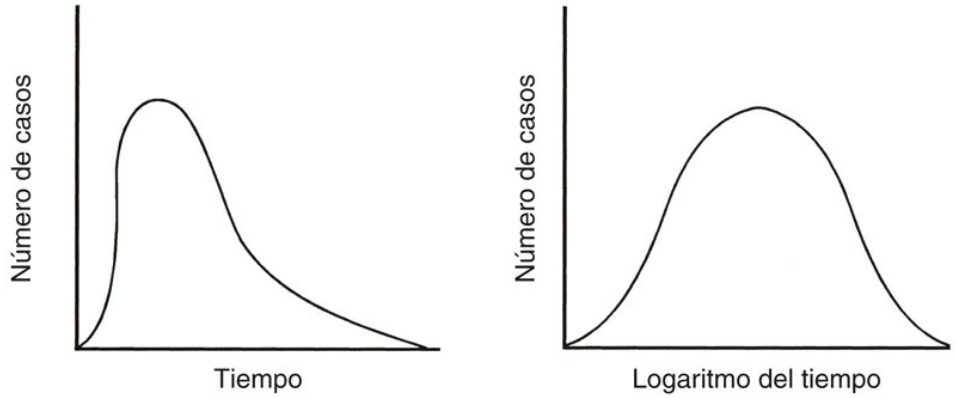


FIG. 2.15 Número de casos representados frente al tiempo y el logaritmo del tiempo.

Las tres variables críticas al investigar un brote o una epidemia son:

1. ¿Cuándo tuvo lugar la exposición?
2. ¿Cuándo comenzó la enfermedad?
3. ¿Cuál fue el periodo de incubación de la enfermedad?

Si conocemos dos de ellas podemos calcular la tercera.

Tasa de ataque

La tasa de ataque se define como:

$$\frac{\text{Número de personas en riesgo en las que se produce una cierta enfermedad}}{\text{Número total de personas en riesgo}}$$

La tasa de ataque es útil para comparar el riesgo de la enfermedad en grupos con exposiciones diferentes. La tasa de ataque puede ser específica para una exposición determinada. Por ejemplo, la tasa de ataque en personas que consumieron un cierto alimento se denomina *tasa de ataque específica de alimento*. Se calcula así:

$$\frac{\text{Número de personas que consumieron un cierto alimento y enfermaron}}{\text{Número total de personas que consumieron dicho alimento}}$$

Por lo general, el *tiempo* no se especifica explícitamente en la tasa de ataque porque la exposición es común y la enfermedad aguda; dado que suele saberse cuánto tiempo transcurre hasta que aparecen la mayoría de los casos tras una exposición, el periodo de tiempo está implícito en la tasa de ataque.

Una persona que adquiere la enfermedad a partir de una exposición (p. ej., a partir de un alimento contaminado) se denomina *caso primario*. Una persona que adquiere la enfermedad tras la exposición a un caso primario se denomina *caso secundario*. La *tasa de ataque secundario* se define, por tanto, como la tasa de ataque en personas susceptibles que no han estado expuestas al agente sospechoso, pero han estado expuestas a un caso primario. Es una buena medida de la propagación de la enfermedad de una persona a otra una vez que la enfermedad se ha introducido en una población, y puede considerarse como una onda que se mueve a partir del caso primario. La tasa de ataque secundario a menudo se calcula en los familiares del caso índice.

La tasa de ataque secundario también posee aplicación en enfermedades no infecciosas cuando se estudia a miembros de una familia para determinar la extensión con la cual una enfermedad se agrupa entre familiares de primer grado de un caso índice (herencia o agrupamiento familiar), lo que puede

proporcionar una hipótesis sobre las contribuciones relativas de los factores genéticos y ambientales a la causa de una enfermedad.

Exploración de la ocurrencia de la enfermedad

Los conceptos expuestos en este capítulo forman la base para el estudio de la ocurrencia de la enfermedad. Cuando parece que una enfermedad ha ocurrido por encima de un nivel endémico (habitual) y queremos investigar su ocurrencia, nos preguntamos:

¿*Quién* fue afectado por la enfermedad?
¿*Cuándo* ocurrió la enfermedad?
¿*Dónde* surgieron los casos?

Es un hecho bien conocido que el riesgo de enfermedad se ve afectado por todos estos factores.

Quién

Las características del hospedador humano se relacionan claramente con el riesgo de enfermedad. Factores como el sexo, la edad y la raza, así como factores de riesgo conductuales (p. ej., tabaquismo), ejercen un efecto importante.

Gonorrea

Como se muestra en la figura 2.16, históricamente la frecuencia de la gonorrea ha sido más elevada en varones, y esta diferencia por sexo se observa al menos desde 1960 (no se expone en este gráfico). Como la probabilidad de permanecer asintomático es más elevada en las mujeres, la enfermedad en ellas probablemente se ha comunicado en una cifra menor de la real. La frecuencia se había ido igualando en los varones y las mujeres a lo largo de las últimas décadas, pero desde 2013 se han observado tasas más altas de gonorrea en hombres que en mujeres. Dichos aumentos en las tasas entre los hombres podrían atribuirse a una mayor transmisión o a una mayor detección de casos (p. ej., a través de un aumento del cribado extragenital) entre homosexuales, bisexuales y otros hombres que tienen relaciones sexuales con hombres.

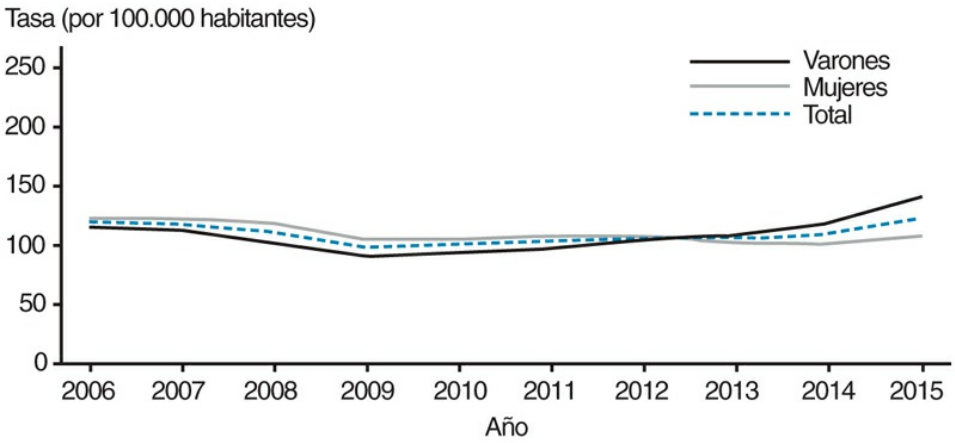


FIG. 2.16 Gonorrea: tasa de casos notificados por sexo, Estados Unidos, 2006-2015. (De Centers for Disease Control and Prevention: *Sexually transmitted disease surveillance 2010*. Atlanta: U.S. Department of Health and Human Services; 2016. http://www.cdc.gov/std/stats10/figures/15.htm. Consultado el 8 de mayo de 2017.)

Tos ferina

La incidencia de tos ferina en Estados Unidos alcanzó su máximo en 2004, con cifras de 8,9 casos por 100.000 habitantes, más del doble de las comunicadas en 2003. En 1994, la incidencia era de 1,8. El número de casos comunicados en 2004 fue el más elevado desde 1959. Aunque los niveles de cobertura de la vacuna frente a la tos ferina son altos en la población infantil de Estados Unidos, esta enfermedad sigue causando morbilidad. Parte de este aumento se debe a las mejoras diagnósticas, así como al mayor reconocimiento y comunicación de los casos. Como se observa en la figura 2.17, las tasas más bajas de tos ferina en Estados Unidos se observaron en 1991. Si bien las tasas de incidencia mostraron dos picos más en 2008 y 2009, posteriormente disminuyeron hasta 2016. Cabe destacar que los lactantes menores de 1 año, que tienen el mayor riesgo de muerte, continúan teniendo la tasa más alta de tos ferina registrada.

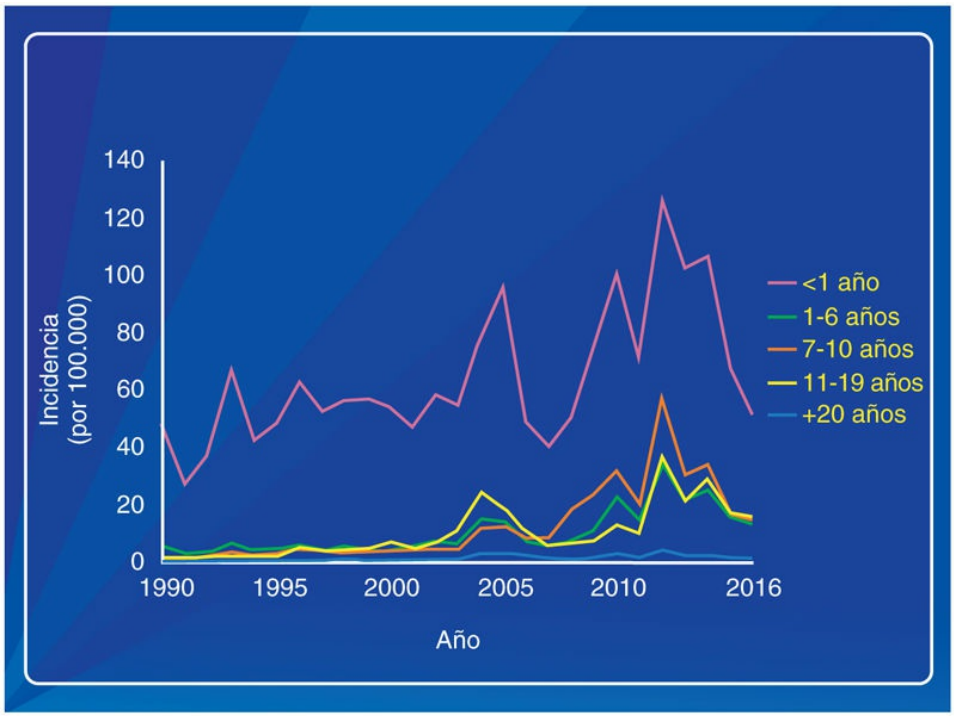


FIG. 2.17 Incidencia de tos ferina por 100.000 habitantes por año y grupo de edad, Estados Unidos, 1990-2016. (De Centers for Disease Control and Prevention. https://www.cdc.gov/pertussis/images/incidence-graph-age.png. Consultado el 13 de junio de 2018.)

La ocurrencia de la tos ferina se relaciona claramente con la edad (fig. 2.18). Aunque la *tasa* más elevada de tos ferina se observa en lactantes de menos de 6 meses de edad (99 por 100.000 habitantes), el *número* de casos comunicados fue más elevado en los niños de 11 a 19 años. Aproximadamente la mitad de los casos de tos ferina comunicados en 2014 y 2015 se produjeron en niños de 10 a 19 años y en adultos de más de 20 años. Aunque la causa específica de este fenómeno es desconocida, podría deberse a la pérdida de la protección 5-10 años después de la vacunación frente a la tos ferina.

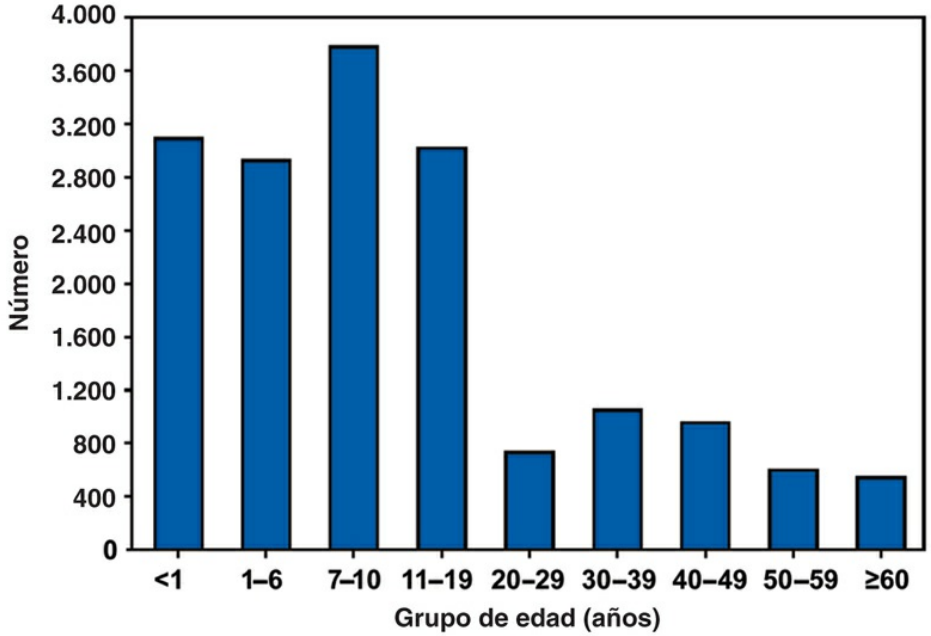


FIG. 2.18 Número de casos de tos ferina declarados por grupo de edad, Estados Unidos, 2009. (De Centers for Disease Control and Prevention. Summary of notifiable diseases, United States, 2009. *Morb Mortal Wkly Rep*. 2011;58:1–100.)

Cuándo

Algunas enfermedades se producen con cierta periodicidad. Por ejemplo, la meningitis aséptica presenta valores máximos anualmente de modo consistente (fig. 2.19). A menudo dicha variación temporal presenta un patrón estacional. Por ejemplo, los cuadros diarreicos son más frecuentes durante los meses de verano y las enfermedades respiratorias son más frecuentes durante los meses de invierno. La cuestión del *cuándo* también se plantea cuando se estudian las tendencias de la incidencia de una enfermedad a lo largo del tiempo. Por ejemplo, en Estados Unidos, tanto la incidencia como el número de fallecidos debidos al síndrome de inmunodeficiencia adquirida (SIDA) aumentaron durante muchos años, pero comenzaron a disminuir en 1996, debido principalmente a los nuevos tratamientos y a los esfuerzos en educación sanitaria.

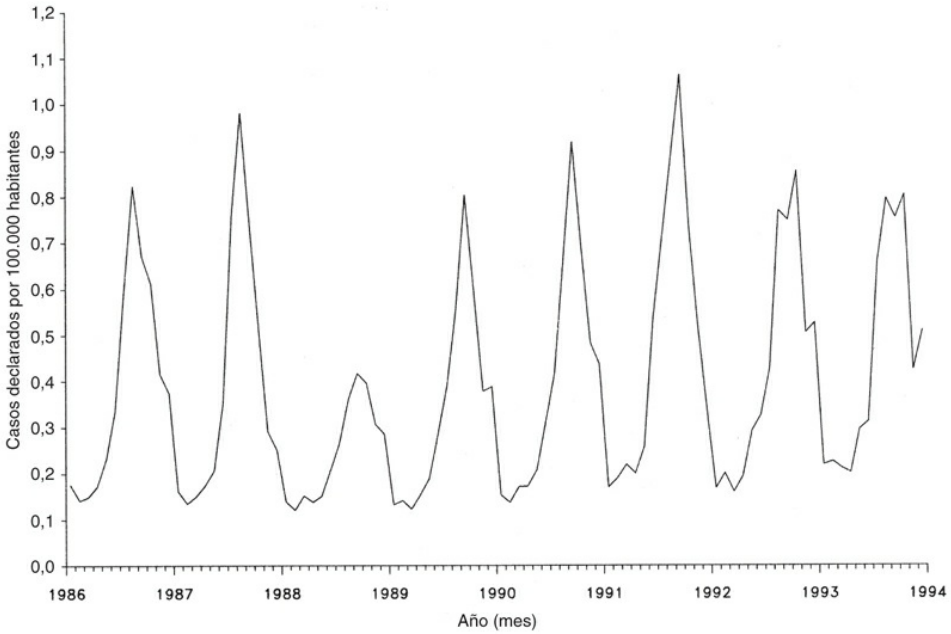


FIG. 2.19 Casos declarados de meningitis aséptica al mes por 100.000 habitantes, Estados Unidos, 1986-1993. (De Centers for Disease Control and Prevention. Summary of notifiable diseases, United States, 1993. *Morb Mortal Wkly Rep.* 1994;42:22.)

Dónde

La enfermedad no se distribuye de forma aleatoria en el tiempo o en el espacio. Por ejemplo, en la figura 2.20 se expone la distribución geográfica de la enfermedad de Lyme en Estados Unidos en 2015, en la que cada punto representa un caso de enfermedad de Lyme. Se observa un claro agrupamiento de casos a lo largo de la costa noreste, en la parte norte-central del país y en la región de la costa del Pacífico. Los estados en los que se han comunicado ciclos enzoóticos establecidos de *Borrelia burgdorferi,* el agente causal, representan el 95% de los casos. La distribución de la enfermedad está estrechamente relacionada con la del vector, la garrapata del ciervo.

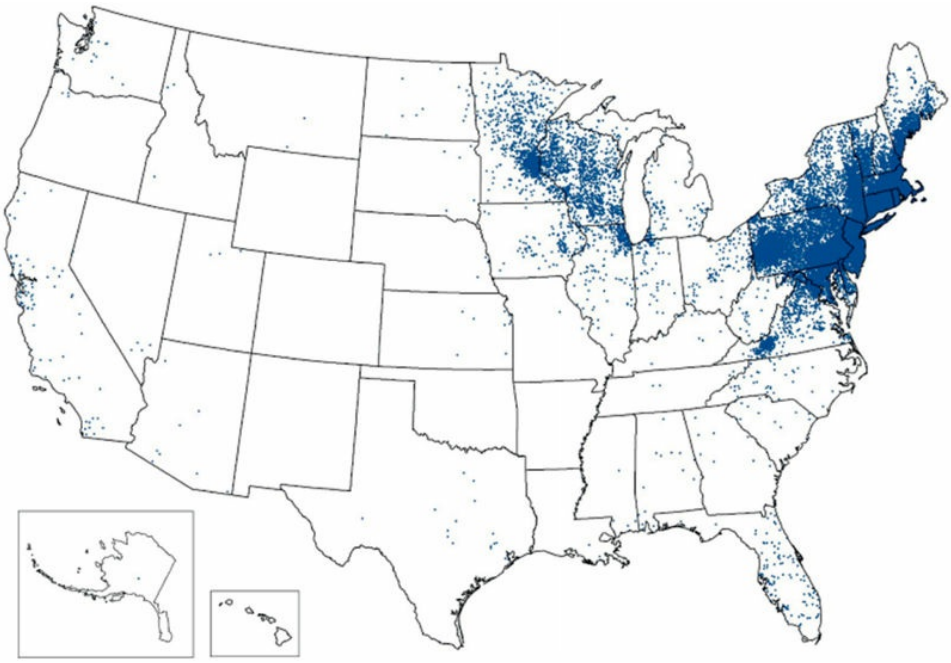
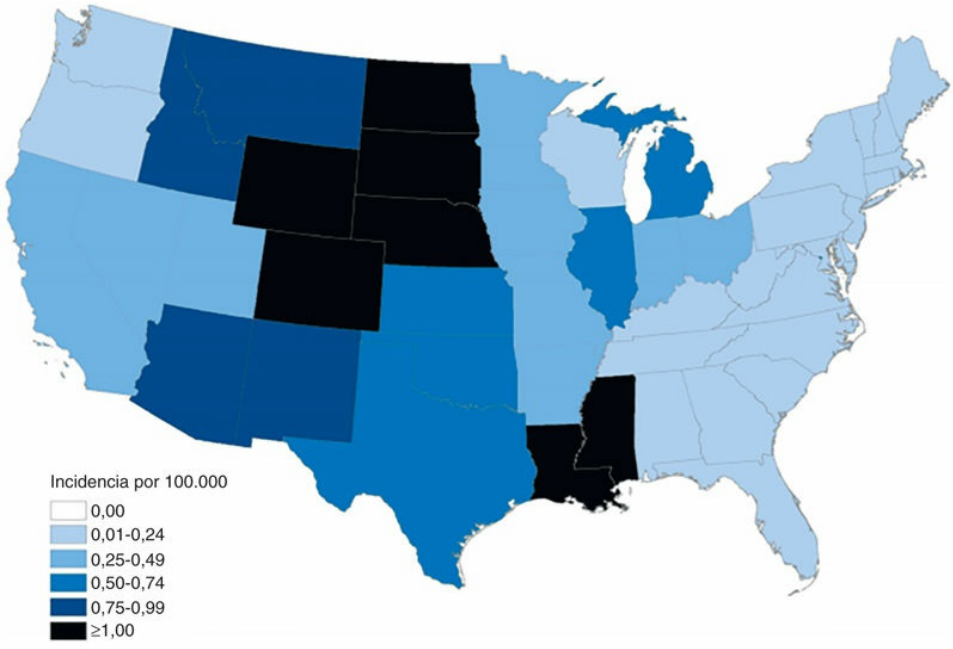


FIG. 2.20 Casos notificados de enfermedad de Lyme en Estados Unidos, 2015. Cada punto representa un caso de enfermedad de Lyme y se coloca aleatoriamente en el condado de residencia del paciente. La presencia de un punto en un estado no significa necesariamente que la enfermedad de Lyme se haya adquirido en ese estado. Las personas viajan entre estados y el lugar de residencia a veces es diferente del lugar donde el paciente se infectó. (De Centers for Disease Control and Prevention. Summary of notifiable diseases, United States, 2015. https://www.cdc.gov/lyme/resources/reportedcasesoflymedisease_2015. pdf. Consultado el 8 de mayo de 2017.)

Un ejemplo espectacular de propagación de la enfermedad se observa con el virus del Nilo occidental (VNO) en Estados Unidos[15]. El VNO fue identificado y aislado por primera vez en 1937 en la región del Nilo occidental de Uganda, y durante muchos años solo se encontraba en el hemisferio oriental. El ciclo básico de la enfermedad es pájaro-mosquito-pájaro. Los mosquitos son infectados cuando pican a pájaros infectados. Cuando los mosquitos que pican a pájaros y seres humanos se infectan, suponen una amenaza para las personas. La mayoría de las infecciones humanas son subclínicas, pero en los últimos años aproximadamente 1 de cada 150 infecciones causan meningitis o encefalitis. El riesgo de sufrir una enfermedad neurológica aumenta de modo importante en las personas de más de 50 años. Otros síntomas son fiebre, náuseas y vómitos, exantema, cefalea y debilidad muscular. La relación casos-mortalidad, o la proporción de personas que sufren la enfermedad (casos) y que posteriormente fallecen debido a la enfermedad, puede ser de hasta el 14%. La edad avanzada es un factor de riesgo importante de mortalidad debida al VNO. Un estudio ha encontrado que la mortalidad es nueve veces más frecuente en pacientes de edad avanzada en comparación con pacientes más jóvenes. El tratamiento es de soporte y la prevención se realiza principalmente a través del control de los mosquitos y el uso de repelentes de insectos y mosquiteras. El seguimiento de la distribución de la enfermedad depende de la vigilancia de

los casos humanos y el control de casos de enfermedad o de muertes por la enfermedad en los pájaros y otros animales. La vigilancia se estudia con mayor detalle en el capítulo 3.

El VNO se identificó por primera vez en la ciudad de Nueva York en 1999. En la figura 2.21 se muestra la incidencia media anual de los casos de enfermedad neuroinvasiva por VNO notificados por los estados a los CDC de 1999 a 2015. Durante el mismo periodo, la epidemia de casos de enfermedad por el VNO alcanzó su máximo anual en los meses de septiembre (fig. 2.22). Aún queda mucho por aprender sobre esta enfermedad para facilitar su tratamiento, prevención y control.



Source: ArboNET, Arboviral Diseases Branch, Centers for Disease Control and Prevention

FIG. 2.21 Incidencia media anual por estado de casos de enfermedad neuroinvasiva por virus del Nilo occidental notificados a los Centers for Disease Control and Prevention, 1999-2015. (De Centers for Disease Control and Prevention. https://www.cdc.gov/westnile/resources/pdfs/data/6-WNV-Neuro-Incidence-by-State-Map_1999-2015_07072016.pdf. Consultado el 8 de mayo de 2017.)

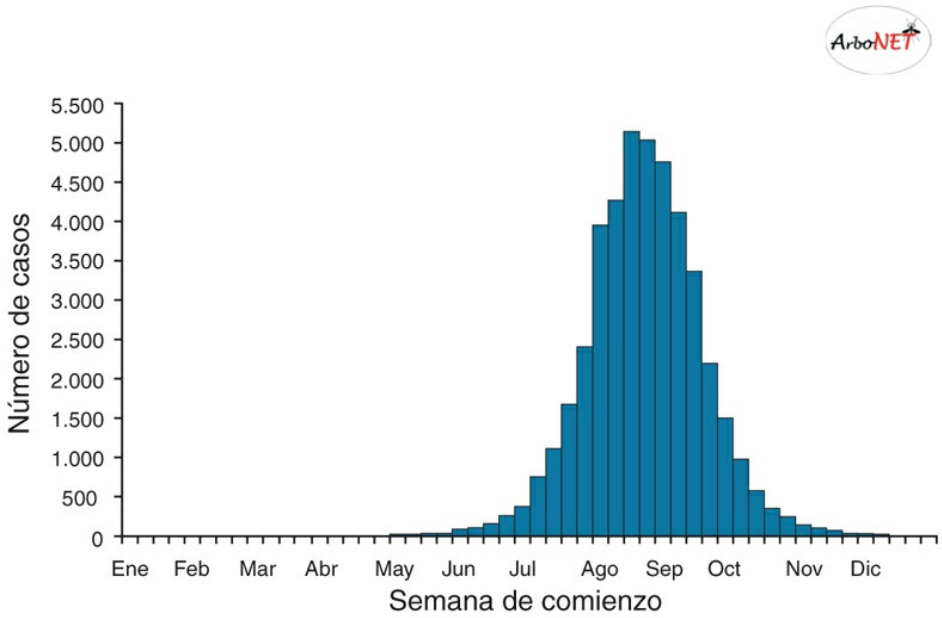


FIG. 2.22 Número de casos de enfermedad por el virus del Nilo occidental notificados a los Centers for Disease Control and Prevention por semana de comienzo de la enfermedad, 1995-2015 (De Centers for Disease Control and Prevention. https://www.cdc.gov/westnile/resources/pdfs/data/4-WNV-Week-Onset_for-PDF_1999-2015_07072016.pdf. Consultado el 8 de mayo de 2017.)

Investigación del brote

Las características que acabamos de exponer constituyen el núcleo central en la investigación de prácticamente todos los brotes. Los pasos para investigar un brote por lo general siguen este patrón (cuadro 2.2).

Cuadro 2.2 Pasos en la investigación de un brote agudo

La investigación de un brote agudo puede ser principalmente deductiva (es decir, razonando a partir de premisas o proposiciones probadas previamente) o inductiva (es decir, razonando a partir de hechos particulares hasta lograr una conclusión general), o puede ser una combinación de ambas.

Las consideraciones importantes que hay que tener en cuenta al investigar un brote agudo de enfermedades infecciosas son la determinación de que realmente el brote ha tenido lugar y la definición de la cantidad de población en riesgo, la determinación de la medida de la propagación y el reservorio y la caracterización del microorganismo.

Los pasos utilizados habitualmente se enumeran a continuación, pero dependiendo del brote, el orden exacto puede ser diferente.

1. Definir el brote y validar la existencia de un brote.
    a. Definir el «numerador» (casos).
        1) Características clínicas: ¿se conoce la enfermedad?
        2) ¿Cuáles son sus aspectos serológicos o culturales?
        3) ¿Se conocen las causas parcialmente?
    b. Definir el «denominador»: ¿cuál es la población con riesgo de presentar la enfermedad (es decir, susceptible)?
    c. Determinar si el número observado de casos supera claramente el número esperado.
    d. Calcular las tasas de ataque.
2. Estudiar la distribución de los casos en función de:
    a. Tiempo
    b. Lugar
    } Busque interacciones tiempo-lugar
3. Buscar combinaciones (interacciones) de variables relevantes.
4. Plantear hipótesis basadas en lo siguiente:
    a. El conocimiento existente sobre la enfermedad (en caso de existir).
    b. Analogías con enfermedades de etiología conocida.
    c. Hallazgos de la investigación del brote.
5. Comprobar las hipótesis.
    a. Analizar aún más los datos existentes (estudios de casos y controles).

b. Mejorar las hipótesis y obtener datos adicionales que puedan ser necesarios.
6. Recomendar medidas de control.
   a. Control del brote actual.
   b. Prevención de brotes similares futuros.
7. Preparar un informe escrito de la investigación y los hallazgos.
8. Comunicar los hallazgos a las partes implicadas en la creación e implementación de normativas y al público.

Tabulación cruzada

Cuando nos encontramos con varios agentes causales posibles, como a menudo es el caso en los brotes de enfermedades de origen alimentario, un método muy útil para determinar cuál de los posibles agentes probablemente sea la causa es la llamada *tabulación cruzada*. Esto lo vamos a ilustrar en un ejemplo de un brote de enfermedad estreptocócica de origen alimentario comunicado por los CDC en una cárcel de Florida hace algunos años[16].

En agosto de 1974, un brote de faringitis por estreptococos β-hemolíticos del grupo A afectó a 325 de 690 internos. En un cuestionario entregado a 185 internos seleccionados aleatoriamente, el 47% comunicaron haber presentado faringitis entre el 16 de agosto y el 22 de agosto. En un segundo cuestionario, las tasas de ataque específicas de alimento para víveres que fueron servidos a internos seleccionados aleatoriamente demostraron una asociación entre dos alimentos y el riesgo de presentar faringitis: la bebida y la ensalada de huevo servidas en el almuerzo del 16 de agosto (tabla 2.2).

Tabla 2.2

Tasas de ataque específicas de alimento de los productos consumidos el 16 de agosto de 1974 en la cárcel del condado de Dade, Miami

	COMIERON			NO COMIERON			
Producto consumido	Enfermos	Total	% de enfermos (tasa de ataque)	Enfermos	Total	% de enfermos (tasa de ataque)	*p*
Bebida	179	264	67,8	22	50	44,0	<0,010
Sándwiches de ensalada de huevo	176	226	77,9	27	73	37,0	<0,001

De Centers for Disease Control and Prevention: Outbreak of foodborne streptococcal disease. *MMWR*. 1974;23:365.

En la tabla 2.2 se calcula la tasa de ataque para cada una de las exposiciones

sospechosas (bebida y ensalada de huevo) en aquellos que comieron o bebieron los alimentos (estuvieron expuestos) y en aquellos que no lo hicieron (no estuvieron expuestos). Para la bebida y la ensalada de huevo, las tasas de ataque son claramente superiores entre los que bebieron o comieron el alimento sospechoso que entre los que no lo hicieron. Sin embargo, esta tabla no nos permite determinar qué alimento fue el responsable del brote, si la ensalada de huevo o la bebida.

Para responder a esta pregunta, empleamos la técnica de tabulación cruzada. En la tabla 2.3 examinamos de nuevo las tasas de ataque de los que comieron ensalada de huevo en comparación con los que no lo hicieron, pero esta vez lo hacemos separadamente para aquellos que bebieron la bebida y para aquellos que no lo hicieron.

Tabla 2.3

Análisis mediante tabulación cruzada de la ensalada de huevo y la bebida consumidas el 16 de agosto de 1974 en la cárcel del condado de Dade, Miami

	COMIERON ENSALADA DE HUEVO				NO COMIERON ENSALADA DE HUEVO			
	Enfermos	Sanos	Total	% de enfermos (tasa de ataque)	Enfermos	Sanos	Total	% de enfermos (tasa de ataque)
Tomaron bebida	152	49	201	75,6	19	53	72	26,4
No tomaron bebida	12	3	15	80,0	7	21	28	25,0

De Centers for Disease Control and Prevention: Outbreak of foodborne streptococcal disease. *MMWR*. 1974;23:365.

Observando los datos por columnas, vemos que tanto entre los que comieron ensalada de huevo como entre los que no lo hicieron, el hecho de beber la bebida no aumentaba la incidencia de enfermedad estreptocócica (75,6% frente a 80% y 26,4% frente a 25%, respectivamente). Sin embargo, analizando los datos en las filas, observamos que comer la ensalada de huevo aumentaba la tasa de ataque de la enfermedad, tanto en los que bebieron la bebida (75,6% frente a 26,4%) como en los que no la bebieron (80% frente a 25%). Por tanto, la ensalada de huevo está claramente implicada como fuente de las infecciones. En el capítulo 15 se aborda con más detalle el análisis y la interpretación de la tabulación cruzada.

Este ejemplo demuestra el uso de la tabulación cruzada en un brote de una enfermedad infecciosa transmitida por alimentos; no obstante, el método tiene gran aplicabilidad en cualquier enfermedad en la que se sospechen múltiples factores etiológicos. Se aborda con mayor detalle en el capítulo 15.

A veces varios agentes son responsables de un brote. Un ejemplo es un brote de enfermedad gastrointestinal que se produjo en un crucero el mismo día que una tormenta, lo que provocó que miles de millones de litros de escorrentía se contaminaran con aguas residuales que se habían liberado en el lago donde se realizó el crucero. La tabulación cruzada mostró que los pasajeros que consumieron hielo tenían una tasa de ataque de más del doble que la de los que no consumieron hielo. Las muestras de heces fueron positivas para múltiples agentes, incluyendo *Shigella sonnei* y *Giardia*[17].

Conclusión

Este capítulo ha revisado algunos conceptos básicos que subyacen en el enfoque epidemiológico de las enfermedades contagiosas agudas. Muchos de estos conceptos pueden aplicarse igual de bien a las enfermedades no transmisibles que en este momento no parecen ser de origen principalmente infeccioso. Además, en un número cada vez mayor de enfermedades crónicas que inicialmente se consideraba que no eran infecciosas, la infección parece desempeñar un cierto papel. Así, la infección por el virus de la hepatitis B es una causa importante del cáncer hepático primario. Las infecciones por papilomavirus y *Helicobacter pylori* son necesarias para la aparición del cáncer cervical y el cáncer gástrico, respectivamente, y el virus de Epstein-Barr se ha relacionado con la enfermedad de Hodgkin. El límite entre la epidemiología de las enfermedades infecciosas y no infecciosas se ha borrado en muchas áreas. Además, incluso para las enfermedades que no son de origen infeccioso puede existir un componente inflamatorio, los patrones de diseminación comparten una gran parte de las mismas dinámicas y los aspectos metodológicos para su estudio son similares. Muchos de estos temas se analizan en detalle en la sección II.

Bibliografía

1. Kipling R. Just-So Stories: The Elephant's Child, 1902. Reprinted by Everyman's Library Children's Classics. New York: Alfred A Knopf; 1992:79.
2. Mims CA. The Pathogenesis of Infectious Disease. 3rd ed. London: Academic Press; 1987.
3. Drüeke TB, Locatelli F, Clyne N, et al. CREATE Investigators. Normalization of hemoglobin level in patients with chronic kidney disease and anemia. *N Engl J Med*. 2006;355(20):2071–2084.
4. Singh AK, Szczech L, Tang KL, et al. CHOIR Investigators. Correction of anemia with epoetin alfa in chronic kidney disease. *N Engl J Med*. 2006;355(20):2085–2098.
5. Hearing on ensuring kidney patients receive safe and appropriate anemia management care. Hearing before the Subcommittee on Health of the Committee on Ways and Means, U.S. House of Representatives, 110th Congress, First Session, June 26, 2007, Serial No. 110-51 [fecha de última consulta: 6 de mayo de 2017]. https://www.gpo.gov/fdsys/pkg/CHRG-110hhrg49981/html/CHRG-110hhrg49981.htm.
6. Imai M, Watanabe T, Hatta M, et al. Experimental adaptation of an influenza H5 HA confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets. *Nature*. 2012;486:420–428.
7. Herfst S, Schrauwen EJ, Linster M, et al. Airborne transmission of influenza A/H5N1 virus between ferrets. *Science*. 2012;336:1534–1541.
8. Bartsch SM, Lopman BA, Ozawa S, et al. Global economic burden of norovirus gastroenteritis. *PLoS ONE*. 2016;11(4):e0151219.
9. Centers for Disease Control, Prevention. Outbreaks of gastroenteritis associated with noroviruses on cruise ships—United States. *2002. MMWR Morb Mortal Wkly Rep*. 2002;51:1112–1115.
10. Lu J, Fang L, Sun L, et al. Association of GII.P16-GII.2 recombinant norovirus strain with increased norovirus outbreaks, Guangdong, China. *2016. Emerg Infect Dis*.

2017;23(7):1188–1190.
11. Panum PL. Observations Made During the Epidemic of Measles on the Faroe Islands in the Year 1846. New York: Delta Omega Society, Distributed by the American Public Health Association; 1940.
12. Frank PF, Stollerman GH, Miller LF. Protection of a military population from rheumatic fever. *JAMA*. 1965;193:775.
13. Centers for Disease Control and Prevention. U.S. Multi-State Measles Outbreak, December 2014–January 2015. Atlanta, GA: US Department of Health and Human Services; 2015. http://emergency.cdc.gov/han/han00376.asp.
14. Armenian HK, Lilienfeld AM. The distribution of incubation periods of neoplastic diseases. *Am J Epidemiol*. 1974;99(2):92–100.
15. Petersen LR, Marfin AA. West Nile virus: a primer for the clinician. *Ann Intern Med*. 2002;137:173–179.
16. Outbreak of foodborne streptococcal disease. *MMWR*. 1974;23:365.
17. Serdarevic F, Jones RC, Weaver KN, et al. Multi-pathogen waterborne disease outbreak associated with a dinner cruise on Lake Michigan. *Epidemiol Infec*. 2012;140:621–625.

Preguntas de repaso del capítulo 2

1. *Que una enfermedad sea endémica significa:*

a. Que ocurre claramente por encima de la expectativa normal.

b. Que está presente habitualmente en las poblaciones humanas.

c. Que afecta a un gran número de países simultáneamente.

d. Que muestra un patrón estacional.

e. Que es prevalente en los animales.

Las preguntas 2 y 3 se basan en la información ofrecida a continuación:

La primera tabla muestra el número total de personas que comieron uno de los dos alimentos especificados que posiblemente estaban infectados por estreptococos del grupo A. La segunda tabla muestra el número de personas enfermas (con faringitis aguda) que comieron cada una de las diferentes combinaciones especificadas

de los alimentos.

Número total de personas que comieron cada combinación especificada de alimentos

	Comieron atún	No comieron atún
Comieron ensalada de huevo	75	100
No comieron ensalada de huevo	200	50

Número total de personas que comieron cada combinación especificada de alimentos y que posteriormente enfermaron (con faringitis aguda)

	Comieron atún	No comieron atún
Comieron ensalada de huevo	60	75
No comieron ensalada de huevo	70	15

2. *¿Cuál es la tasa de ataque de faringitis en las personas que comieron ensalada de huevo y atún?*

a. 60/75.

b. 70/200.

c. 60/135.

d. 60/275.

e. Ninguna de las anteriores.

3. *Según los resultados mostrados en las tablas precedentes, ¿cuál de los siguientes alimentos (o combinaciones de alimentos) es más probable que sea la causa de la infección?*

a. Solo el atún.

b. Solo la ensalada de huevo.

c. Ni el atún ni la ensalada de huevo.

d. Tanto el atún como la ensalada de huevo.

e. No puede calcularse a partir de los datos presentados.

4. *En el estudio de un brote de una enfermedad infecciosa, resulta útil dibujar la curva epidémica porque:*

a. Ayuda a determinar qué tipo de brote se ha producido (p. ej., de una sola fuente, de persona a persona).

b. Muestra si se ha producido inmunidad grupal.

c. Ayuda a determinar el periodo de incubación medio.

d. *a* y *c*.

e. a, b y c.

5. *¿Cuál de las siguientes respuestas es característica de un brote con vehículo común y exposición única?*

a. Casos secundarios frecuentes.

b. Mayor gravedad al aumentar la edad.

c. Explosivo.

d. Los casos comprenden personas que se han expuesto y que no se han expuesto.

e. Todas las anteriores.

6. *¿Cuál de las siguientes enfermedades generalizadas recientes se considera pandémica?*

a. Enfermedad por el virus del Ébola.

b. Síndrome respiratorio de Oriente Medio por coronavirus (MERS-CoV).

c. Gripe por el virus H1N1 (gripe porcina).

d. Sarampión.

CAPÍTULO 3

La ocurrencia de la enfermedad: I. Vigilancia de la enfermedad y medidas de la morbilidad

Palabras clave

vigilancia
incidencia
historia natural de la enfermedad
persona-tiempo
prevalencia
sistemas de información geográfica

Debemos todos los grandes avances del conocimiento a aquellos que se esfuerzan en descubrir cuánto había de todo.

—James Maxwell, físico (1831-79)

Si usted puede medir aquello de lo que habla, y puede expresarlo con un número, sabe algo sobre el tema, pero si no puede medirlo, su conocimiento es escaso e insatisfactorio.

—William Thomson, Lord Kelvin, ingeniero, matemático y físico (1824-1907)

Objetivos de aprendizaje

- Describir el importante papel de la epidemiología en la vigilancia de las enfermedades.
- Comparar las diferentes medidas de la morbilidad, como las tasas de incidencia, la incidencia acumulada, las tasas de ataque, la prevalencia y el concepto personas-tiempo en riesgo.
- Ilustrar por qué los datos sobre la incidencia son necesarios para medir el riesgo.
- Analizar la interrelación entre la incidencia y la

prevalencia.

- Describir las limitaciones al definir los numeradores y los denominadores de las medidas de la incidencia y la prevalencia.

En el capítulo 2 analizamos cómo se transmiten las enfermedades. Resulta claro a partir de dicha exposición que, con el fin de estudiar la transmisión de las enfermedades en poblaciones humanas, necesitamos ser capaces de medir la frecuencia de la ocurrencia de la enfermedad y las muertes debidas a la enfermedad. En este capítulo describiremos la vigilancia de la enfermedad en las poblaciones humanas y su importancia como fuente de información sobre la morbilidad de la enfermedad. A continuación, analizaremos cómo utilizamos las tasas y las proporciones para expresar el grado de morbilidad debida a la enfermedad, y en el capítulo 4 nos ocuparemos de expresar la extensión de la mortalidad en términos cuantitativos.

Vigilancia

La vigilancia tiene un papel fundamental en la salud pública. Los Centros para el Control y la Prevención de Enfermedades (CDC, Centers for Disease Control and Prevention) definieron la vigilancia epidemiológica como la «recogida sistemática continua, el análisis y la interpretación de datos sanitarios esenciales para la planificación, implementación y evaluación de prácticas de salud pública estrechamente integradas con la difusión oportuna de estos datos a quien necesite conocerlos»[1]. La vigilancia puede servir para estudiar los cambios en la frecuencia de la enfermedad o para evaluar los cambios en los niveles de los factores de riesgo de enfermedades específicas. Gran parte de nuestra información sobre la morbilidad y la mortalidad de la enfermedad proviene de programas de vigilancia sistemática de la enfermedad. La vigilancia se emplea comúnmente en las enfermedades infecciosas, pero en los últimos años se ha vuelto cada vez más importante para controlar los cambios en otros tipos de enfermedades, como malformaciones congénitas, enfermedades no transmisibles, patologías por toxinas ambientales y para lesiones o enfermedades tras desastres naturales, como huracanes o terremotos. Es el método principal a través del cual las agencias federales en Estados Unidos, como la Agencia de Protección Ambiental (EPA, Environmental Protection Agency), identifican los contaminantes de preocupación emergente (CEC, *contaminants of emerging concern*). La vigilancia también se utiliza para controlar el grado de cobertura de las campañas de vacunación y la protección de una población y para estudiar la prevalencia de microorganismos resistentes a fármacos, como el paludismo y la tuberculosis (TB).

Un elemento importante de esta y de otras definiciones de vigilancia es proporcionar a los responsables de establecer políticas sanitarias una guía para desarrollar e implementar las mejores estrategias para programas de prevención y control de enfermedades. Con el fin de que los países o estados desarrollen acciones de salud pública coordinadas, resultan esenciales los mecanismos de intercambio de información. Por tanto, son necesarias definiciones de enfermedades y criterios diagnósticos estandarizados que puedan aplicarse en diferentes países o para fines de vigilancia de la salud pública dentro de un país. El CDC define un caso de vigilancia como «el conjunto de criterios uniformes utilizados para definir una enfermedad para la salud pública», que tiene la intención de ayudar a los funcionarios de salud pública a registrar y notificar casos[2]. Esto es diferente de la definición clínica empleada por los médicos para realizar un diagnóstico clínico para iniciar el tratamiento y satisfacer las necesidades individuales de los pacientes. Los formularios utilizados para recopilar y notificar datos sobre diferentes enfermedades también deben ser estandarizados.

Vigilancia pasiva y activa

La *vigilancia pasiva* hace referencia a la vigilancia en la que se utilizan los datos disponibles sobre enfermedades comunicables o en los que la comunicación de enfermedades es obligatoria o solicitada por el gobierno o las autoridades sanitarias locales, estando encargado de la comunicación el proveedor de asistencia sanitaria o el responsable de salud del distrito. Este tipo de comunicación también se denomina *declaración pasiva*. La exactitud y la calidad de los datos comunicados dependen, por tanto, en gran medida de esta persona y del personal a su cargo, que a menudo se hacen cargo de esta misión sin fondos o recursos adicionales. Como resultado, es probable que se produzca infradeclaración y falta de exactitud de la comunicación. Para minimizar este problema, los instrumentos de comunicación deben ser simples y breves. Los ejemplos de enfermedades notificables incluyen infecciones de transmisión sexual comunes (sífilis, gonorrea, virus de inmunodeficiencia humana/síndrome de inmunodeficiencia adquirida [VIH/SIDA]). Cuando se utiliza la declaración pasiva, los brotes locales pueden ser pasados por alto debido a que el número relativamente menor de casos verificados se diluye en el amplio denominador de la población total de una provincia o país. Sin embargo, un sistema de declaración pasiva es relativamente barato y relativamente fácil de desarrollar inicialmente. El seguimiento de los brotes de gripe mediante la evaluación de las búsquedas en Google o las redes sociales son ejemplos de cómo esto puede ocurrir en las comunidades. Además, como muchos países poseen sistemas de declaración pasiva para diversas enfermedades comunicables, por lo general infecciosas, la declaración pasiva permite realizar comparaciones internacionales que pueden identificar regiones que necesitan ayuda urgente, confirmar nuevos casos y proporcionar intervenciones apropiadas para el control y el tratamiento.

La *vigilancia activa* hace referencia a un sistema en el que se contrata personal específicamente para que se encargue del proyecto del programa de vigilancia. Son contratados para realizar visitas de campo periódicas a instalaciones sanitarias, como clínicas, centros de salud de atención primaria y hospitales, para identificar casos nuevos de una o varias enfermedades o fallecidos debidos a la enfermedad (búsqueda de casos). La vigilancia activa puede consistir en realizar entrevistas a médicos y pacientes, revisar historiales médicos y, en los países en vías de desarrollo y en las áreas rurales, inspeccionar pueblos y ciudades para detectar casos bien periódicamente, de modo rutinario o una vez se haya comunicado un caso índice. La comunicación suele ser más exacta cuando la vigilancia es activa que cuando es pasiva, ya que la vigilancia activa es llevada a cabo por personal que ha sido contratado y formado específicamente para desempeñar esa misión.

Cuando se utiliza la vigilancia pasiva, a menudo se pide al personal existente (frecuentemente a los médicos) que comunique los casos nuevos. Sin embargo, a menudo se encuentran saturados por sus responsabilidades primarias de proporcionar asistencia sanitaria y administrar los servicios de

salud. Para ellos, rellenar los informes sobre los casos nuevos es una carga adicional que a menudo es considerada una tarea secundaria respecto a sus responsabilidades principales. Además, con la comunicación activa los brotes locales por lo general suelen ser identificados más fácilmente. Pero el mantenimiento de la comunicación activa es más caro que el de la pasiva y con frecuencia es más difícil de desarrollar inicialmente.

La vigilancia puede presentar problemas adicionales en los países en vías de desarrollo. Por ejemplo, las áreas que hay que vigilar pueden ser de difícil alcance y puede ser complicado mantener la comunicación entre dichas áreas y las autoridades centrales que deben tomar decisiones políticas y distribuir los recursos necesarios para el seguimiento y el control y la prevención de la enfermedad. Además, las definiciones de enfermedad utilizadas en los países desarrollados en ocasiones pueden ser inapropiadas o inutilizables en los países en vías de desarrollo debido a la falta de pruebas de laboratorio y de otros recursos sofisticados necesarios para una evaluación diagnóstica completa de los casos sospechosos. El resultado puede ser, por tanto, una infracomunicación de casos clínicos observados. En casos de epidemias de enfermedades, la Organización Mundial de la Salud (OMS) y varios países desarrollados, incluido Estados Unidos, movilizan recursos a los países en vías de desarrollo para ayudar a los funcionarios locales de salud pública en la búsqueda de casos y la recopilación de datos. Esto fue evidente en el brote de ébola en África occidental en 2014 y en la epidemia del virus Zika en 2015 en América Latina y el Caribe.

Un ejemplo de las dificultades en la vigilancia de las enfermedades utilizando datos de mortalidad es el problema de diferenciar las estimaciones de mortalidad por paludismo, uno de los mayores azotes en la actualidad, especialmente en países pobres en vías de desarrollo. En 2004 se produjo un repunte mundial en las muertes por paludismo. Desde entonces, los fallecimientos debidos al paludismo han disminuido de manera importante, especialmente en África subsahariana, lo que se ha atribuido a la expansión exitosa de actividades de control del vector, como el uso de mosquiteras tratadas con insecticidas para prevenir la infección y las mejoras del tratamiento de los infectados. Murray y cols. publicaron un análisis en 2014 en el que comunicaron que la carga global debido a la mortalidad por paludismo era de aproximadamente 854.000 muertes, una cifra alrededor de un 46% más elevada que la estimada en el World Malaria Report de 2014 de la OMS, que era de aproximadamente 584.000[3]. Esta disparidad en las estimaciones pone de manifiesto las dificultades a la hora de obtener datos fiables en ausencia de un sistema de vigilancia, un registro vital y pruebas diagnósticas estandarizados.

La vigilancia también puede llevarse a cabo para valorar cambios en los niveles de los factores de riesgo ambientales de la enfermedad. Por ejemplo, pueden realizarse estudios de la concentración de partículas contaminantes en el aire o de la radiación atmosférica, en especial tras la comunicación de un accidente. Un ejemplo único de esto es la explosión de la central nuclear de

Fukushima Daiichi en Fukushima, Japón, en 2011. Un terremoto de magnitud 9.0, seguido de un tsunami, inhabilitó los generadores de emergencia necesarios para el enfriamiento continuo, lo que finalmente terminó en un accidente nuclear con explosión química de hidrógeno-aire y liberación masiva de materiales radiactivos al medio ambiente[4]. Dicha vigilancia puede proporcionar signos de alerta temprana acerca de una posible elevación en la tasa de enfermedades asociadas con dicho agente ambiental. Por tanto, la vigilancia en busca de cambios en las tasas de las enfermedades o en los niveles de factores de riesgo ambientales puede servir de medida de la gravedad del accidente y apuntar a posibles soluciones para reducir dichos peligros en el futuro.

Etapas de la enfermedad en un individuo y en una población

Consideremos ahora los niveles de una enfermedad en una población en un periodo de tiempo y cómo se desplazan los individuos de un nivel de la enfermedad a otro en la población.

En la figura 3.1 se muestra la línea temporal del desarrollo de una enfermedad en un individuo. Un individuo se encuentra sano (es decir, sin enfermedad) y en algún momento tiene lugar el comienzo biológico de la enfermedad. La persona a menudo no es consciente del momento en el tiempo en el que comienza la enfermedad. Posteriormente aparecen los síntomas, lo que hace que el paciente busque asistencia médica. En ciertas situaciones, la hospitalización puede ser necesaria con fines diagnósticos y/o terapéuticos. En cualquier caso, en algún momento se realiza el diagnóstico y se inicia el tratamiento. A continuación, pueden producirse varios resultados: la curación, el control de la enfermedad, la discapacidad o la muerte. (Esto se examinará con mayor detalle en el capítulo 18, en la sección «La historia natural de la enfermedad».)

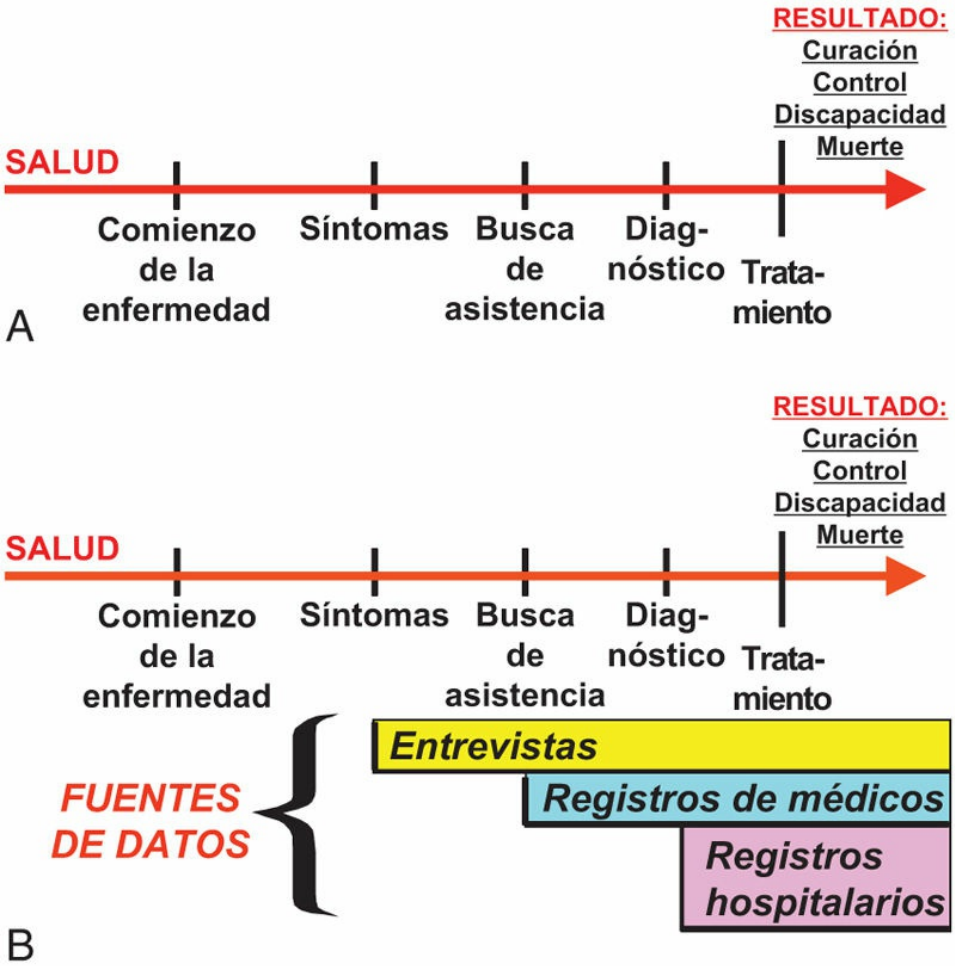


FIG. 3.1 (A) La historia natural de la enfermedad. (B) La historia natural de la enfermedad y algunas fuentes de datos relacionados con cada intervalo.

En la figura 3.2A-D se muestra la progresión de la enfermedad en una población en relación con los niveles de la enfermedad y la asistencia médica. El rectángulo exterior representa la población total (v. fig. 3.2A) y el rectángulo menor, el subgrupo más pequeño de enfermos (v. fig. 3.2B). A medida que una persona enferma, se desplaza desde el grupo de los enfermos al de los que buscan asistencia y al subgrupo de los hospitalizados, desde el rectángulo exterior hasta los rectángulos progresivamente más pequeños del diagrama, como muestran las flechas curvas (v. fig. 3.2C). Como se observa en la figura 3.2D, en todos estos rectángulos se producen fallecimientos, representados por las flechas rectas pequeñas, pero la tasa de mortalidad es proporcionalmente mayor en los grupos con enfermedades más graves, como en los hospitalizados.

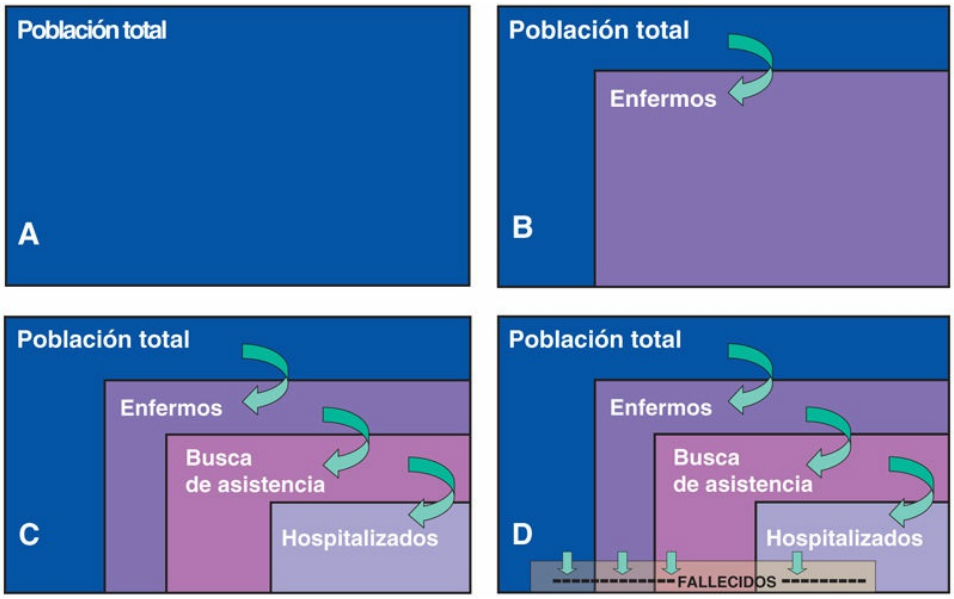


FIG. 3.2 (A-C) La población: progresión de salud a diversos grados de gravedad de la enfermedad. (D) La población: ocurrencia de muertes en cada grupo. (Modificado de White KL, Williams TF, Greenberg BG. The ecology of medical care. *N Engl J Med.* 1961;265:885–892.)

¿Qué fuentes de datos pueden utilizarse para obtener información acerca de la enfermedad de una persona? Durante el periodo de la enfermedad que precisa hospitalización, las historias clínicas y los registros hospitalarios resultan de utilidad (v. fig. 3.1B). Si la hospitalización no es necesaria, los registros del médico de atención primaria pueden ser la mejor fuente. Si queremos información acerca de la enfermedad incluso antes de que se busque asistencia médica, podemos obtener dicha información del paciente, con un cuestionario o una entrevista. Si el paciente no puede proporcionar esta información, podemos obtenerla de un familiar o de alguien que conozca el estado de salud del paciente. En esta figura no se muestran los registros de las aseguradoras médicas, que en ocasiones pueden proporcionar información muy útil.

La fuente de datos a partir de la cual se identifican los casos influye claramente sobre las tasas que calculamos para expresar la frecuencia de la enfermedad. Por ejemplo, los registros hospitalarios no incluyen datos sobre los pacientes que obtuvieron asistencia sanitaria en las consultas de los médicos. Por tanto, cuando valoramos tasas para la frecuencia de ocurrencia de una cierta enfermedad, debemos identificar las fuentes de los casos y determinar cómo fueron identificados los casos. Cuando interpretamos las tasas y las comparamos con las tasas comunicadas en otras poblaciones y en otros momentos del tiempo, debemos tener en cuenta las características de las fuentes a partir de las cuales se obtuvieron los datos.

La ocurrencia de la enfermedad puede medirse utilizando tasas o proporciones. Las *tasas* nos dicen la rapidez con la que una enfermedad se está produciendo en una población; las *proporciones* nos informan sobre la fracción de la población que está afectada. Veamos cómo utilizamos las tasas y las proporciones para expresar la extensión de una enfermedad en una

comunidad o en otra población. En este capítulo estudiaremos las medidas de la enfermedad o morbilidad; en el capítulo 4 se analizan las medidas de la mortalidad.

Medidas de la morbilidad

Tasa de incidencia

La tasa de incidencia de una enfermedad se define como el número de casos nuevos de una enfermedad que se producen durante un periodo de tiempo específico en una población con riesgo de desarrollar una enfermedad.

$$\text{Tasa de incidencia por } 1.000 = \frac{\text{N.º de casos nuevos de una enfermedad que se producen en la población durante un periodo de tiempo específico}}{\text{N.º de personas con riesgo de sufrir la enfermedad durante dicho periodo de tiempo}} \times 1.000$$

En esta tasa, el resultado se ha multiplicado por 1.000 para poder expresar la incidencia por 1.000 personas. La elección de esta cifra es más o menos arbitraria; podríamos haber empleado 10.000, 1 millón o cualquier otra cifra. Sin embargo, esta elección generalmente está influenciada por la frecuencia de la enfermedad; por ejemplo, para una enfermedad frecuente, como el resfriado común, la incidencia generalmente se define como un porcentaje; para las enfermedades raras, como la anemia aplásica, se multiplica por 100.000 o incluso por 1.000.000.

El elemento crítico en la definición de la tasa de incidencia es el de casos *NUEVOS* de la enfermedad. La tasa de incidencia es una medida de los acontecimientos; la enfermedad se identifica en una persona que desarrolla la enfermedad y que no padecía la enfermedad previamente. Como la tasa de incidencia es una medida de los acontecimientos (es decir, la transición desde un estado de salud a uno de enfermedad), es, por tanto, una medida del riesgo. El riesgo puede ser valorado en cualquier grupo de población, como en un grupo de edad particular, en los varones o en las mujeres, en un grupo profesional o en un grupo que haya estado expuesto a cierto agente ambiental, como la radiación o una toxina química. Por ejemplo, en la figura 3.3 se expone la tendencia en la incidencia del cáncer de tiroides en los niños de Bielorrusia, Ucrania y Rusia de 1986 a 1994, obtenida de datos de

vigilancia tras una explosión en el reactor de Chernóbil[5]. Las tasas de incidencia más elevadas se observaron en las áreas más contaminadas: Gómel en el sur de Bielorrusia y partes del norte de Ucrania. Sin embargo, un problema en la interpretación de estos datos es la posibilidad de que el aumento observado pudiera deberse al cribado intensivo que se inició tras el accidente. Dicho cribado podría haber identificado tumores tiroideos que de otro modo podrían no haber sido detectados y, por lo tanto, podrían no haber sido atribuidos a la exposición común (el reactor). Sin embargo, en la actualidad existe un consenso general acerca de que el aumento observado en cánceres tiroideos en niños y en adolescentes de áreas expuestas a la lluvia radiactiva de Chernóbil era, de hecho, real.

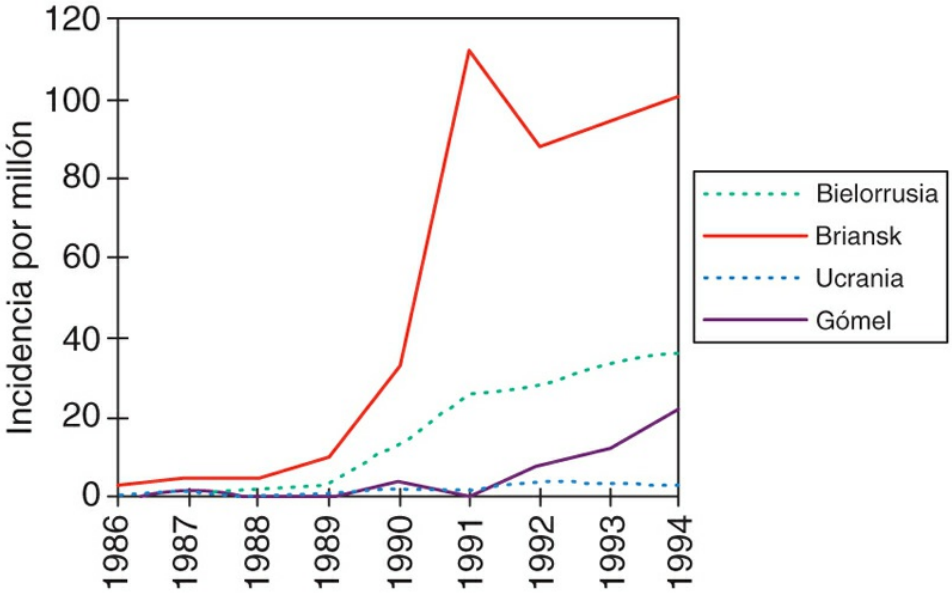


FIG. 3.3 Tendencias en la incidencia de cáncer de tiroides infantil en Bielorrusia, Ucrania y Rusia, 1986-94. (De Bard D, Verger P, Hubert P. Chernobyl, 10 years after: Health consequences. *Epidemiol Rev.* 1997;19:187–204.)

El denominador de la tasa de incidencia representa el número de personas con riesgo de desarrollar la enfermedad. Para que una tasa de incidencia sea significativa, todo individuo incluido en el denominador debe tener el potencial de convertirse en parte del grupo representado en el numerador. Por tanto, si calculamos la incidencia del cáncer uterino, el denominador debe incluir únicamente a mujeres sin antecedentes de histerectomía, porque las mujeres histerectomizadas y los varones no pueden ser nunca parte del grupo representado en el numerador, es decir, ambos grupos no tienen riesgo de padecer un cáncer de útero. Aunque este punto parece evidente, no siempre está tan claro, por lo que volveremos a este tema más adelante en la exposición.

Otro aspecto importante del denominador es el tiempo. Las medidas de la incidencia pueden utilizar dos tipos de denominadores: personas con riesgo que son observadas durante un periodo de tiempo definido; o, cuando toda la

población no es observada durante todo el periodo de tiempo, personas-tiempo (o unidades de tiempo durante las que se observa a una persona). Consideremos cada una de estas dos situaciones

Personas con riesgo que son observadas durante un periodo de tiempo definido

En el primer tipo de denominador de la tasa de incidencia, especificamos un periodo de tiempo y debemos conocer que todos los individuos del grupo representado por el denominador han sido seguidos durante *dicho periodo completo*. La elección del periodo de tiempo es arbitraria: podríamos calcular la incidencia en 1 semana, 1 mes, 1 año, 5 años, etc. Lo importante es que, con independencia del periodo de tiempo utilizado en el cálculo, este debe ser especificado claramente, y todos los individuos incluidos en el cálculo deben haber sido observados y, por supuesto, con riesgo de sufrir el desenlace de interés durante todo el periodo. La incidencia calculada utilizando un periodo de tiempo durante el cual se considera que todos los individuos de la población tienen riesgo respecto del resultado se denomina también *proporción de incidencia acumulada,* que es una medida del riesgo.

Cuando todas las personas no son observadas durante todo el periodo de tiempo, personas-tiempo o unidades de tiempo durante las que se observa a una persona

A menudo, todos los individuos del denominador no pueden ser seguidos durante todo el tiempo especificado por diversos motivos, como la pérdida de seguimiento o el fallecimiento por una causa diferente a la que se está estudiando. Cuando diferentes individuos son observados durante diferentes periodos de tiempo, calculamos una tasa de incidencia (denominada también *densidad de incidencia*), en la que el denominador consiste en la suma de las unidades de tiempo en las que el individuo se encontraba en riesgo y era observado. Esto se denomina *personas-tiempo* y a menudo se expresa en términos de personas-meses o personas-años (pa) de observación.

Consideremos el concepto pa: una persona con riesgo que es observada durante un año = 1 pa. Una persona con riesgo que es observada durante 5 años = 5 pa. Y 5 personas con riesgo, cada una de las cuales es observada durante solo 1 año, también es 5 pa.

Asumamos que tenemos un estudio de 5 años y cinco personas han sido observadas durante todo el periodo (indicado por una flecha para cada persona en la fig. 3.4). En cada uno de los 5 años del estudio, los cinco participantes son observados, por lo que tenemos 5 pa de observación en cada uno de los 5 años, logrando un total de 25 pa de observación en todo el estudio.

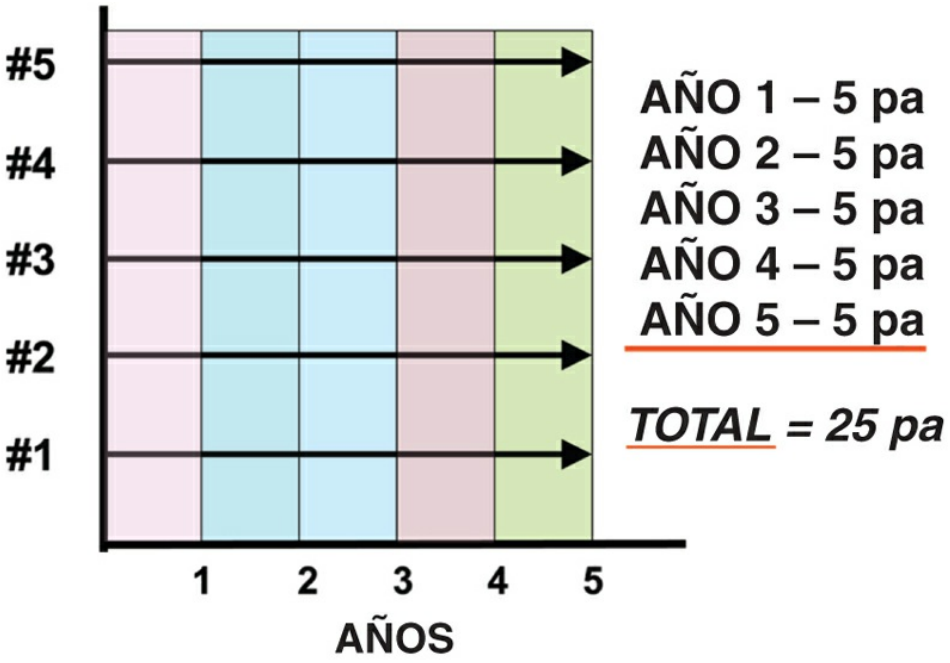
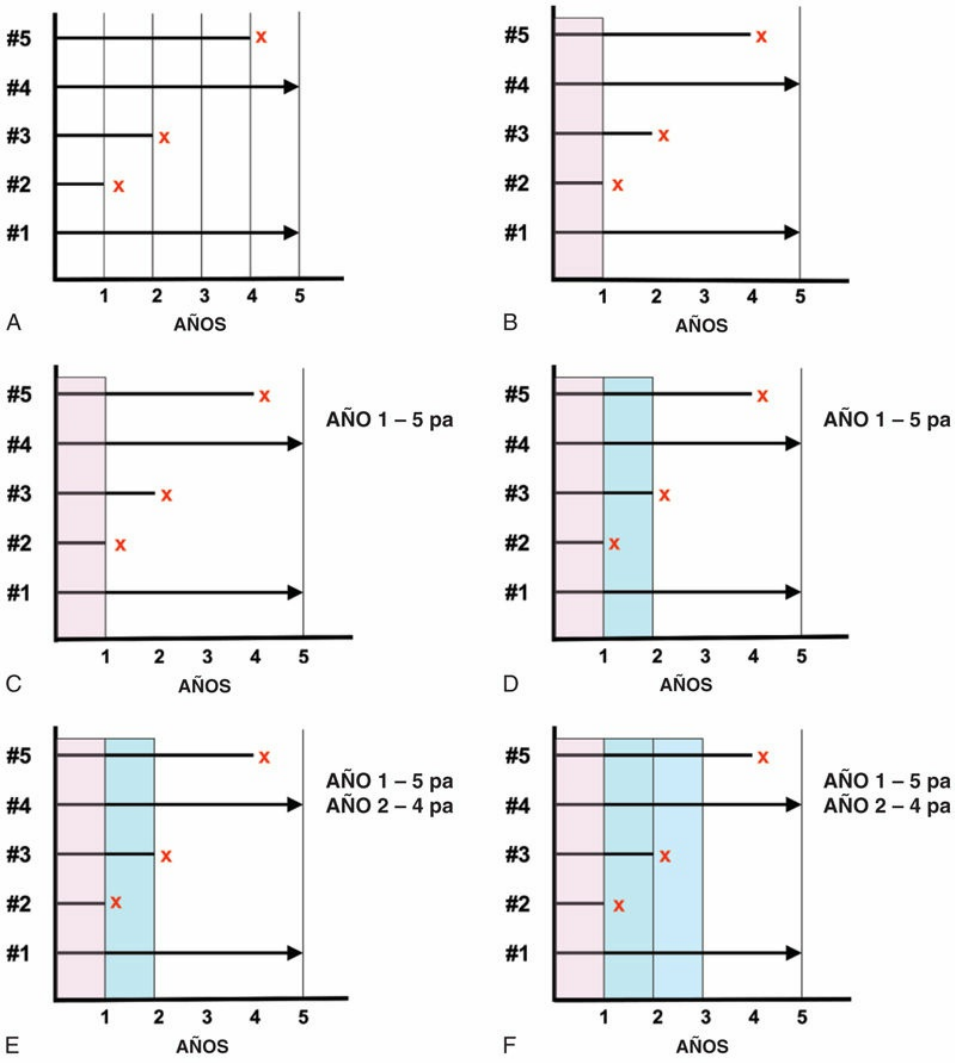


FIG. 3.4 Cuando todos los miembros de la población estudiada son observados durante todo el periodo: personas-años *(pa)* de observación.

Ahora consideremos la situación en la que las cinco personas con riesgo no son observadas durante los 5 años del estudio, sino que son observadas durante diferentes periodos de tiempo (fig. 3.5A). En este diagrama, las dos flechas representan a dos personas que fueron observadas durante todos los 5 años. Las líneas temporales de las otras tres personas finalizan con una «x» roja, que indica el punto en el que terminó la observación de cada individuo, bien porque el acontecimiento de interés ocurrió, o porque se perdió el seguimiento de la persona o por otros problemas.


#5
#4
#3
#2
#1
1 2 3 4 5
AÑOS
A
B
C
D
E
F
AÑO 1 – 5 pa
AÑO 2 – 4 pa

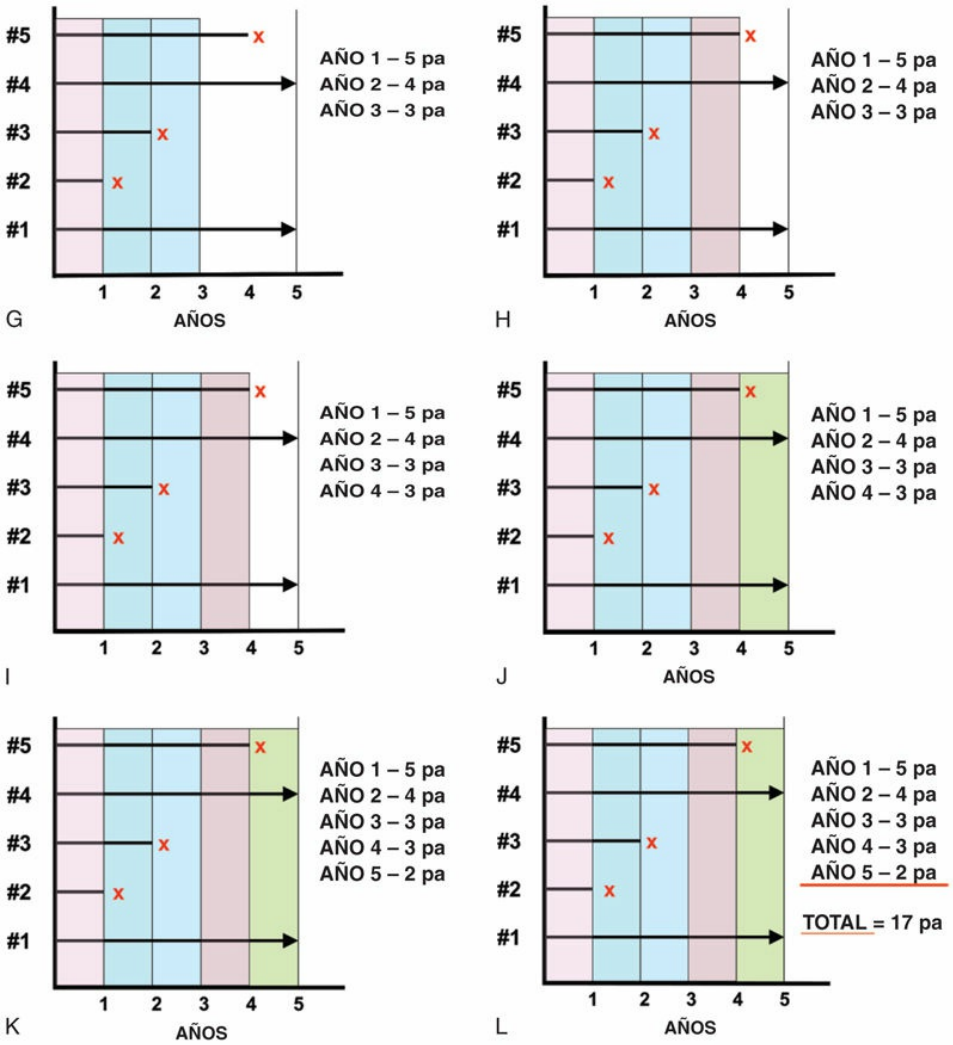


FIG. 3.5 (A-L) Pero ¿qué pasa si las personas en riesgo de la población son observadas durante diferentes periodos de tiempo? Cálculo del concepto personas-tiempo como personas-años *(pa)* observados. (V. explicación en pág. 47 del texto.)

¿Cómo calculamos el número total de pa observados en este estudio? Miremos el primer año del estudio (v. fig. 3.5B). Todas las cinco personas fueron observadas durante el primer año, por lo que tenemos 5 pa de observación en el primer año (v. fig. 3.5C).

Fijémonos a continuación en el segundo año del estudio (v. fig. 3.5D). El participante n.° 2 solo fue observado durante el primer año, por lo que en el segundo año solo tenemos cuatro participantes, cada uno de los cuales contribuyó en 1 año de seguimiento al estudio, para un total de 4 pa (v. fig. 3.5E).

Si observamos el tercer año del estudio, veremos que el participante n.° 3 fue observado únicamente los 2 primeros años del estudio (v. fig. 3.5F). Por tanto, tan solo tres participantes fueron observados en el tercer año, lo que genera 3 pa de observación durante el tercer año (v. fig. 3.5G). Todos estos participantes también fueron observados el cuarto año del estudio (v. fig. 3.5H) y de nuevo contribuyeron con 3 pa de observación durante el cuarto año del estudio (v. fig. 3.5I).

Por último, fijémonos en el quinto año del estudio (v. fig. 3.5J). Observamos que el participante n.° 5 solo fue observado durante los primeros 4 años del estudio. Como resultado, tan solo quedaron dos participantes que fueran observados en el quinto año del estudio. Ellos contribuyeron a 2 pa de observación durante el quinto año (v. fig. 3.5K). Como se observa en la figura 3.5L, durante los 5 años de estudio se consiguieron 5 + 4 + 3 + 3 + 2 pa de observación, lo que hace un total de 17 pa de observación. (Si todos los 5 participantes hubiesen sido observados durante los 5 años del estudio, se habrían conseguido 25 pa de observación, como muestra la fig. 3.4.) Por tanto, *si las personas en riesgo son observadas durante diferentes periodos de tiempo,* la tasa de incidencia es:

$$\text{Tasa de incidencia por } 1.000 = \frac{\text{N.º de casos NUEVOS de una enfermedad que se producen en la población durante un periodo de tiempo específico}}{\text{Personas-tiempo total (la suma de los periodos de tiempo de observación de cada persona que ha sido observada durante todo el periodo de tiempo o parte del mismo)}} \times 1.000$$

En el capítulo 6 se analiza con más detalle el concepto personas-tiempo.

Identificación de casos nuevos para calcular la incidencia

En términos prácticos, cuando deseamos calcular la incidencia, ¿cómo identificamos a todos los casos nuevos en una población durante un periodo de tiempo específico? En ciertas situaciones puede ser posible vigilar a toda la población a lo largo del tiempo con pruebas que pueden detectar casos de nueva aparición de una enfermedad. Sin embargo, esto a menudo no es posible y, en su lugar, se identifica una población y se realizan pruebas de cribado para la enfermedad a nivel basal (los casos prevalentes se definen en la siguiente sección) (fig. 3.6). Aquellos que no presentan la enfermedad a nivel basal son seguidos durante el tiempo especificado (p. ej., un año). Posteriormente se vuelven a examinar para detectar si han desarrollado la

enfermedad en cuestión (fig. 3.7). Todos los casos identificados claramente han desarrollado la enfermedad durante el periodo de 1 año, ya que los que están siendo observados no presentaban la enfermedad al principio del año. Por tanto, estos casos son casos nuevos o incidentes y sirven como numerador de la tasa de incidencia.

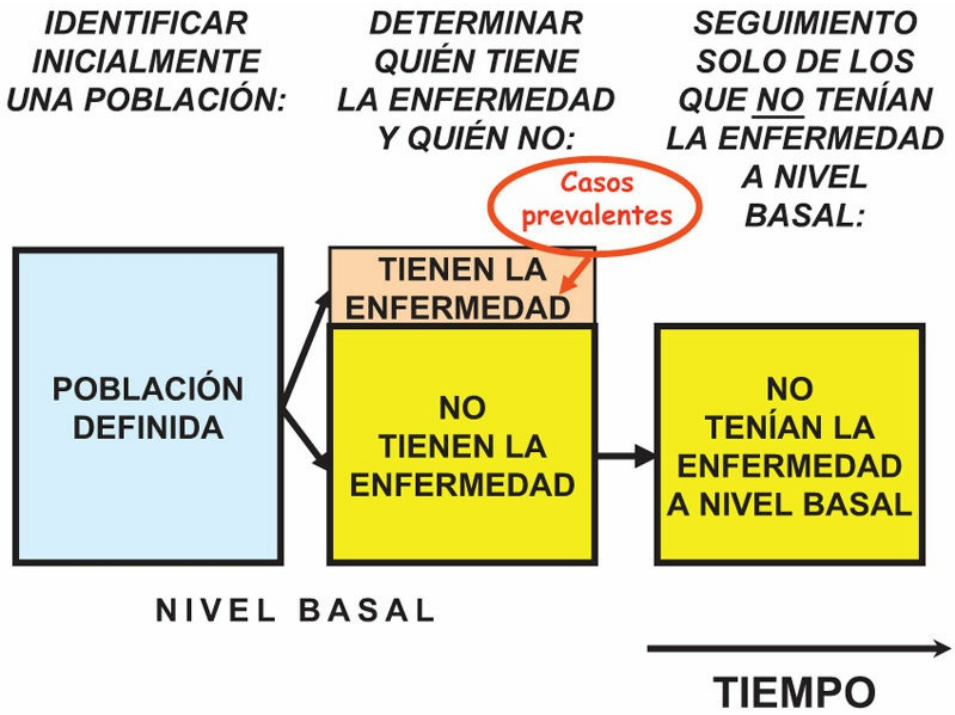


FIG. 3.6 Identificación de casos recién detectados de una enfermedad. Paso 1: cribado de casos prevalentes a nivel basal. La explicación se expone en la página 50 del texto.

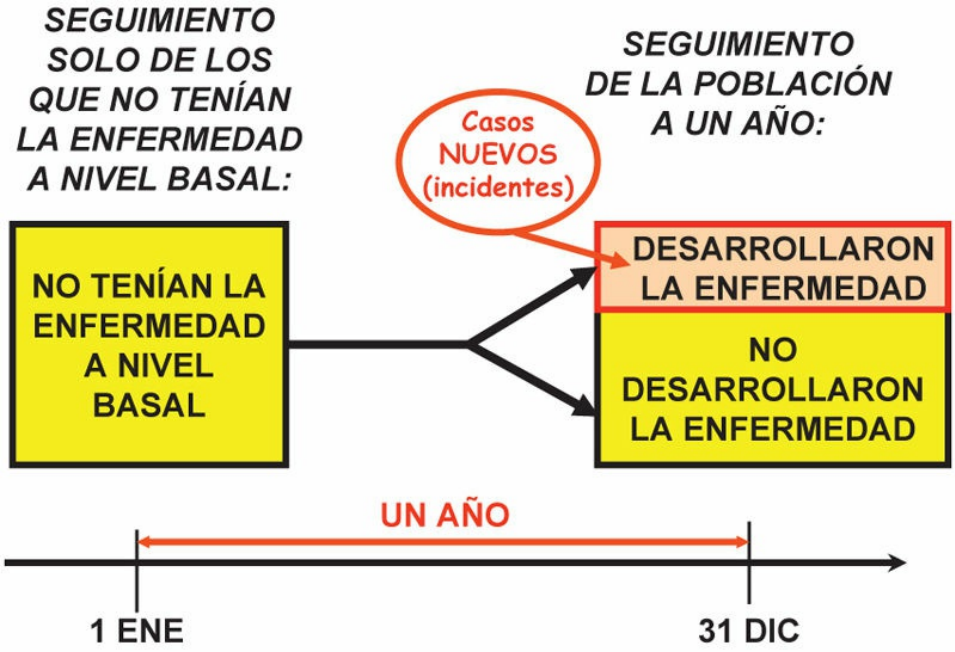


FIG. 3.7 Identificación de casos recién detectados de una enfermedad. Paso 2: seguimiento y nuevo cribado al año para identificar casos que aparecieron durante el año.

Aunque en la mayoría de las situaciones es necesario expresar la incidencia especificando un denominador, en ocasiones el número de casos por sí solo puede ser informativo. Por ejemplo, en la figura 3.8 se expone el número de casos de TB comunicados en Estados Unidos de 1982 a 2015. El número de casos comunicados en un año en Estados Unidos (desde que comenzó la comunicación de casos) alcanzó el mínimo histórico en 2015. A pesar de una pequeña disminución desde 1980 hasta 1985, el número comunicado de casos de TB aumentó aproximadamente un 20% entre 1985 y 1992. Gran parte del aumento de la TB expuesto aquí se asoció con la identificación en rápida expansión de infecciones por el VIH. Sin embargo, incluso antes de que el SIDA y el VIH fuesen reconocidos como problemas de salud pública importantes, la TB seguía siendo un problema grave, a menudo desatendido, especialmente en ciertas áreas urbanas de Estados Unidos. Desde 1992 hasta 2008, el número total de casos de TB disminuyó del 2% al 7% anual. Este es un ejemplo en el que un gráfico que solo representa los números de casos sin un denominador puede ser muy útil cuando no existen motivos para sospechar un cambio significativo en el denominador durante un periodo de tiempo determinado.

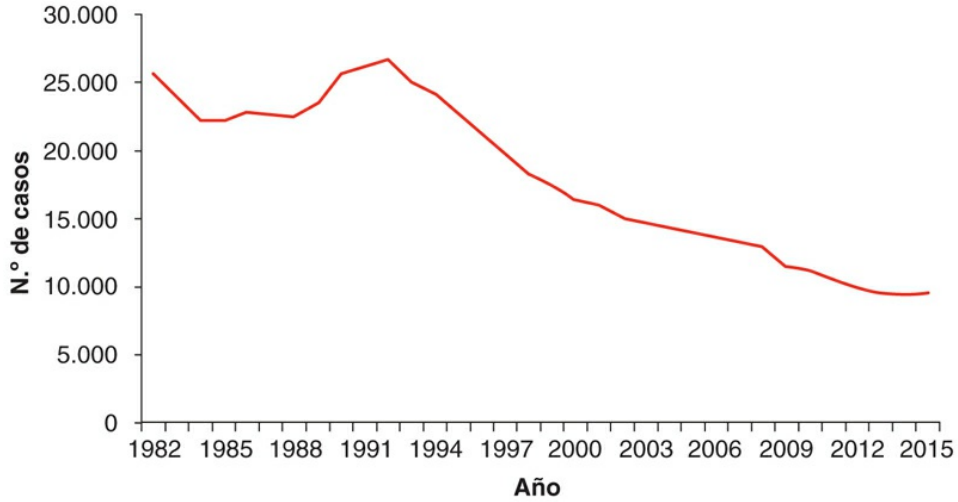


FIG. 3.8 Casos notificados de tuberculosis, Estados Unidos, 1982-2015 (a 9 de junio de 2016). (De Centers for Disease Control and Prevention [CDC]. *Reported Tuberculosis in the United States, 2015.* Atlanta, GA: US Department of Health and Human Services, CDC; 2016. Consultado el 15 de mayo de 2017.)

En general, sin embargo, nuestro objetivo a la hora de calcular la incidencia es ser capaces de hacerlo con la información necesaria tanto para el numerador como para el denominador, de modo que podamos realizar comparaciones válidas. En la figura 3.9 se presenta la tendencia en la tasa de incidencia de cánceres seleccionados en Estados Unidos en varones *(izquierda)* y en mujeres *(derecha)* desde 1975 hasta 2013. Como se observa aquí, la incidencia de cáncer de pulmón ha estado disminuyendo en los varones y estabilizándose en las mujeres. Tras importantes aumentos en la incidencia durante muchos años, el cáncer de próstata en los varones ha estado disminuyendo desde 2001. El cáncer de mama en las mujeres en Estados Unidos ha disminuido entre 1998 y 2003, seguido de un ligero aumento desde 2004 hasta 2013. Tras haberse mantenido estable durante varios años, el cáncer de colon y de recto ha estado disminuyendo tanto en varones como en mujeres.

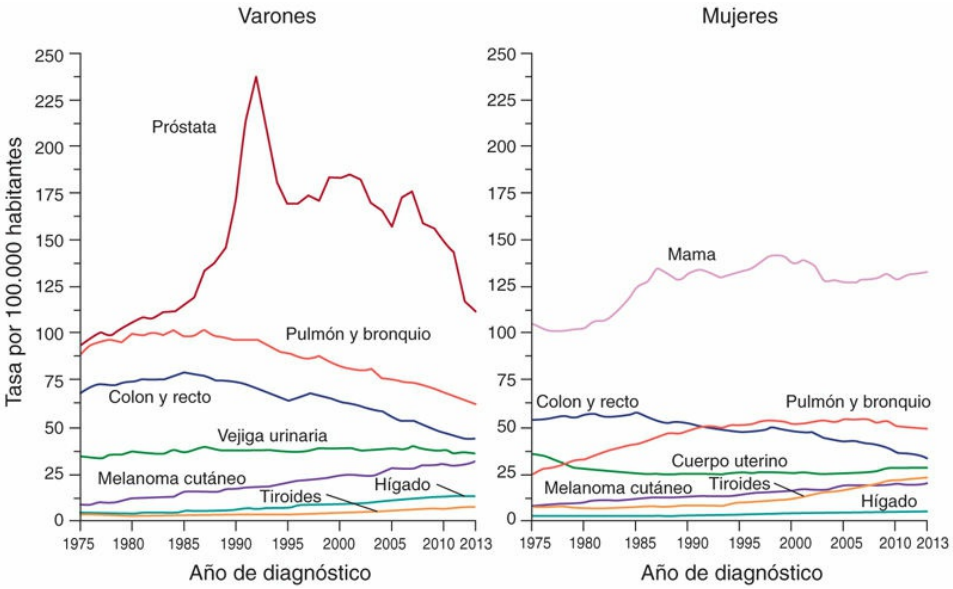


FIG. 3.9 Tasas de incidencia anual de cáncer ajustadas por edad entre varones y mujeres para cánceres seleccionados, Estados Unidos, 1975-2013 (ajustadas por edad respecto a la población estándar de Estados Unidos del año 2000). (De Siegel R, Miller K, Jemal A. Cancer statistics, 2017. *CA Cancer J Clin.* 2017;67:7–30.)

Tasa de ataque

En ocasiones, el tiempo asociado con el denominador puede ser especificado implícitamente en vez de explícitamente. Por ejemplo, en el capítulo 2 expusimos la investigación de un brote de una enfermedad de origen alimentario, en el que hablamos de la *tasa de ataque*, que se define como el número de personas expuestas a un alimento sospechoso que enfermaron dividido por el número de personas expuestas a dicho alimento. La tasa de ataque no especifica explícitamente el intervalo de tiempo porque sabemos que en muchos brotes de enfermedades de origen alimentario la mayoría de los casos se producen a las pocas horas o unos pocos días después de la exposición. Por tanto, los casos que se presentan meses después no son considerados parte del mismo brote. Sin embargo, en muchas situaciones, el conocimiento actual de la biología y la historia natural de la enfermedad no define claramente una franja temporal, por lo que el tiempo debe especificarse explícitamente. Una consideración más es que la tasa de ataque no es verdaderamente una tasa, sino una proporción. Como en el término *tasa* se incluye una unidad de tiempo en el cálculo, una tasa de ataque de origen alimentario no es un término preciso, ya que realmente nos informa sobre la *proporción* de todas las personas que comieron un cierto alimento que enfermaron. Sin embargo, el término *tasa de ataque* se ha utilizado tradicionalmente durante mucho tiempo. Más adelante seguiremos analizando el uso de proporciones en la evaluación de la ocurrencia de una enfermedad.

Prevalencia

La prevalencia se define como el número de personas afectadas presentes en la población en un momento específico dividido por el número de personas en la población en dicho momento, es decir, qué proporción de la población presenta la enfermedad en dicho momento.

$$\text{Prevalencia por } 1.000 = \frac{\text{N.º de casos de una enfermedad presentes en la población en un momento específico}}{\text{N.º de personas en la población en dicho momento específico}} \times 1.000$$

Por ejemplo, si estamos interesados en conocer la prevalencia de la artritis en una comunidad determinada en una fecha concreta, podríamos visitar todas las casas en dicha comunidad y, por medio de entrevistas o de exploraciones físicas, determinar cuántas personas tienen artritis ese día. Este número sería el numerador de la prevalencia. El denominador sería la población de la comunidad en dicha fecha.

¿Cuál es la diferencia entre *incidencia* y *prevalencia*? La prevalencia puede considerarse como una instantánea o una rebanada de la población en un momento en el tiempo en el que determinamos quién tiene la enfermedad y quién no. Pero haciendo esto no estamos determinando cuándo apareció la enfermedad. Algunos individuos pueden haber desarrollado la artritis ayer, otros la semana pasada, otros el último año y algunos hace 10 o 20 años. Por tanto, cuando estudiamos una comunidad para estimar la prevalencia de una enfermedad, por lo general no tenemos en cuenta la duración de la enfermedad. Por tanto, el numerador de la prevalencia incluye una mezcla de personas con diferentes duraciones de la enfermedad, y, como resultado, no tenemos una medida del riesgo. Si queremos evaluar el riesgo, debemos emplear la incidencia, porque, a diferencia de la prevalencia, incluye solo acontecimientos o casos nuevos y un periodo de tiempo específico durante el que se produjeron esos acontecimientos.

En las publicaciones médicas y sobre salud pública, el término *prevalencia* a menudo se usa de dos formas:

- *Prevalencia puntual.* La prevalencia de la enfermedad en un momento dado del tiempo. Este es el uso del término *prevalencia* que acabamos de exponer.

- *Prevalencia de periodo.* ¿Cuántas personas han presentado la enfermedad en cualquier momento durante un cierto periodo de tiempo? El periodo de tiempo al que se hace referencia puede elegirse arbitrariamente (p. ej., un mes, un año o un periodo de 5 años). Algunas personas pueden haber desarrollado la enfermedad durante dicho periodo y otras puede que ya tuvieran antes la enfermedad y fallecieron o se curaron durante dicho periodo. La cuestión importante es que toda persona representada en el numerador presentaba la enfermedad en algún momento durante el periodo especificado.

En la tabla 3.1 se exponen los dos tipos de prevalencia, así como la incidencia acumulada, empleando preguntas sobre el asma.

Tabla 3.1

Ejemplos de prevalencia puntual y de periodo y de incidencia acumulada en estudios de asma con entrevistas

Pregunta de la entrevista	Tipo de medida
¿Tiene asma en la actualidad?	Prevalencia puntual
¿Ha tenido asma en los últimos [*n*] años?	Prevalencia de periodo
¿Ha tenido asma alguna vez?	Incidencia acumulada

Volviendo a la prevalencia puntual, en términos prácticos, es prácticamente imposible estudiar toda una ciudad en un solo día. Por tanto, aunque conceptualmente estamos pensando en términos de un solo momento en el tiempo, en realidad, el estudio llevaría mucho más tiempo. Cuando vemos la palabra *prevalencia* utilizada sin ningún término que la modifique, generalmente se refiere a la prevalencia puntual; para el resto del capítulo, utilizaremos el término *prevalencia* con el significado de prevalencia puntual.

Consideremos la incidencia y la prevalencia. En la figura 3.10 se exponen cinco casos de una enfermedad en una comunidad en 2017. El primer caso de la enfermedad se produjo en 2016 y el paciente falleció en 2017.

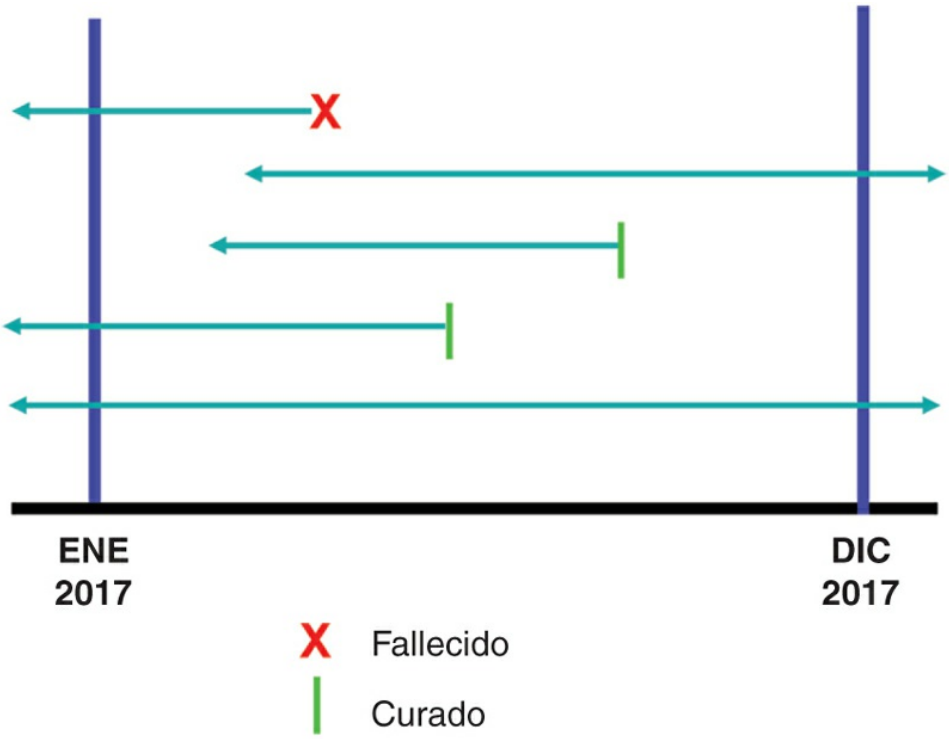


FIG. 3.10 Ejemplo de incidencia y prevalencia: I.

El segundo caso se produjo en 2017 y continuó hasta 2018. El tercer caso se trataba de una persona que enfermó en 2017 y se curó en 2017. El cuarto caso se produjo en 2016 y el paciente se curó en 2017. El quinto caso ocurrió en 2016 y continuó a lo largo de 2017 y 2018.

En este ejemplo, consideraremos únicamente los casos (numeradores) e ignoraremos los denominadores. ¿Cuál es el numerador para la incidencia en 2017 en este ejemplo? Sabemos que la incidencia solo tiene en cuenta los casos nuevos y, como dos de los cinco casos se produjeron en 2017, el numerador para la incidencia en 2017 es 2.

¿Cuál sería el numerador para la prevalencia puntual en 2017? Esto depende de cuándo se haga el estudio de prevalencia (fig. 3.11). Si realizamos el estudio en mayo, el numerador será 5. Si realizamos el estudio en julio, el numerador será 4. Sin embargo, si realizamos el estudio en septiembre, el numerador será 3, y si lo realizamos en diciembre, será 2. Por tanto, la prevalencia dependerá del momento del año en el que se realice el estudio.

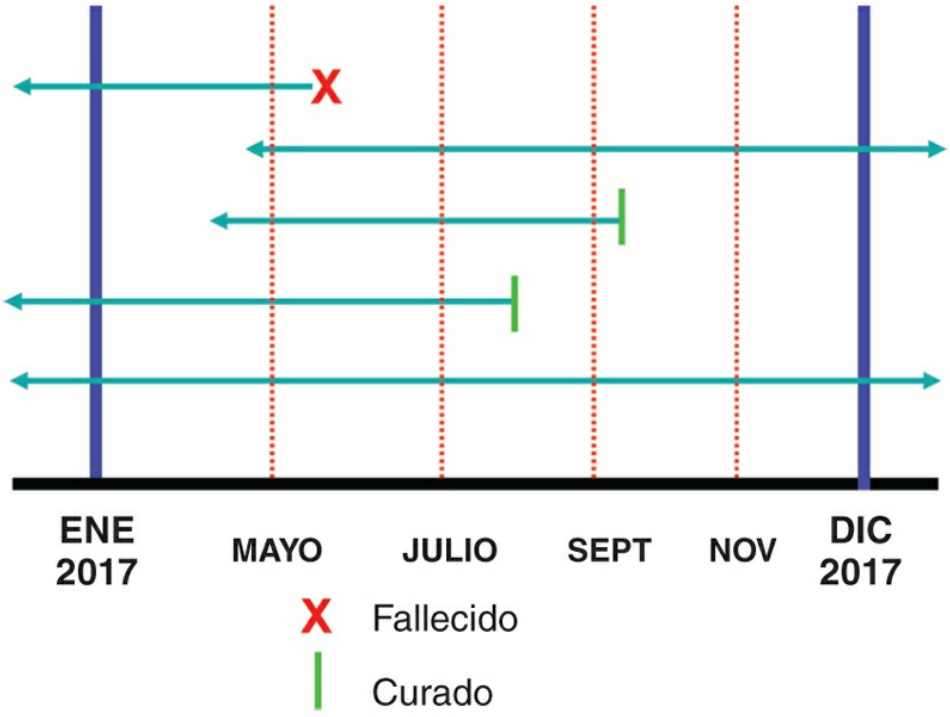


FIG. 3.11 Ejemplo de incidencia y prevalencia: II.

En la figura 3.12A-D se expone la relación dinámica entre la incidencia y la prevalencia. El matraz representa una comunidad (v. fig. 3.12A) y las cuentas en el matraz representan los casos prevalentes de una enfermedad en la comunidad. ¿Cómo podemos añadir casos o aumentar la prevalencia? Como se observa en la figura 3.12B, podemos hacerlo a través de la incidencia: añadiendo casos nuevos. ¿Qué pasaría si pudiéramos eliminar cuentas del matraz y disminuir la prevalencia? ¿Cómo podríamos hacerlo? Como se observa en la figura 3.12C, podría conseguirse si se produjesen fallecimientos o curaciones. Claramente, estos dos resultados suponen una gran diferencia para el paciente; sin embargo, con respecto a la prevalencia, las curaciones y las muertes producen el mismo efecto: reducen el número de personas enfermas en la población y, por tanto, disminuyen la prevalencia. Así, lo que se produce es la situación dinámica expuesta en la figura 3.12D. La adición continua de casos nuevos (incidencia) aumenta la prevalencia, mientras que los fallecimientos o las curaciones la reducen.

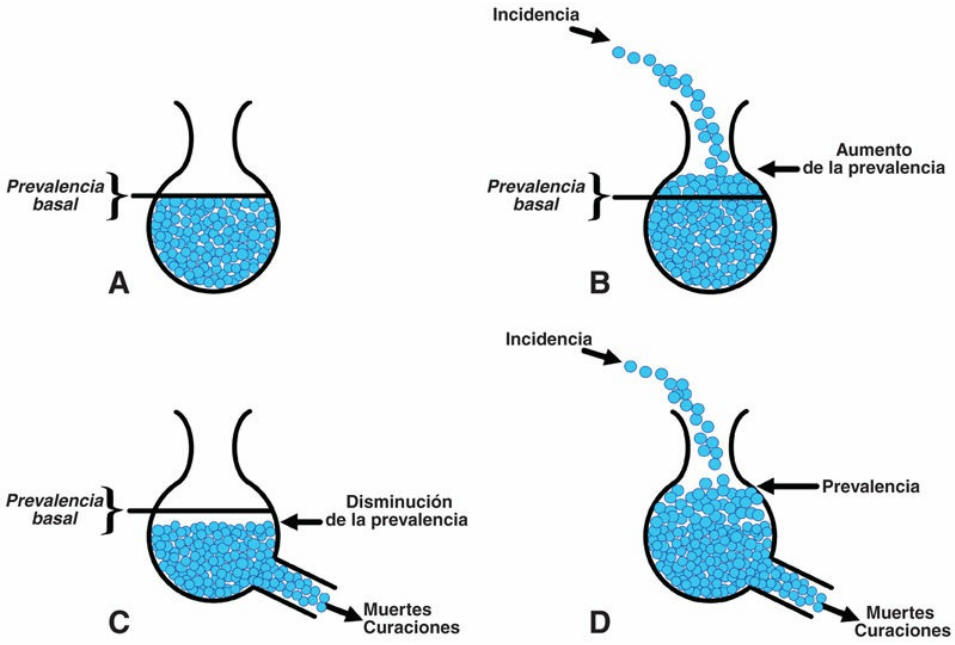


FIG. 3.12 Relación entre la incidencia y la prevalencia. (A) Nivel de prevalencia en la población. (B) Aumento de la prevalencia debido a la incidencia. (C) Disminución de la prevalencia debido a muertes y/o curaciones. (D) Impacto global sobre la prevalencia de la incidencia, las muertes y/o las curaciones.

Este efecto de disminución de la prevalencia debido a los fallecimientos o las curaciones es un aspecto importante en la salud pública y en la medicina clínica. Por ejemplo, cuando se dispuso de insulina por primera vez, ¿qué ocurrió con la prevalencia de la diabetes? La prevalencia aumentó porque la diabetes no se curaba, sino que únicamente se controlaba. Muchos pacientes diabéticos que antes habrían fallecido ahora sobrevivían, por lo que la prevalencia aumentó. Esta paradoja aparente a menudo se presenta en los programas de salud pública: tras introducir una nueva intervención de asistencia sanitaria (p. ej., el tratamiento antirretroviral de gran actividad para la infección por el VIH) aumenta la supervivencia (menos pacientes fallecen por VIH/SIDA) o contribuye a la detección precoz de la enfermedad en más personas, y el efecto neto es un aumento aparente de la prevalencia. Puede ser difícil convencer a algunas personas de que un programa es exitoso si en realidad la prevalencia de la enfermedad objetivo del programa aumenta. Sin embargo, esto es justo lo que ocurre cuando se evita la muerte y la enfermedad no se cura o no se erradica.

Hemos afirmado que la prevalencia no es una medida del riesgo. Entonces, ¿por qué molestarnos en estimar la prevalencia? La prevalencia es una medida importante y útil de la carga de una enfermedad para la comunidad, que proporciona información a los responsables de la toma de decisiones sobre la asignación de recursos. Por ejemplo, ¿cuántas personas en la comunidad tienen artrosis? Esta información podría ayudarnos a determinar, por ejemplo, cuántas clínicas son necesarias, qué tipos de servicios de rehabilitación son necesarios y cuántos profesionales sanitarios son necesarios y de qué tipo. La prevalencia es, por tanto, valiosa para planificar servicios

sanitarios. Cuando utilizamos la prevalencia también queremos hacer proyecciones futuras y anticiparnos a los cambios que probablemente tengan lugar en la carga de la enfermedad. Sin embargo, si queremos encontrar la causa o etiología de la enfermedad, debemos explorar la relación entre la exposición y el riesgo de la enfermedad, y, para hacer esto, necesitamos datos sobre la incidencia.

No obstante, los datos sobre la prevalencia en ocasiones pueden ser muy útiles; pueden ser sugestivos, si no confirmatorios, en estudios sobre la etiología de ciertas enfermedades. Por ejemplo, el asma es una enfermedad de niños en la que es difícil medir su incidencia porque el momento exacto de comienzo de la enfermedad (inicio) a menudo es difícil de establecer, dadas las dificultades para definir la enfermedad y determinar los síntomas iniciales. Por este motivo, cuando estamos interesados en las tendencias temporales y en la distribución geográfica del asma, la prevalencia es la medida empleada con mayor frecuencia. La información sobre la prevalencia del asma a menudo se obtiene de declaraciones del paciente, como entrevistas o cuestionarios.

La figura 3.13 muestra la prevalencia ajustada de enfermedad renal crónica (ERC) en estadios 3 y 4 en adultos de Estados Unidos, por presencia o ausencia de diabetes, de la Encuesta nacional de examen de salud y nutrición (NHANES, *National Health and Nutrition Examination Survey*) 1988–94 a 2011–12; la NHANES realiza estudios transversales anuales con muestras representativas de la población general de Estados Unidos. La ERC se definió como una tasa de filtración glomerular estimada (FGe) de 15 a 59 ml/min/1,73 m^2, estimada con la ecuación de la colaboración epidemiológica de la enfermedad renal crónica (EPI-ERC) a partir de mediciones calibradas únicas de creatinina sérica. Podemos observar un aumento inicial en la prevalencia ajustada de las etapas 3 y 4 de la ERC, que se detuvo alrededor de la década de 2000 entre los individuos no diabéticos, mientras que la prevalencia continuó aumentando en los diabéticos. Esto podría explicarse en parte por la mayor supervivencia de los pacientes diabéticos con ERC.

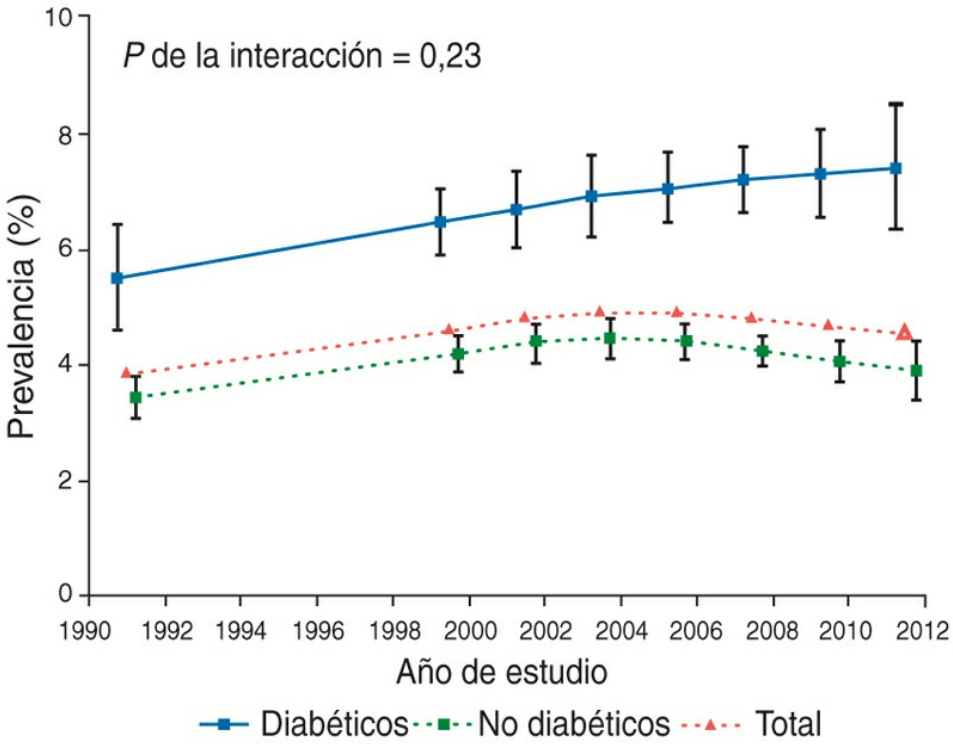


FIG. 3.13 Prevalencia ajustada de la enfermedad renal crónica en estadios 3 y 4 (tasa estimada de filtración glomerular de 15 a 59 ml/min/1,73 m^2 calculada con la ecuación de la Chronic Kidney Disease Epidemiology Collaboration) en adultos de Estados Unidos, por edad, 1990–2012. (De Murphy D, McCulloch CE, Lin F, et al. Centers for Disease Control and Prevention Chronic Kidney Disease Surveillance Team. Trends in prevalence of chronic kidney disease in the United States. *Ann Intern Med.* 2016;165:473–481.)

Otro ejemplo del valor de los datos de prevalencia se expone en la figura 3.14. Uno de los problemas de salud pública más importantes y desafiantes en la actualidad en Estados Unidos y en otros países desarrollados es la prevalencia cada vez mayor de la obesidad. La obesidad se asocia con una morbilidad y mortalidad importantes y es un factor de riesgo para enfermedades como la artritis, la hipertensión, la diabetes tipo 2, la enfermedad coronaria y el accidente cerebrovascular. En esta figura, en la que se expone la prevalencia autoinformada de la obesidad por estado para 2015, ningún estado comunicó una prevalencia inferior al 20% y 44 presentaban una prevalencia igual o superior al 25%. También podemos observar que la prevalencia más alta de obesidad se encuentra en los estados del sur.

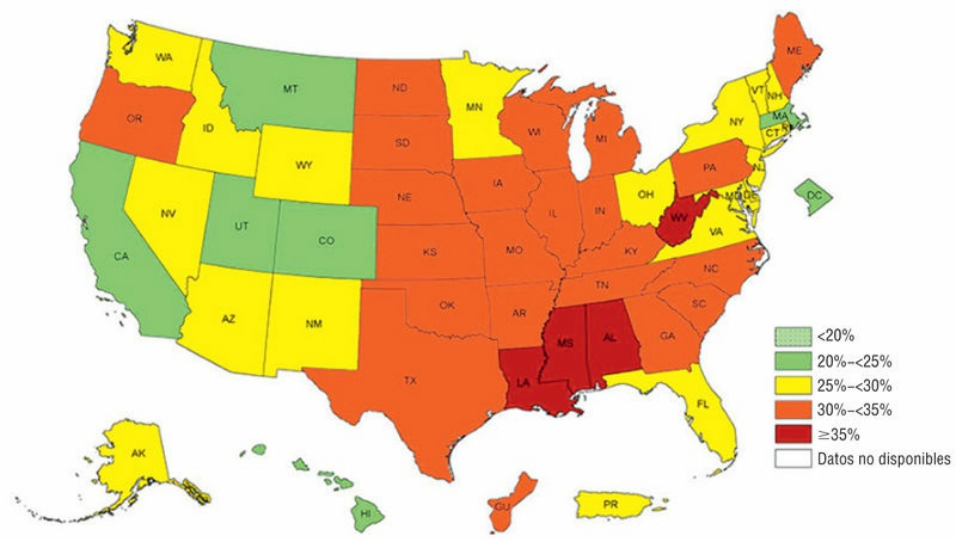


FIG. 3.14 Prevalencia de obesidad autoinformada entre adultos de Estados Unidos, por estado y territorio, Sistema de vigilancia de factores de riesgo conductuales, 2015. La obesidad se definió por un índice de masa corporal ≥30 o ~13,6 kg de sobrepeso para una persona de 1,63 m. (Modificado de Centers for Disease Control and Prevention. https://www.cdc.gov.ezp.welch.jhmi.edu/obesity/data/prevalence-maps.html. Consultado el 16 de mayo de 2017.)

Una limitación de estos datos (anteriores y posteriores a 2011) es que están basados en alturas y pesos comunicados por los que responden por teléfono. Se ha observado que los que responden encuestas, especialmente en las encuestas telefónicas sobre obesidad, por lo general infravaloran su peso, exageran su altura o ambos. En este estudio, los participantes fueron clasificados en función de su índice de masa corporal (IMC), que se define como el peso de una persona en kilogramos dividido por el cuadrado de la estatura en metros (IMC = peso [kg]/estatura^2 [metros^2]). Un IMC de 25 o superior se clasifica como sobrepeso y un IMC igual o superior a 30 se considera obesidad. El resultado probablemente es una infraestimación de la prevalencia de la obesidad si nos basamos en el IMC, por lo que la verdadera prevalencia de la obesidad por estado probablemente sea más elevada que la que se expone en la figura 3.14. Teniendo en cuenta la tendencia descrita anteriormente y la observada en la figura 3.14, será necesario un gran esfuerzo de salud pública y un gran compromiso para solucionar este problema de salud pública que sigue empeorando ininterrumpidamente. Además, el uso aislado del IMC para definir la obesidad tiene sus propias limitaciones. El IMC no distingue entre el exceso de grasa, músculo o masa ósea, y no proporciona ninguna información sobre la distribución de la grasa en cada individuo.

En el cuadro 3.1 se enumeran algunas fuentes posibles de datos estadísticos sobre la morbilidad. Cada una posee sus limitaciones, debido principalmente a que la mayoría de estas fuentes no están diseñadas con fines de investigación, sino más bien con fines administrativos o de facturación. Por tanto, pueden caracterizarse por datos incompletos o ambiguos y, en ocasiones, puede que solo se refieran a poblaciones muy seleccionadas que

pueden no ser representativas de la población en la que quisiéramos generalizar los hallazgos.

Cuadro 3.1 Algunas fuentes de datos estadísticos sobre la morbilidad

1. Declaración de enfermedades: enfermedades contagiosas, registros de cáncer.
2. Datos acumulados como resultado de tramitación de seguros y planes de asistencia médica con prepago:
    a. Seguros sanitarios y de accidentes de grupos.
    b. Planes de asistencia médica prepago.
    c. Planes de seguros de discapacidad estatales.
    d. Compañías de seguros de vida.
    e. Planes de seguros hospitalarios: Blue Cross.
    f. Railroad Retirement Board.
3. Asistencia pública y planes de asistencia médica financiados con impuestos:
    a. Asistencia pública, ayuda a ciegos, ayuda a discapacitados.
    b. Planes de asistencia médica estatales o federales.
4. Hospitales y clínicas.
5. Registros de absentismo: industria y colegios.
6. Exploraciones físicas previas al empleo o periódicas en la industria y los colegios.
7. Programas de búsquedas de casos.
8. Registros del personal militar:
    a. Fuerzas armadas.
    b. Administraciones de veteranos.
9. Estudios de morbilidad en muestras de población (p. ej., National Health Survey, National Cancer Surveys).

Problemas con las medidas de la incidencia y la prevalencia

Problemas con los numeradores

El primer problema es definir quién tiene la enfermedad. Algunas enfermedades son difíciles de diagnosticar y, cuando surge dicha dificultad diagnóstica, con frecuencia se convoca a grupos de expertos para crear criterios diagnósticos. Hay cinco conjuntos de criterios diagnósticos para la bacteriemia por *Staphylococcus aureus* asociada a la asistencia sanitaria (AAS), cada uno con diferentes niveles de rigurosidad (tabla 3.2). La figura 3.15 muestra la proporción de pacientes clasificados como que presentan bacteriemia por *S. aureus* (BSA) AAS de acuerdo con las cinco definiciones

diferentes en 2.638 pacientes empleando datos del norte de Dinamarca. Se observa que la estimación de la prevalencia se ve muy afectada por el grupo de criterios que se utilice.

Tabla 3.2

Definiciones 1-5 de la bacteriemia por *Staphylococcus aureus* asociada a la asistencia sanitaria

Nivel más alto de rigurosidad	Definición	Criterios:
		Hemocultivo realizado dentro de los 2 días del ingreso y lo siguiente:
	1.	• Cualquier ingreso hospitalario en los últimos 30 días
	2.	• Cualquier ingreso hospitalario en los últimos 30 días o • Contacto hospitalario ambulatorio que incluye cirugía o contacto con clínicas de oncología, hematología o nefrología en los últimos 30 días
Nivel más bajo de rigurosidad	3.	• Cualquier ingreso hospitalario en los últimos 30 días • Cualquier tipo de contacto hospitalario ambulatorio en los últimos 30 días
	4.	• Cualquier ingreso hospitalario en los últimos 90 días o • Cualquier tipo de contacto hospitalario ambulatorio en los últimos 30 días
	5.	• Cualquier ingreso hospitalario en los últimos 90 días o • Cualquier tipo de contacto hospitalario ambulatorio en los últimos 30 días o • Terapia antibiótica o inmunosupresora 30 días antes del ingreso

Modificada de Smit J, Søgaard M, Schønheyder HC, et al. Classification of healthcare-associated *Staphylococcus aureus* bacteremia: influence of different definitions on prevalence, patient characteristics, and outcome. *Infect Control Hosp Epidemiol*. 2016;37(2):208–211.

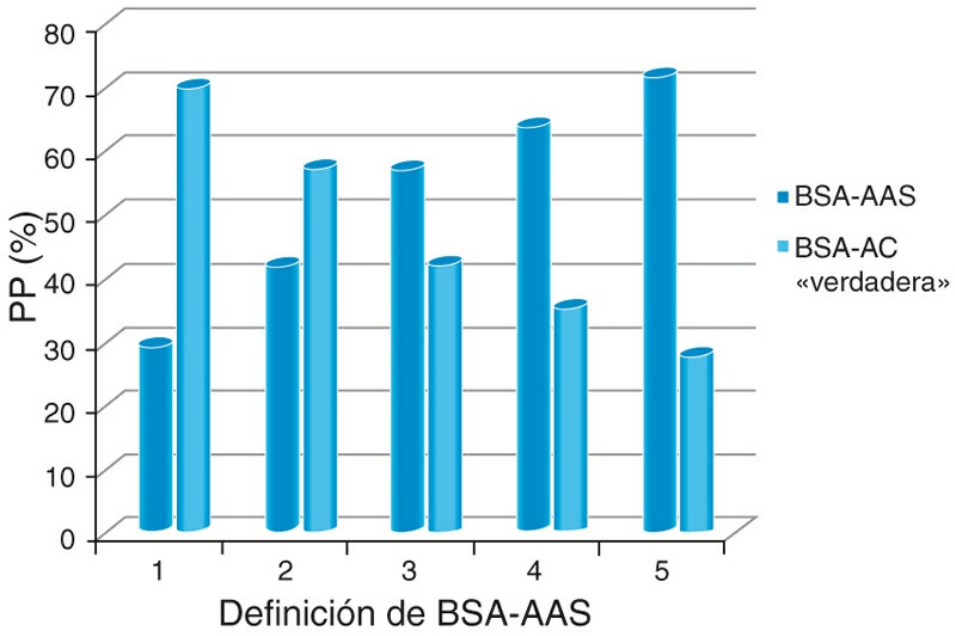


FIG. 3.15 Proporción de prevalencia *(PP)* de pacientes diagnosticados de bacteriemia por *Staphylococcus aureus* asociada a la asistencia sanitaria *(BSA-AAS)* y BSA adquirida en la comunidad *(AC)* verdadera según las cinco definiciones diferentes. Modificado de Smit J, Søgaard M, Schønheyder HC, et al. Classification of healthcare-associated *Staphylococcus aureus* bacteremia: influence of different definitions on prevalence, patient characteristics, and

Otro ejemplo es una cohorte de 1.879 varones y mujeres de 65 años o más que fueron incluidos en el Canadian Study of Health and Aging (CSHA)[6]. Se calculó la proporción de individuos diagnosticados de demencia utilizando seis sistemas de clasificación empleados comúnmente. En función del sistema diagnóstico utilizado, la proporción de individuos con demencia variaba del 3,1% al 29,1% (fig. 3.16). Esta gran variación en la estimación de la prevalencia posee implicaciones importantes tanto para la investigación como para la provisión de los servicios sanitarios apropiados. Cuando se comunican los resultados de cualquier estudio sobre morbilidad, resulta fundamental que se especifique claramente la definición precisa utilizada para un caso. La decisión acerca de qué definición se debe usar no siempre es sencilla. A menudo depende en gran medida del propósito específico para el que el estudio ha sido realizado.

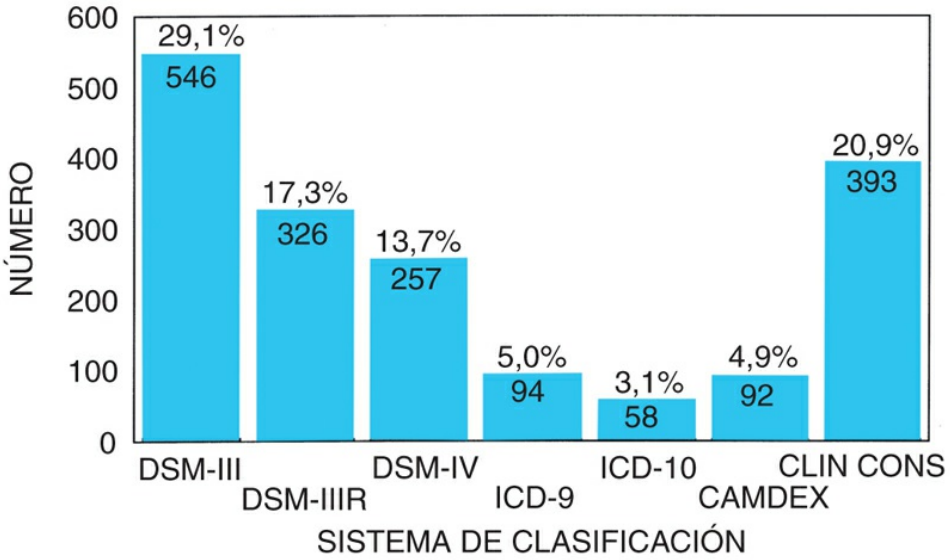


FIG. 3.16 Número de personas con demencia y prevalencia (%) en la cohorte del Canadian Study of Health and Aging (*n* = 1.879), diagnosticada con diferentes sistemas de clasificación. Las diferentes abreviaturas se refieren a manuales diagnósticos utilizados comúnmente para trastornos médicos. *CAMDEX,* Cambridge Mental Disorders of the Elderly Examination; *CLIN CONS,* Clinical Consensus; *DSM,* Diagnostic and Statistical Manual of Mental Disorders; *ICD,* International Classification of Diseases. (Datos de Erkinjuntti T, Østbye T, Steenhuis R, Hachinski V. The effect of different diagnostic criteria on the prevalence of dementia. *N Engl J Med.* 1997;337:1667–1674.)

El siguiente aspecto relacionado con los numeradores es el de determinar qué personas deberían ser incluidas en el numerador. ¿Cómo encontramos los casos? Podemos utilizar datos disponibles habitualmente o, como se ha expuesto anteriormente en este capítulo, podemos realizar un estudio diseñado específicamente para obtener datos para estimar la incidencia o la prevalencia. En muchos de estos estudios, los datos se obtienen a partir de

entrevistas; algunas de las posibles limitaciones de los datos obtenidos a través de entrevistas se exponen en el cuadro 3.2. Lo ideal sería contar con pruebas de laboratorio u otras pruebas confirmatorias. Sin embargo, a menudo no se dispone de estas pruebas, y, a pesar de estas limitaciones, los datos obtenidos a través de entrevistas son muy valiosos para proporcionar información acerca de los casos nuevos.

Cuadro 3.2 Algunas posibles fuentes de error en los estudios mediante entrevistas

1. Problemas debidos a dificultades en el diagnóstico:
    a. El que responde puede tener la enfermedad, pero puede encontrarse asintomático y no ser consciente de la enfermedad.
    b. El que responde puede tener la enfermedad, puede haber presentado síntomas, pero puede que no haya solicitado asistencia médica, por lo que puede que no conozca el nombre de la enfermedad.
    c. El que responde puede tener la enfermedad, puede haber recibido atención médica, pero puede que no se haya llegado al diagnóstico o no se le haya transmitido o puede haberlo entendido incorrectamente.
    d. El que responde puede no recordar con precisión un episodio de enfermedad o hechos y exposiciones relacionadas con la enfermedad.
2. Problemas asociados con el que responde en el estudio:
    a. El que responde puede estar implicado en litigios acerca de la enfermedad y puede elegir no responder o puede alterar su respuesta.
    b. El que responde puede ser reacio a proporcionar información exacta si tiene dudas de que ciertas respuestas pueden no ser del agrado del entrevistador o pueden suscitar un posible estigma.
    c. El que responde se encuentra demasiado enfermo como para poder responder. Como resultado, el que responde no es incluido en el estudio o se entrevista a un representante, como un familiar o un amigo. Los representantes, sin embargo, a menudo poseen información incompleta acerca de las exposiciones pasadas del entrevistado.
3. Problemas asociados con el entrevistador:
    a. El que responde puede proporcionar información, pero el entrevistador puede no registrarla o hacerlo de forma incorrecta.
    b. El entrevistador puede no plantear la pregunta que debe

preguntar o puede que la plantee incorrectamente.
c. El entrevistador puede estar sesgado al conocer la hipótesis que se quiere estudiar y puede interrogar a un grupo de participantes con mayor intensidad que a otro.

Problemas con los denominadores

Hay muchos factores que pueden afectar a los denominadores utilizados. Puede producirse un recuento selectivo insuficiente de ciertos grupos de la población. Por ejemplo, en muchos recuentos de la población se ha pasado por alto a los varones jóvenes de grupos étnicos minoritarios y a inmigrantes recientes. Con frecuencia queremos determinar si un cierto grupo posee un riesgo de sufrir una enfermedad superior al esperado, de modo que puedan dirigirse las medidas preventivas apropiadas a dicho grupo. Por tanto, nos interesan las tasas de enfermedad en diferentes grupos étnicos en vez de en toda la población en conjunto. Sin embargo, existen diferentes formas de clasificar a la población por grupo étnico, como el idioma, el país de origen, la herencia o el grupo étnico de los progenitores. Cuando distintos estudios emplean definiciones diferentes, la comparación de los resultados es difícil. Lo más importante en cualquier estudio es que la definición de trabajo se establezca con claridad, de modo que el lector pueda juzgar si los resultados son realmente comparables.

En una sección anterior hemos afirmado que, para que una tasa tenga sentido, toda la población del grupo representado en el denominador debe tener la posibilidad de pasar al grupo representado en el numerador. Este aspecto no es sencillo. Por ejemplo, la histerectomía es una de las intervenciones quirúrgicas realizadas más comúnmente en Estados Unidos. Este punto suscita dudas acerca de las tasas de mortalidad por cáncer cervical o endometrial. Si en el denominador incluimos a mujeres que han sido sometidas a una histerectomía, claramente ellas no presentan riesgo de padecer un cáncer cervical o endometrial. En la figura 3.17 se exponen las tasas de mortalidad por cáncer cervical en Estados Unidos; corregidas y no corregidas en función de si las mujeres estaban histerectomizadas o no. Observamos que las tasas corregidas son más elevadas. ¿Por qué? Porque en las tasas corregidas se han eliminado del denominador a las mujeres que habían sido sometidas a una histerectomía. Por tanto, el denominador se reduce y la tasa aumenta. Sin embargo, en este caso la tendencia a lo largo del tiempo no varía significativamente ya utilicemos las tasas corregidas o las no corregidas.

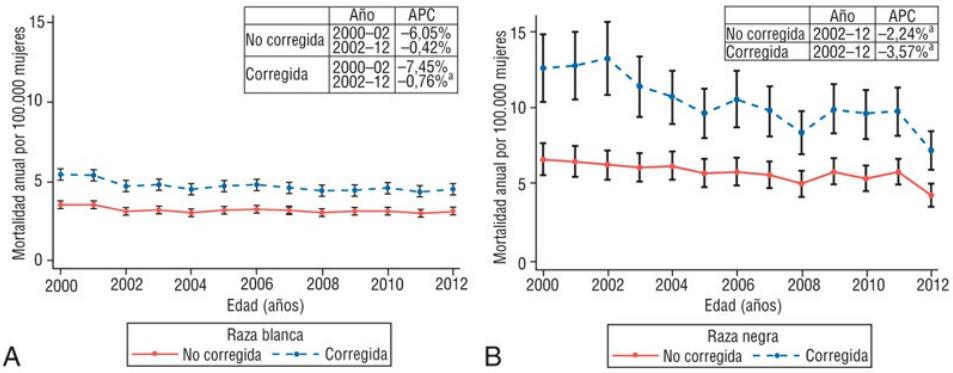


FIG. 3.17 Tendencias en las tasas de mortalidad por cáncer de cuello uterino estandarizadas por edad, sin corregir y corregidas por la prevalencia de histerectomía, desde 2000 a 2012 para (A) mujeres blancas y (B) negras. [a]El valor de la variación porcentual anual (APC, *annual percentage change*) *P* fue significativo para α <0,05. (De Beavis AL, Gravitt PE, Rositch AF. Hysterectomy-corrected cervical cancer mortality rates reveal a larger racial disparity in the United States. *Cancer.* 2017;123:1044–1050.)

Problemas con los datos hospitalarios

Los datos de los registros hospitalarios (en la actualidad a menudo registros médicos electrónicos) son una de las fuentes de información más importantes en los estudios epidemiológicos. Sin embargo, en el cuadro 3.3 se exponen algunos de los problemas que surgen al utilizar datos hospitalarios con fines de investigación. En primer lugar, los ingresos hospitalarios son selectivos. Pueden ser selectivos debido a las características del personal, la gravedad de la enfermedad, los trastornos médicos asociados y las normas de admisión, que varían de un hospital a otro. En segundo lugar, los registros hospitalarios no están diseñados para la investigación sino para la asistencia del paciente. Los registros pueden ser incompletos, ilegibles o pueden haberse perdido. La calidad diagnóstica de los registros hospitalarios, de los registros médicos y de los servicios clínicos puede variar. Por tanto, si queremos combinar pacientes de diferentes hospitales, podemos tener problemas de comparabilidad. En tercer lugar, si queremos calcular tasas, tenemos un problema para definir los denominadores, porque la mayoría de los hospitales estadounidenses carecen de zonas de captación, es decir, áreas geográficas que exijan que todas las personas de dichas áreas se hospitalicen en un hospital particular, y que ninguno de fuera del área de influencia sea ingresado en ese hospital.

Cuadro 3.3 Algunas limitaciones de los datos hospitalarios

1. Los ingresos hospitalarios son selectivos en relación con:
    a. Las características del personal.
    b. La gravedad de la enfermedad.

c. Los trastornos asociados.
d. Las políticas de ingreso.

2. Los registros hospitalarios no están diseñados para la investigación. Pueden ser:
   a. Incompletos, ilegibles o pueden haberse perdido.
   b. De calidad diagnóstica variable.
3. La(s) población(es) en riesgo (denominador) por lo general está(n) mal definida(s).

En clave de humor, en el cuadro 3.4 se muestran algunas notas que fueron dictadas por los médicos para su inclusión en las historias clínicas de los pacientes.

Cuadro 3.4 Algunas notas dictadas por médicos para su inclusión en las historias clínicas de los pacientes

«El paciente tiene dos hijos adolescentes, pero ninguna otra anomalía.»
«Al segundo día la rodilla se encontraba mejor y al tercer día había desaparecido por completo.»
«El paciente se encontraba alerta e inconsciente.»
«Cuando la paciente se mareó sus ojos rodaron por la habitación.»
«La exploración rectal puso de manifiesto una glándula tiroides de tamaño normal.»
«Cuando fue ingresado, su corazón se había parado y se encontraba mejor.»

Relación entre la incidencia y la prevalencia

Hemos dicho que la incidencia es una medida del riesgo, no así la prevalencia, ya que esta última no tiene en cuenta la duración de la enfermedad. Sin embargo, entre la incidencia y la prevalencia existe una relación importante: en una situación estable, en la que las tasas no cambian y las entradas son iguales a las salidas, y cuando la prevalencia no es demasiado elevada, es de aplicación la siguiente ecuación:

$$\text{Prevalencia} = \text{Incidencia} \times \text{Duración de la enfermedad}$$

Esto se demuestra en el siguiente ejemplo hipotético. Mediante radiografías torácicas se realizó un cribado de TB en 2.000 personas: 1.000 son habitantes con ingresos altos de Hitown y 1.000 son habitantes con ingresos bajos de Lowtown (tabla 3.3). Los hallazgos radiológicos fueron positivos en 100 de los habitantes de Hitown y en 60 de los habitantes de Lowtown. ¿Podemos concluir, por tanto, que el riesgo de TB es más elevado en los habitantes de

Hitown que en los de Lowtown? Claramente no, ya que lo que estamos midiendo con una radiografía de tórax es la prevalencia puntual de la enfermedad; desconocemos desde hace cuánto tiempo tienen la enfermedad las personas con hallazgos radiológicos positivos (tabla 3.4). De hecho, podemos considerar un escenario hipotético que podría explicar la prevalencia más alta en los habitantes de Hitown que no se relaciona con un riesgo más elevado en dicha población (tabla 3.5). Hemos expuesto que la prevalencia = incidencia × duración. Asumamos que la población de Lowtown tiene un riesgo mucho más elevado (incidencia) de TB que la población de Hitown: 20 casos/año en los habitantes de Lowtown en comparación con 4 casos/año en los habitantes de Hitown. Pero, por diversos motivos, como el peor acceso a la asistencia sanitaria y el peor estado nutricional, los habitantes de Lowtown sobreviven con sus enfermedades, de media, durante únicamente 3 años, mientras que los habitantes de Hitown sobreviven, de media, 25 años. En este ejemplo, la prevalencia es, por tanto, más alta en la población de Hitown que en la de Lowtown no porque el riesgo de la enfermedad sea más alto en los habitantes de Hitown sino porque la supervivencia de los afectados en Hitown es más prolongada; así, la prevalencia de la enfermedad (incidencia × duración) es más elevada en los habitantes de Hitown que en los de Lowtown.

Tabla 3.3

Ejemplo hipotético de cribado mediante radiografía torácica: I. Poblaciones estudiadas y número de radiografías positivas

Población estudiada	Número con radiografías positivas
1.000 Hitown	100
1.000 Lowtown	60

Tabla 3.4

Ejemplo hipotético de cribado mediante radiografía torácica: II. Prevalencia puntual

Población estudiada	Número con radiografías positivas	Prevalencia puntual por 100.000 habitantes
1.000 Hitown	100	100
1.000 Lowtown	60	60

Tabla 3.5

Ejemplo hipotético de cribado mediante radiografía torácica: III. Prevalencia, incidencia y duración

Población estudiada	Prevalencia puntual por 100.000 habitantes	Incidencia (ocurrencia/año)	Duración (años)
Hitown	100	4	25
Lowtown	60	20	3

En la figura 3.18 se muestra el porcentaje de todos los nacimientos extramaritales en Nueva Zelanda, desde 1962 a 1979. Este hecho preocupó mucho por el aparente aumento estable de los nacimientos extramaritales. Sin embargo, como se observa en la figura 3.19, realmente no se había producido un aumento en la tasa de nacimientos extramaritales; se había producido una disminución de los nacimientos totales, debido principalmente a la disminución de los nacimientos en las mujeres casadas. Como resultado, los nacimientos extramaritales representaban un porcentaje mayor del total de nacimientos, aunque la tasa de nacimientos extramaritales no hubiera aumentado a lo largo del periodo de 17 años.

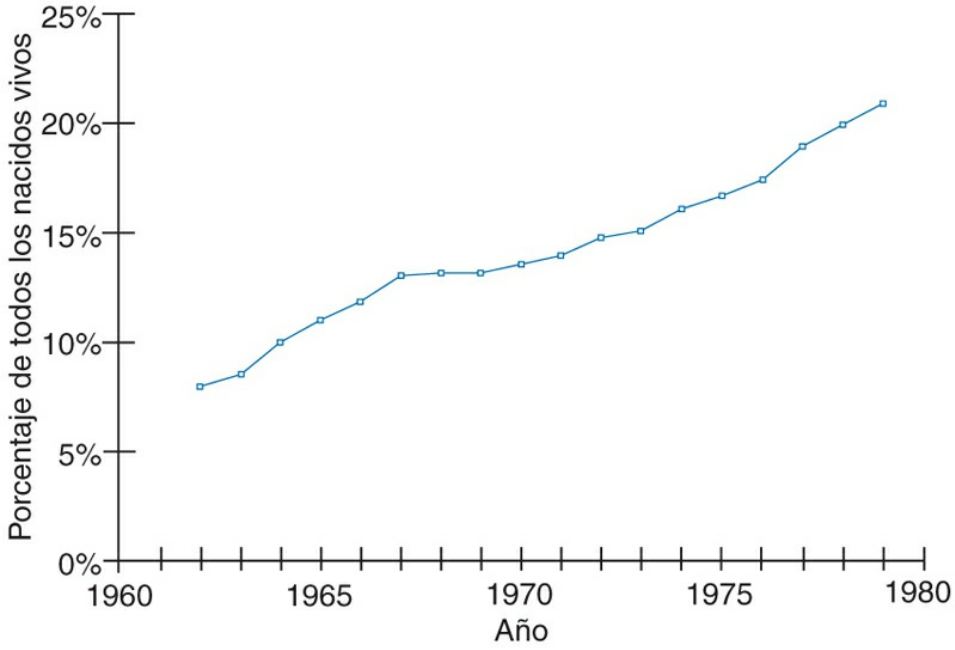


FIG. 3.18 Porcentaje de nacimientos extramaritales en Nueva Zelanda, 1962-79, según datos del departamento de estadística. (Modificado de Benfield J, Kjellstrom T. New Zealand ex-nuptial births and domestic purposes benefits in a different perspective.*N Z Nurs J.* 1981;74:28–31.)

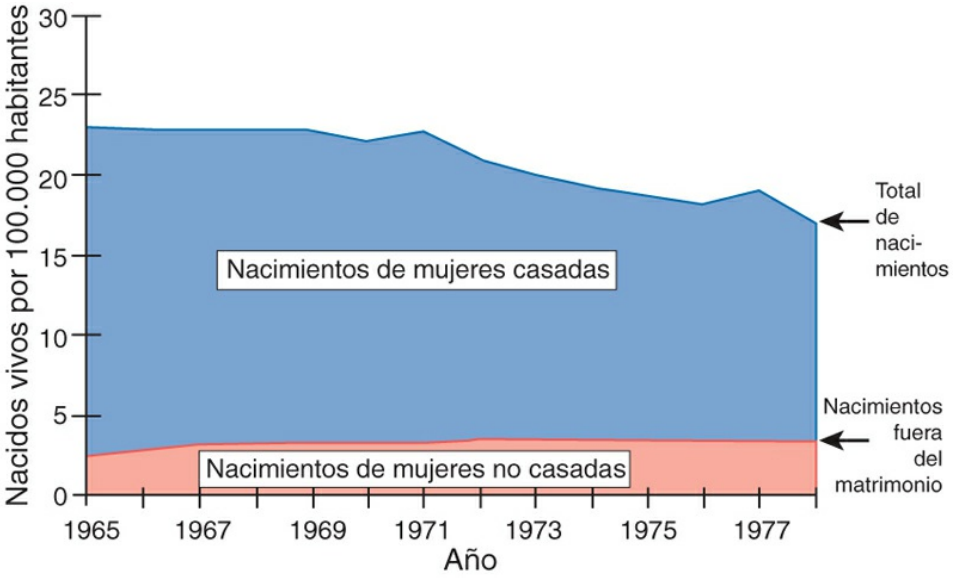


FIG. 3.19 Nacimientos de mujeres casadas y no casadas en Nueva Zelanda, 1965-78, según datos del departamento de estadística. (Modificado de Benfield J, Kjellstrom T. New Zealand ex-nuptial births and domestic purposes benefits in a different perspective. *N Z Nurs J.* 1981;74:28–31.)

Este ejemplo nos enseña dos cosas: en primer lugar, una proporción no es una tasa, y volveremos a este punto en nuestro análisis de la mortalidad; en segundo lugar, un nacimiento puede ser considerado como un acontecimiento, al igual que el desarrollo de una enfermedad, y pueden calcularse las tasas apropiadas. Al estudiar el nacimiento de niños con malformaciones, hay quien prefiere hablar de la prevalencia de malformaciones al nacer en vez de la incidencia de malformaciones al nacer, porque la malformación estaba claramente presente (aunque a menudo no se había diagnosticado), incluso antes del parto. Además, como un porcentaje de casos con malformaciones sufren un aborto antes del parto, las estimaciones de la frecuencia de malformaciones al nacimiento es probablemente una infravaloración significativa de la incidencia verdadera. Por tanto, el término «prevalencia en el nacimiento» es el que suele utilizarse.

En la figura 3.20 se muestran las tasas de incidencia del cáncer de mama en mujeres por edad y la distribución del cáncer de mama en mujeres por edad. Ignórese por ahora el gráfico de barras y considérese la línea curva. El patrón es el de una incidencia que aumenta de modo continuo con la edad, con un cambio en la pendiente de la curva entre los 45 y los 50 años. Este cambio se observa en muchos países. Se ha sugerido que algo ocurre cerca del momento de la menopausia y que el cáncer de mama premenopáusico y el posmenopáusico pueden ser enfermedades diferentes. Obsérvese que, incluso a edades avanzadas, la incidencia o el riesgo de cáncer de mama sigue aumentando.

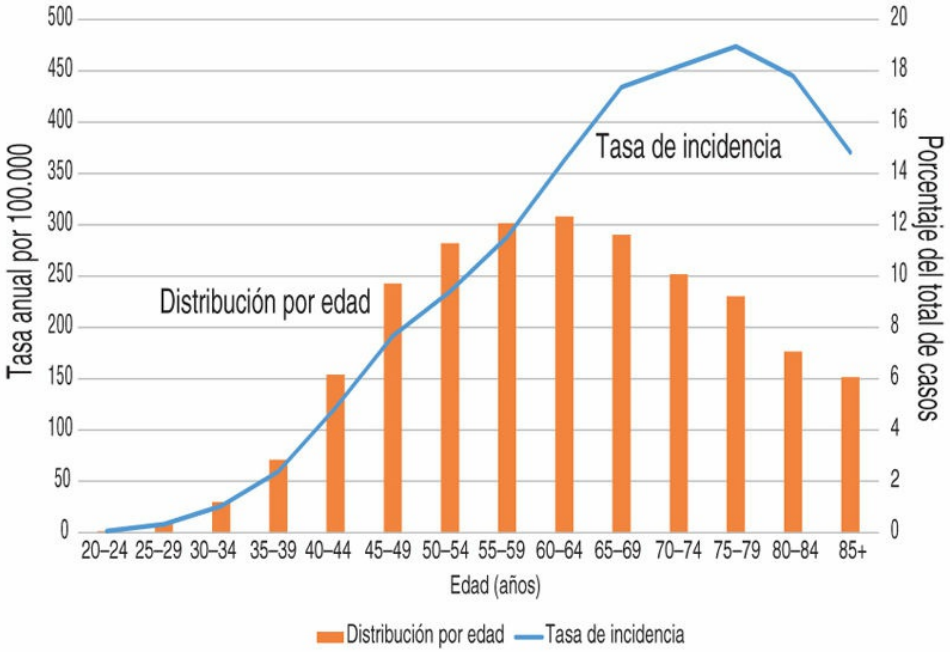


FIG. 3.20 Tasas de incidencia del cáncer de mama en mujeres blancas y distribución de casos por edad, 2000-13. (Datos de Surveillance, Epidemiology, and End Results [SEER] Program [www.seer.cancer.gov] SEER*Stat Database: Incidence—SEER 18 Regs Research. Por cortesía del Dr. Louise Brinton y Mr. Jake Thistle.)

Fijémonos ahora en el gráfico de barras, la distribución de los casos de cáncer de mama por edad. Si la incidencia está aumentando tan espectacularmente con la edad, ¿por qué en el grupo de mujeres de edad más avanzada se producen menos del 5% de los casos? La respuesta es que existen muy pocas mujeres vivas en dicho grupo de edad, por lo que, aunque posean un riesgo muy elevado de presentar un cáncer de mama, el grupo es tan pequeño que contribuye únicamente en una pequeña proporción al número total de casos de cáncer de mama que se observa en todos los grupos de edad. El hecho de que se vean tan pocos casos de cáncer de mama en los grupos de edad avanzada ha contribuido a la falsa impresión pública de que el riesgo de cáncer de mama es bajo en estos grupos y que, por tanto, no es importante realizar mamografías en las ancianas. Este es un error grave. La necesidad de cambiar la idea del público sobre este asunto es un desafío importante de salud pública. Vemos, entonces, la importancia que tiene reconocer la diferencia entre la distribución de la enfermedad o la proporción de casos y la tasa de incidencia o el riesgo de la enfermedad.

Sistema de información geográfica

Un método de estudio de las diferencias geográficas o espaciales en la distribución de casos, ya sea la incidencia o la prevalencia, es trazar los casos en un mapa. El mapeo de la distribución geográfica de los casos se remonta al trabajo del Dr. John Snow, un médico inglés, en medio de un brote de cólera en el distrito de Soho de Londres, Inglaterra, en 1854. La figura 3.21 muestra el mapa en el que Snow trazó las muertes relacionadas con el cólera y las

bombas de agua de la ciudad. Snow usó el mapa para mostrar que los casos de cólera estaban centrados alrededor de una bomba de agua en Broad Street, que era operada por una compañía que tomaba agua de una parte del río Támesis contaminada por aguas residuales. Cuando se clausuró la bomba, la incidencia de casos de cólera disminuyó bruscamente, lo que demuestra la teoría de Snow de que la fuente de la epidemia era el agua contaminada.

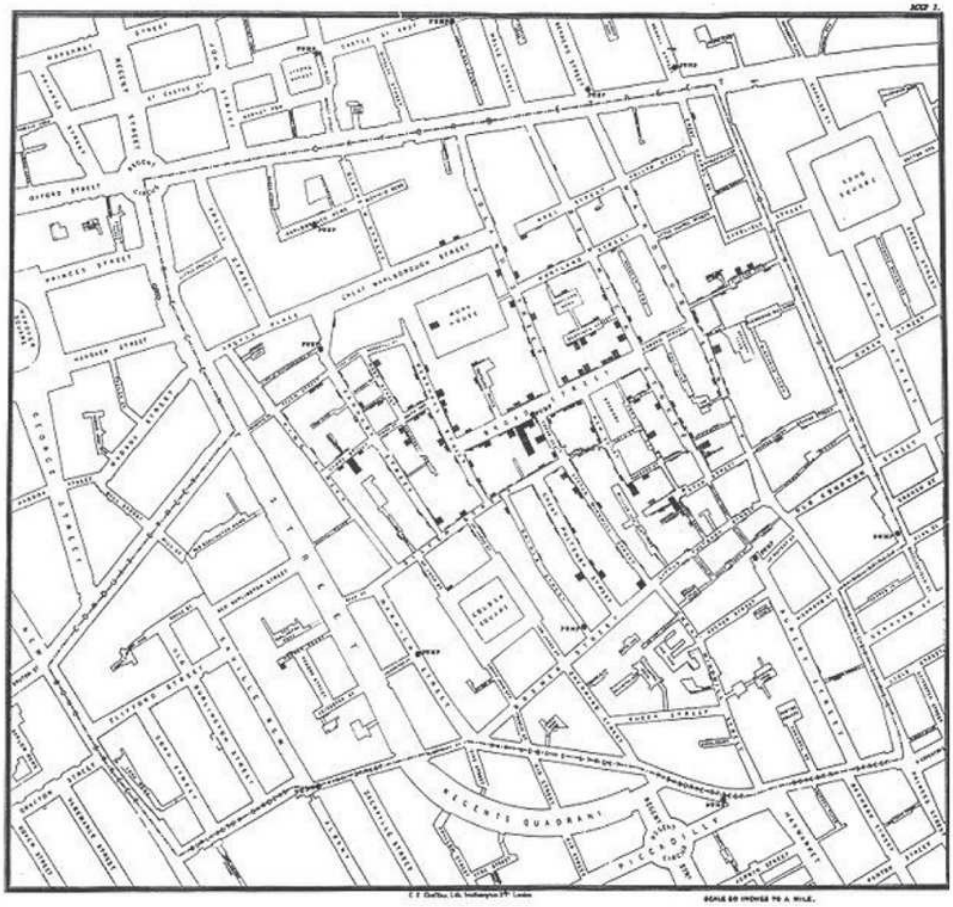


FIG. 3.21 Mapa del distrito de Soho en Londres, Inglaterra, que muestra los agrupamientos de muertes por cólera alrededor de la bomba de Broad Street. (Publicado por C.F. Cheffins, Lith, Southhampton Buildings, London, England, 1854, en: Snow J. *On the Mode of Communication of Cholera.* 2.ª ed. New Burlington Street, London: John Churchill; 1855. https://en.wikipedia.org/wiki/File:Snow-cholera-map-1.jpg.)

La ciencia del uso de mapas continuó evolucionando. El Sistema de Información Geográfica (SIG) utiliza una variedad de información sobre la distribución geográfica de las enfermedades y cómo esta se relaciona con el entorno en el que viven las personas. Posteriormente identifica *grupos* de enfermedades. Esto ayuda a los responsables de formular políticas a identificar y priorizar los problemas de salud y la asignación de recursos. Sin embargo, muchos agrupamientos aparentes son debidos únicamente al azar, y la investigación de estos grupos de casos y descartar una etiología ambiental en lo que parece ser una proximidad mayor de la esperada de

casos de una enfermedad en el tiempo y en el espacio supone un desafío, y ahí es donde el papel de los modernos métodos epidemiológicos y estadísticos espaciales entran en juego.

La figura 3.22 muestra un mapa de la ciudad de Baltimore en el que se exponen los delitos violentos por cada 100 residentes (en gradiente de color) por distrito censal entre 2006 y 2010, y el número de puntos de venta de alcohol (círculos). Usando modelos geoespaciales, los investigadores encontraron que el aumento en el número de puntos de venta de alcohol estaba asociado con un aumento del crimen violento en Baltimore. Tales hallazgos tienen implicaciones políticas y de salud pública muy importantes. Las políticas de código de zonificación deben tener en cuenta cómo afecta a los delitos violentos la presencia de puntos de venta de alcohol en áreas residenciales.

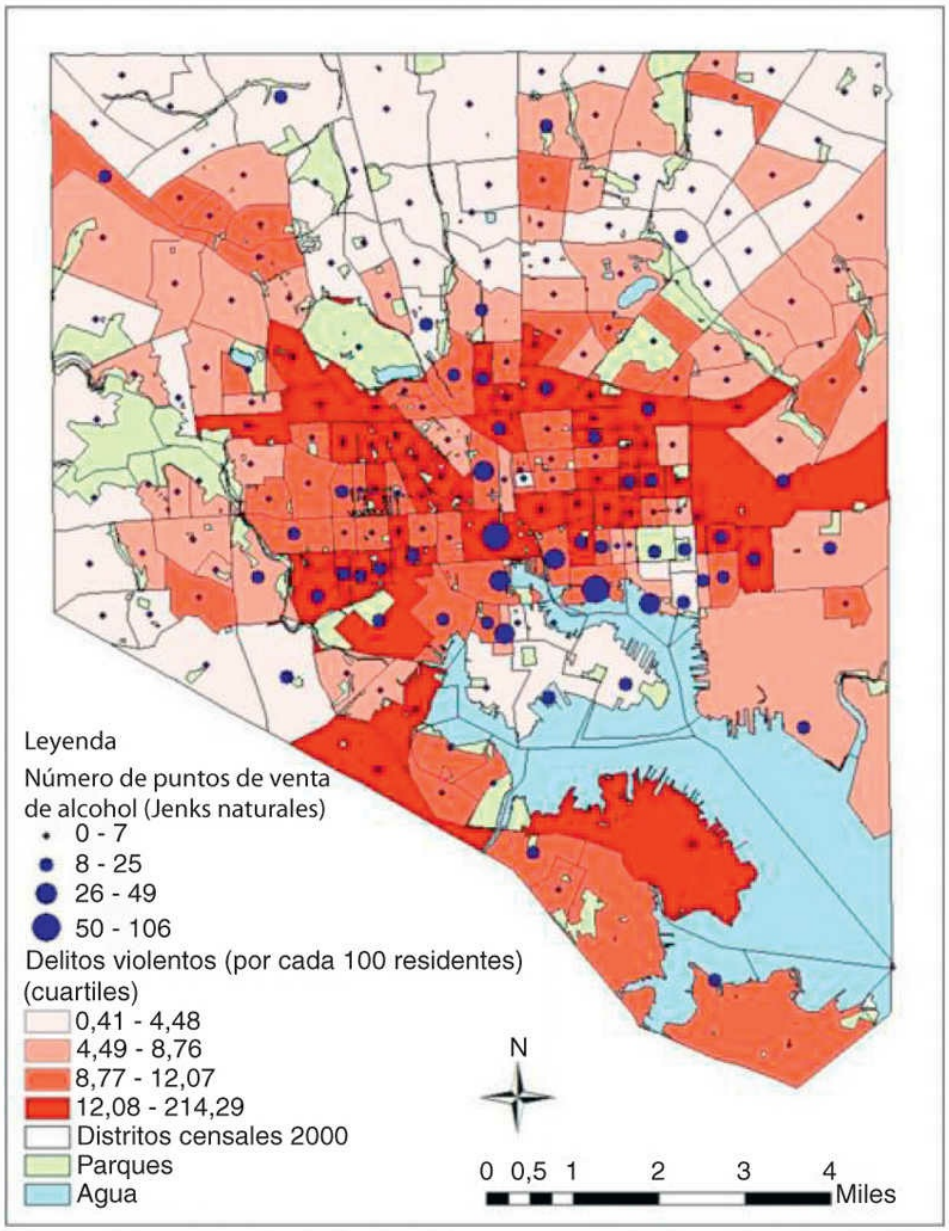


FIG. 3.22 Crímenes violentos (por cada 100 residentes) por distrito censal, ciudad de Baltimore, 2006–10. (Modificado de Jennings JM, Milam AJ, Greiner A, Furr-Holden CD, Curriero FC, Thornton RJ. Neighborhood alcohol outlets and the association with violent crime in one mid-Atlantic City: the implications for zoning policy. *J Urban Health.* 2014;91:62–71.)

Conclusión

En este capítulo hemos hecho hincapié en el importante papel que tiene la epidemiología en la vigilancia de las enfermedades en las poblaciones humanas y la importancia de la vigilancia de la morbilidad en la planificación y el desarrollo de los servicios sanitarios. Esto es un desafío, especialmente en los países en vías de desarrollo, muchos de los cuales carecen de las infraestructuras para la obtención de estadísticas vitales y otros datos rutinarios en poblaciones numerosas (representativas). Hemos revisado diferentes abordajes para medir la morbilidad y hemos visto que una tasa implica la especificación de un numerador, un denominador con las personas en riesgo y el tiempo, ya sea de manera explícita o implícita. En el siguiente capítulo, nos ocuparemos de la medida de la mortalidad. En el capítulo 5 analizaremos cómo utilizamos las pruebas diagnósticas y de cribado para identificar individuos enfermos (incluidos en el numerador) y diferenciarlos de aquellos en la población que no están enfermos. En el capítulo 18 estudiaremos cómo se utiliza la epidemiología para la evaluación de los programas de cribado.

Bibliografía

1. Thacker S, Berkelman RL. Public health surveillance in the United States. *Epidemiol Rev*. 1988;10:164–190.
2. Centers for Disease Control and Prevention. National Notifiable Diseases Surveillance System (NNDSS) [fecha de última consulta: 1 de junio de 2017]. https://wwwn.cdc.gov/nndss/case-definitions.html.
3. Murray CJL, Ortblad KF, Guinovart C, et al. Global, regional, and national incidence and mortality for HIV, tuberculosis, and malaria during 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013. *Lancet*. 2014;384(9947):1005–1070.
4. International Atomic Energy Agency. The Fukushima Daiichi accident; 2015 [fecha de última consulta: 15 de mayo de 2017]. http://www.pub.iaea.org/books/IAEABooks/10962/The-Fukushima-Daiichi-Accident.
5. Bard D, Verger P, Hubert P. Chernobyl, 10 years after: health consequences. *Epidemiol Rev*. 1997;19:187–204.
6. Erkinjuntti T, Østbye T, Steenhuis R, et al. The effect of different diagnostic criteria on the prevalence of dementia. *N Engl J Med*. 1997;337:1667–1674.

Preguntas de repaso del capítulo 3

1. *En la primera exploración de 1.000 varones y 1.000 mujeres de 30 a 35 años en Oxford, Massachusetts, se encontró cefalea de tipo migraña en 5 varones y en 10 mujeres. La conclusión de que las mujeres presentan un riesgo de sufrir una migraña dos veces superior al de los varones en este grupo de edad es:*

a. Correcta.

b. Incorrecta, porque se ha empleado una proporción para comparar las tasas en varones y en mujeres.

c. Incorrecta, debido al fallo de no reconocer el efecto de la edad en los dos grupos.

d. Incorrecta, porque no se aportan datos para realizar una comparación o grupo control.

e. Incorrecta, debido a que no se diferencia entre la incidencia y la prevalencia.

2. *Un estudio sobre prevalencia realizado desde el 1 de enero hasta el 31 de diciembre de 2012 identificó 1.000 casos de esquizofrenia en una ciudad de 2 millones de personas. La tasa de incidencia de esquizofrenia en esta población es de 5/100.000 habitantes/año. ¿Qué porcentaje de los 1.000 casos fueron un nuevo diagnóstico en 2012?*

3. *¿Cuál de las siguientes respuestas es una ventaja de la vigilancia activa?*

a. Requiere menos personal para el proyecto.

b. Es relativamente barata.

c. Es más precisa, ya que reduce la carga de realización de informes por parte de las personas que prestan asistencia sanitaria.

d. Se basa en diferentes definiciones de la enfermedad para tener en cuenta todos los casos.

e. Los sistemas de comunicación pueden organizarse con rapidez.

4. *¿Cuál sería el efecto sobre las tasas de incidencia específicas por edad del cáncer de útero si las mujeres histerectomizadas fuesen excluidas del denominador de los cálculos (asumiendo que en cada grupo de edad hay algunas mujeres que han sido sometidas a una histerectomía)?*

a. Las tasas seguirían siendo las mismas.

b. Las tasas tenderían a disminuir.

c. Las tasas tenderían a aumentar.

d. Las tasas aumentarían en los grupos de mujeres de edad más avanzada y disminuirían en los grupos de mujeres más jóvenes.

e. No es posible determinar si las tasas aumentarían o disminuirían.

5. *De 2008 a 2011 se realizó un estudio en la población adulta no hospitalizada de Estados Unidos. Los resultados de este estudio se muestran a continuación.*

Grupo de edad	**Personas con hipertensión (%)**
18-29 años	4
30-39 años	12
40-49 años	22
50-59 años	43
60-69 años	54
De 70 en adelante	64

Los investigadores afirmaron que en esta población se producía un

aumento del riesgo de hipertensión al avanzar la edad. Según usted, la interpretación de los investigadores:

a. Es correcta.

b. Es incorrecta porque no se basa en tasas.

c. Es incorrecta porque las tasas de incidencia no describen el riesgo.

d. Es incorrecta porque utiliza la prevalencia.

e. Es incorrecta porque los cálculos no están ajustados por edad.

Las preguntas 6 y 7 se basan en la siguiente información:

Población de la ciudad de Atlantis el 30 de marzo de 2012 = 183.000.

N.° de nuevos casos activos de TB que ocurrieron entre el 1 de enero y el 30 de junio de 2012 = 26.

N.° de casos de TB activa según los registros de la ciudad el 30 de junio de 2012 = 264.

6. *La tasa de incidencia de casos activos de TB para el periodo de 6 meses fue de:*

a. 7/100.000 habitantes.

b. 14/100.000 habitantes.

c. 26/100.000 habitantes.

d. 28/100.000 habitantes.

e. 130/100.000 habitantes.

7. *La tasa de prevalencia de TB activa el 30 de junio de 2012 era de:*

a. 14/100.000 habitantes.

b. 130/100.000 habitantes.

c. 144/100.000 habitantes.

d. 264/100.000 habitantes.

e. Ninguna de las anteriores.

8. *La enfermedad X tiene una duración de 15 años y una incidencia baja (5 por 100.000 personas-año). La enfermedad Y tiene una duración de 5 años y una incidencia baja (5 por 100.000 personas-año). Al comparar la enfermedad X con la Y en la misma población, esperaríamos que la enfermedad X tuviera:*

a. Mejor tasa de curación.

b. Menor prevalencia.

c. Mayor prevalencia.

d. Mayor incidencia.

e. Duración media más corta.

9. *Las siguientes estadísticas de salud están disponibles en Internet para el país Z sobre los resultados de dos enfermedades. La enfermedad A tiene una incidencia anual de 225 por 100.000 habitantes y una tasa de mortalidad anual de 150 por 100.000. La enfermedad B tiene una incidencia anual de 500 por 100.000 habitantes y la misma tasa de mortalidad anual que la enfermedad A. Ninguna de las enfermedades A o B tiene cura. ¿Qué concluiría con respecto a la carga de estas enfermedades en el país Z?*

a. La mortalidad proporcional es mayor para la enfermedad A que para la enfermedad B.

b. La tasa de letalidad es mayor para la enfermedad B que para la enfermedad A.

c. La enfermedad A tiene una mayor prevalencia que la enfermedad B.

d. La enfermedad B tiene una mayor prevalencia que la enfermedad A.

e. Los años potenciales de vida perdidos (APVP) son mayores para la enfermedad B que para la enfermedad A.

10. *La infección por el virus chikungunya se introdujo recientemente en la República Dominicana. Durante el primer año después de la introducción, el virus ha infectado a un total de 251.880 personas en la República Dominicana, que tiene una población de 10,4 millones de personas. La infección por el virus chikungunya rara vez es mortal. ¿Cuál de las siguientes afirmaciones es correcta?*

a. La incidencia de la infección por chikungunya es de 251.880 por año.

b. La prevalencia de infecciones por chikungunya es de 251.880.

c. La tasa de mortalidad de chikungunya es de 24,2 por 1.000 personas.

d. La incidencia acumulada de chikungunya a un año es de 24,2 por 1.000 personas.

e. La supervivencia acumulada de las infecciones por chikungunya

es de 24,2 por cada 1.000 personas.

CAPÍTULO 4

La ocurrencia de la enfermedad: II. Mortalidad y otras medidas del impacto de la enfermedad

Palabras clave

tasa de mortalidad
ajuste por edad
tasa de mortalidad estandarizada
calidad de vida
mortalidad proporcional y años potenciales de vida perdidos (APVP)

Usted no muere por haber nacido, ni por haber vivido, ni por envejecer. Usted se muere de algo… La muerte natural no existe: nada de lo que le sucede a un hombre es nunca natural, ya que su presencia pone el mundo en cuestión. Todos los hombres deben morir: pero para cada hombre su muerte es un accidente e incluso, aunque lo sepa y lo consienta, es una violación injustificable.

—Simone de Beauvoir, escrito sobre la muerte de su madre, en Una muerte muy dulce[1]

Objetivos de aprendizaje

- Comparar las diferentes medidas de la mortalidad, como las tasas de mortalidad, la tasa de letalidad, la mortalidad proporcional y los años potenciales de vida perdidos.
- Mostrar cuándo la mortalidad puede aproximarse al riesgo de la enfermedad.
- Introducir aspectos que pueden surgir al comparar la mortalidad en dos o más poblaciones.
- Definir, calcular e interpretar tasas de mortalidad directas e indirectas ajustadas por la edad.

- Introducir otras medidas del impacto de la enfermedad.

La mortalidad tiene un gran interés por varios motivos. El primero de todos es que la muerte es la última experiencia por la que todo ser humano está destinado a pasar. La muerte tiene claramente una tremenda importancia para cada persona, incluyendo preguntas sobre dónde y cómo ocurrirá y si existe alguna forma de retrasarla. Desde el punto de vista del estudio de la ocurrencia de la enfermedad, expresar la mortalidad en términos cuantitativos puede señalar diferencias en el riesgo de morir de una enfermedad entre las personas de diferentes áreas geográficas y subgrupos de la población. Las tasas de mortalidad sirven para medir la gravedad de la enfermedad y nos pueden ayudar a determinar si el tratamiento de una enfermedad se ha vuelto más efectivo con el paso del tiempo. Además, debido al problema que a menudo surge a la hora de identificar nuevos casos de una enfermedad, las tasas de mortalidad pueden servir de sustitutos de las tasas de incidencia cuando la enfermedad estudiada es grave y mortal. Este capítulo abordará la expresión cuantitativa de la mortalidad y el uso de dichas medidas en los estudios epidemiológicos.

Medidas de la mortalidad

En la figura 4.1 se muestra el número de muertes por cáncer desde el año 1969 hasta 2014 en Estados Unidos. Se observa claramente que el *número* absoluto de personas que fallecen de cáncer hasta 2014 está aumentando de manera importante; sin embargo, a partir de este gráfico no podemos concluir que el *riesgo* de morir de cáncer está aumentado, porque los únicos datos que nos ofrece el gráfico son números de fallecimientos (numeradores), pero no contamos con los denominadores (población en riesgo). Si, por ejemplo, el tamaño de la población estadounidense también estuviese aumentando al mismo ritmo, el riesgo de morir por cáncer no cambiaría.

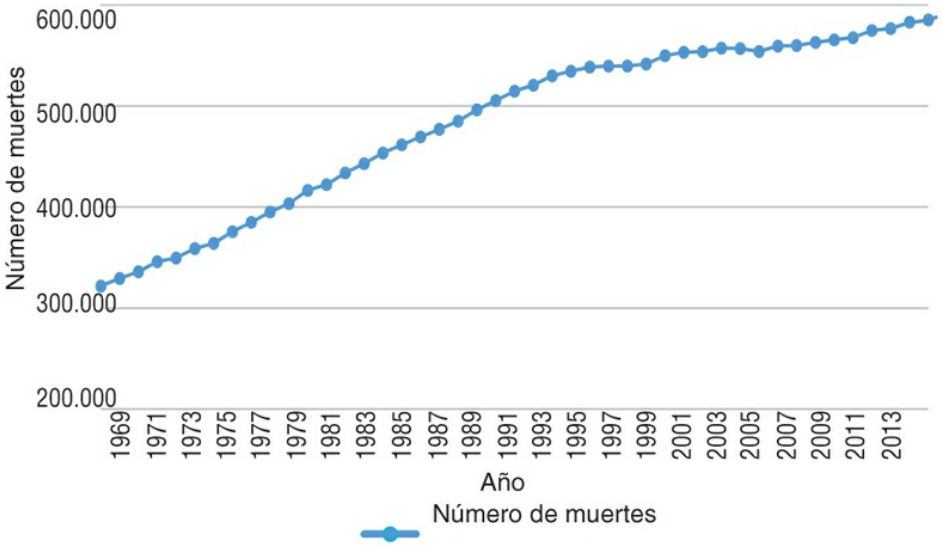


FIG. 4.1 Tendencia de las cifras de muertes por cáncer observadas en hombres y en mujeres en Estados Unidos, 1969-2014. (Datos de Weir HK, Anderson RN, Coleman King SM, et al. Heart disease and cancer deaths—trends and projections in the United States, 1969–2020. *Prev Chronic Dis.* 2016;13:160211.)

Por este motivo, si queremos abordar el riesgo de morir, debemos emplear tasas. En la figura 4.2 se muestran las tasas de mortalidad de diversos tipos de cáncer en varones, de 1930 a 2014. El aumento más espectacular se observa en las muertes por cáncer de pulmón. Este aumento es claramente de proporciones epidémicas y, lo más trágico, es que el cáncer de pulmón es una causa de muerte evitable. Afortunadamente, desde mediados de la década de 1990, la mortalidad por cáncer de pulmón ha disminuido, de modo paralelo a la reducción previa del tabaquismo entre los varones. Otros cánceres también son de interés. La mortalidad ajustada por edad por cáncer de próstata también repuntó a mediados de la década de 1990 y desde entonces se ha reducido. Los cánceres de colon y recto han disminuido a lo largo de muchos años. La tasa de mortalidad por cáncer de estómago ha disminuido espectacularmente desde 1930, aunque la explicación exacta es desconocida.

Se cree que la disminución puede ser el resultado de la mayor disponibilidad de refrigeración, que disminuyó la necesidad de ahumar los alimentos y redujo, por tanto, la exposición humana a los carcinógenos producidos en el proceso de ahumado. Otra causa posible es la mejora en las condiciones de higiene, que puede haber reducido las infecciones por *Helicobacter pylori*, que han sido implicadas en la etiología (causa) del cáncer de estómago.

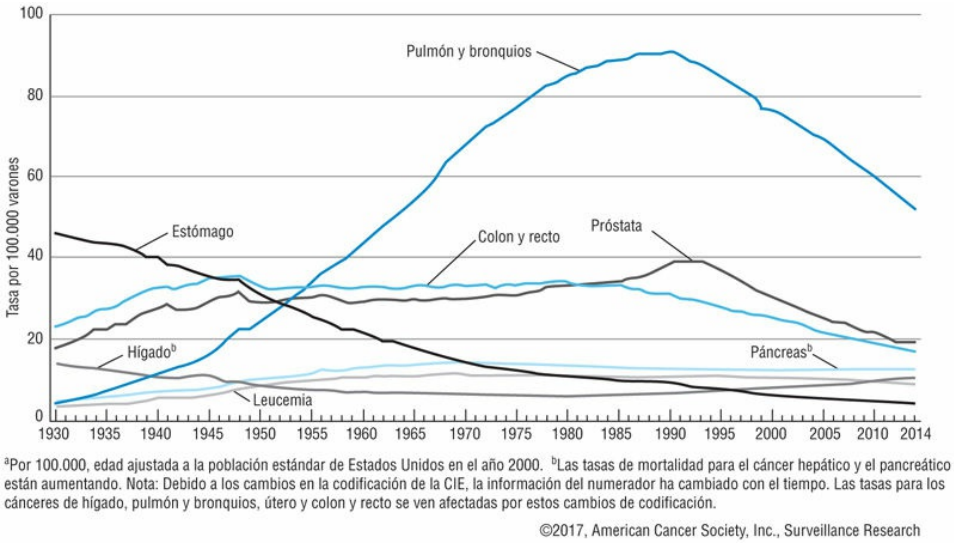
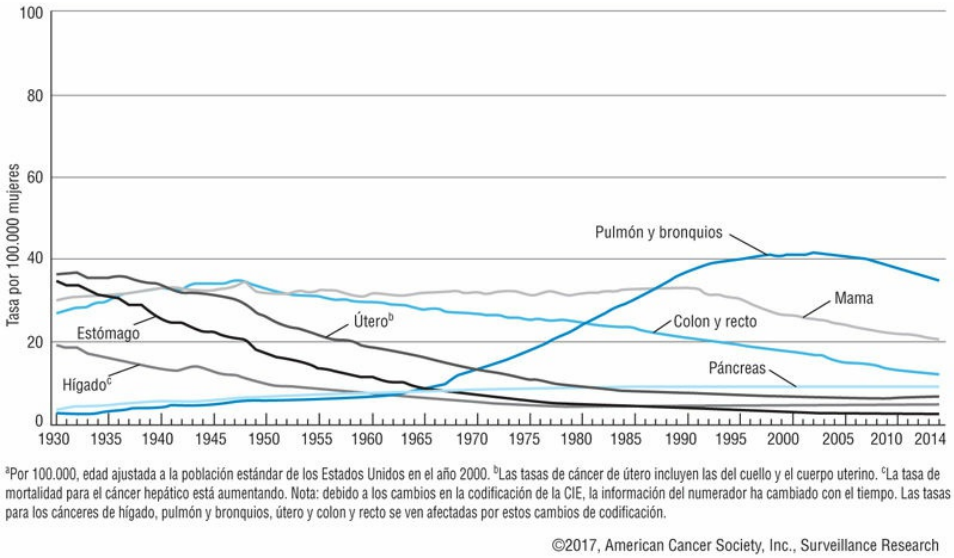


FIG. 4.2 Tasas de mortalidad por cáncer en varones, Estados Unidos, 1930-2014 (ajustadas por edad según la población estándar de Estados Unidos en el año 2000). *CIE*, Clasificación Internacional de Enfermedades. (De American Cancer Society, Cancer Facts & Figures 2017. Basado en US Mortality Volumes 1930 to 1959, US Mortality Data, 1960 to 2014.)

En la figura 4.3 se muestra una presentación similar de la mortalidad por cáncer en las mujeres en el periodo de 1930 a 2014. La mortalidad por cáncer de mama se ha mantenido en prácticamente el mismo nivel durante muchos años, pero ha disminuido desde principios de la década de 1990 hasta 2014. Sería deseable estudiar los cambios en la incidencia del cáncer de mama. Sin embargo, dicho estudio es difícil, porque, debido a las campañas agresivas de educación pública que fomentan la práctica de mamografías y autoexploraciones mamarias, muchos de los cánceres de mama detectados en la actualidad en etapas mucho más tempranas podrían haber pasado desapercibidos hace años. Sin embargo, los datos disponibles indican que la incidencia verdadera del cáncer de mama en las mujeres puede haber aumentado durante muchos años y posteriormente se ha reducido de 2001 a 2014.



[a]Por 100.000, edad ajustada a la población estándar de los Estados Unidos en el año 2000. [b]Las tasas de cáncer de útero incluyen las del cuello y el cuerpo uterino. [c]La tasa de mortalidad para el cáncer hepático está aumentando. Nota: debido a los cambios en la codificación de la CIE, la información del numerador ha cambiado con el tiempo. Las tasas para los cánceres de hígado, pulmón y bronquios, útero y colon y recto se ven afectadas por estos cambios de codificación.

©2017, American Cancer Society, Inc., Surveillance Research

FIG. 4.3 Tasas de mortalidad por cáncer en mujeres, Estados Unidos, 1930-2014 (ajustadas por edad según la población estándar de Estados Unidos en el año 2000). *CIE, Clasificación Internacional de Enfermedades*. (De American Cancer Society, Cancer Facts & Figures 2017. Basado en US Mortality Vol. 1930 to 1959, US Mortality Data, 1960 to 2014. National Center for Health Statistics, Centers for Disease Control and Prevention.)

La mortalidad por cáncer de útero ha disminuido, debido quizá a la detección y diagnóstico precoz. La mortalidad por cáncer de pulmón ha aumentado en las mujeres, y el cáncer de pulmón ha superado al cáncer de mama como causa de muerte en las mujeres. El cáncer de pulmón es casi totalmente prevenible, ya que se debe principalmente a un hábito del estilo de vida (el tabaquismo) adoptado voluntariamente por muchas mujeres y en la actualidad es la principal causa de muerte por cáncer en las mujeres de Estados Unidos.

El tema de la mortalidad relacionada con la edad nos suscita un especial interés. En la figura 4.4 se muestran las tasas de mortalidad por cáncer y por cardiopatías en personas menores de 65 años y en personas de 65 años o mayores. El cáncer es la principal causa de muerte en varones y mujeres menores de 65 años, pero, por encima de esta edad, las cardiopatías superan claramente al cáncer como causa de muerte.

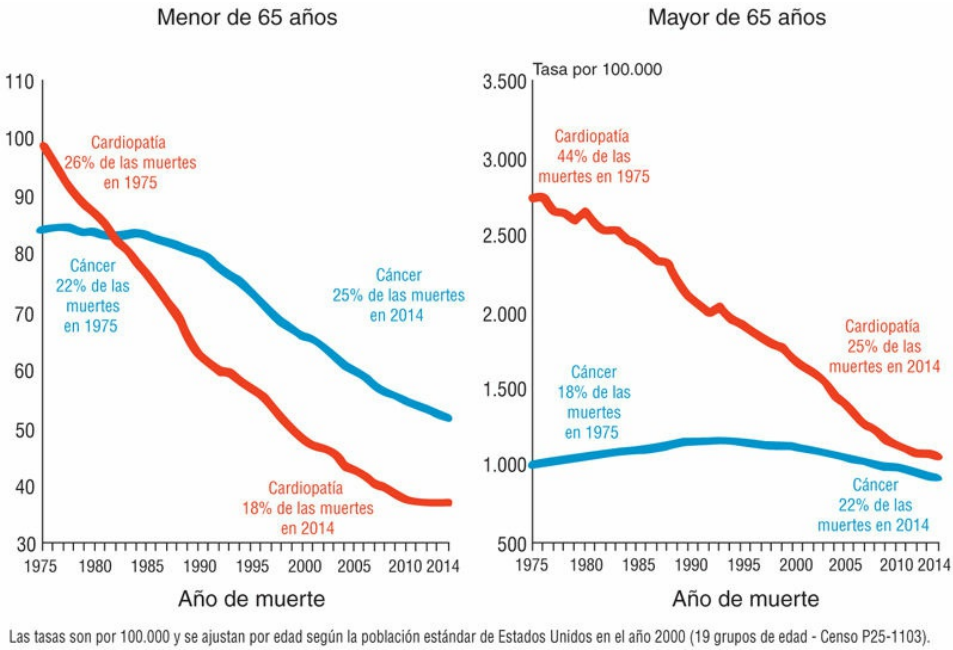


FIG. 4.4 Tasas de mortalidad por cáncer y cardiopatía en individuos de menos de 65 años y de 65 o más años (ajustadas por edad según la población estándar de Estados Unidos en el año 2000). (De Howlader N, Noone AM, Krapcho M, et al, eds. *SEER Cancer Statistics Review, 1975–2014.* Bethesda, MD: National Cancer Institute; https://seer.cancer.gov/csr/1975_2014/, basado en datos presentados al SEER en noviembre de 2016, publicados en la página web del SEER, abril de 2017. Utilizando datos de US Mortality Files, National Center for Health Statistics, Centers for Disease Control and Prevention.)

En la figura 4.5 se exponen las causas de muerte a nivel mundial en los niños menores de 5 años en 2015. Las principales causas de muerte entre los niños menores de 5 años en 2015 fueron las complicaciones del parto prematuro, la neumonía, las complicaciones relacionadas con el parto, la diarrea y las anomalías congénitas. Las muertes neonatales representaron el 45% de las muertes de menores de 5 años en 2015. Las enfermedades infecciosas produjeron más de la mitad de los 5,9 millones de muertes de niños menores de 5 años, con el mayor porcentaje debido a la neumonía, la diarrea y el paludismo.

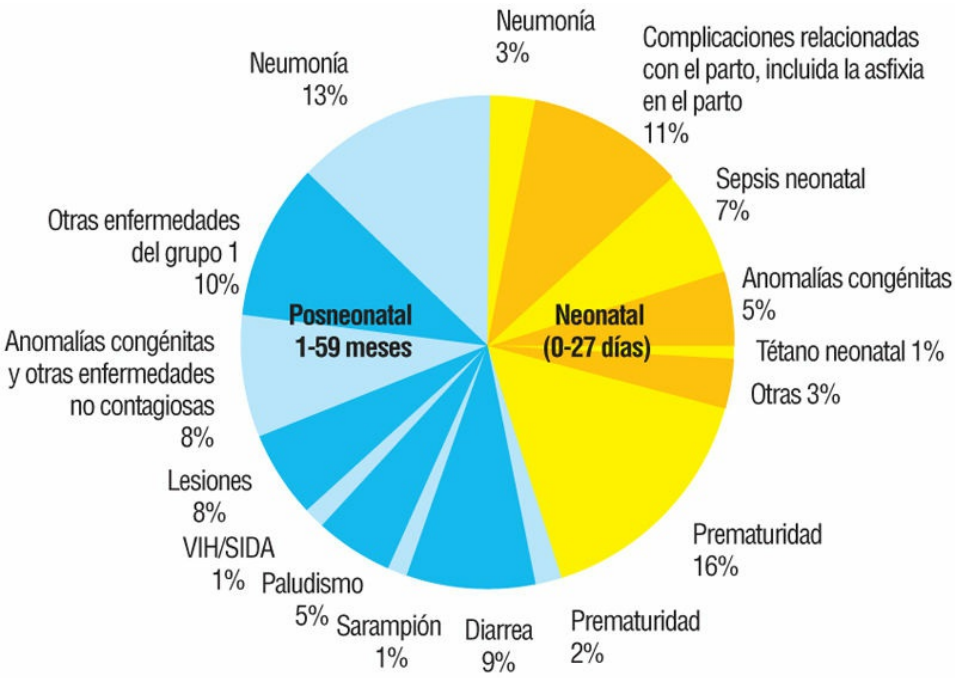


FIG. 4.5 Principales causas de muerte en niños menores de 5 años en 2015. *SIDA*, síndrome de inmunodeficiencia adquirida; *VIH*, virus de la inmunodeficiencia humana. (De World Health Organization. *MCEE Methods and Data Sources for Child Causes of Death 2000–2015. Global Health Estimates Technical Paper WHO/HIS/IER/GHE/2016.1.*)

Tasas de mortalidad

¿Cómo se expresa la mortalidad en términos cuantitativos? Examinemos algunos tipos de tasas de mortalidad. La primera es la tasa de mortalidad anual, o tasa de mortalidad, por todas las causas:

$$\text{Tasa de mortalidad anual por todas las causas (por 100.000 habitantes)} = \frac{\text{N.º total de muertes por todas las causas en 1 año}}{\text{N.º de personas en la población a mitad de año}} \times 100.000$$

Obsérvese que, como la población cambia a lo largo del tiempo, el número de personas en la población a mitad de año suele usarse como una aproximación de la población media.

Los mismos principios mencionados en el análisis de la morbilidad son aplicables al de la mortalidad: para que una tasa de mortalidad tenga sentido,

toda la población del grupo representada en el denominador debe tener la posibilidad de entrar en el grupo representado en el numerador.

Puede que no siempre estemos interesados en una tasa para toda la población; quizá estemos interesados únicamente en un cierto grupo de edad, en los varones o en las mujeres, o en un grupo étnico. Por tanto, si estamos interesados en la mortalidad en los niños menores de 10 años, podemos calcular una tasa específicamente para dicho grupo:

$$\text{Tasa de mortalidad anual por todas las causas en niños menores de 10 años (por 1.000 habitantes)} = \frac{\text{N.º de muertes por todas las causas en un año en niños menores de 10 años}}{\text{N.º de niños menores de 10 años en la población a mitad de año}} \times 1.000$$

Al poner una restricción, como la edad, por ejemplo, la misma restricción debe aplicarse al numerador *y* al denominador, de modo que todas las personas del grupo del denominador estarán en riesgo de entrar en el grupo del numerador. Cuando dicha restricción se aplica a una tasa, se denomina *tasa específica*. La tasa anterior es, por tanto, una *tasa de mortalidad específica por edad*.

También podríamos establecer una restricción en una tasa especificando un diagnóstico y, por tanto, limitar la tasa a las muertes debidas a una cierta enfermedad, es decir, una tasa *específica de causa* o *específica por enfermedad*. Por ejemplo, si estamos interesados en la mortalidad debida al cáncer de pulmón, podríamos calcularla del siguiente modo:

$$\text{Tasa de mortalidad anual por cáncer de pulmón (por 1.000 habitantes)} = \frac{\text{N.º de muertes por cáncer de pulmón en un año}}{\text{N.º de personas en la población a mitad de año}} \times 1.000$$

También podemos plantear restricciones en más de una característica simultáneamente (p. ej., edad y causa de muerte) del siguiente modo:

$$\text{Tasa de mortalidad anual por leucemia en niños menores de 10 años (por 1.000 habitantes)} = \frac{\text{N.º de muertes por leucemia en un año en niños menores de 10 años}}{\text{N.º de niños menores de 10 años en la población a mitad de año}} \times 1.000$$

El tiempo también debe especificarse en toda tasa de mortalidad. La mortalidad puede calcularse en un periodo de 1 año, 5 años o más. El periodo seleccionado es arbitrario, pero debe especificarse con precisión.

Tasa de letalidad

Debemos diferenciar entre *tasa de mortalidad* y *tasa de letalidad*. La tasa de letalidad se calcula del siguiente modo:

$$\text{Tasa de letalidad (porcentaje)} = \frac{\text{N.º de sujetos que mueren durante un periodo de tiempo específico tras el comienzo o el diagnóstico de la enfermedad}}{\text{N.º de sujetos con la enfermedad específica}} \times 100$$

En otras palabras, ¿qué porcentaje de personas *que tienen cierta enfermedad* fallecen en un tiempo determinado tras el diagnóstico de dicha enfermedad? (De manera ideal, quisiéramos utilizar la fecha de comienzo de la enfermedad como el comienzo del periodo de tiempo específico en el numerador. Sin embargo, la fecha de comienzo de la enfermedad a menudo es difícil de estandarizar, ya que muchas enfermedades evolucionan insidiosamente (de modo asintomático) a lo largo de un periodo de tiempo prolongado. Como resultado, en muchas enfermedades crónicas puede ser difícil determinar con precisión cuándo comenzó el proceso de la enfermedad. Por ejemplo, muchos pacientes con artritis no pueden recordar cuándo comenzó el dolor articular. En la práctica, por tanto, a menudo utilizamos los datos del diagnóstico como una medida sustituta de la fecha de comienzo de la enfermedad, porque la fecha exacta del diagnóstico por lo general puede documentarse a partir de los registros médicos disponibles.) Cuando la información se obtiene de la población, se debe destacar que, si la enfermedad en cuestión es grave, la fecha en la que se realizó el diagnóstico puede haber sido una fecha que marcó la vida del paciente y, por tanto, difícil de olvidar.

¿Cuál es la diferencia entre la tasa de letalidad y la tasa de mortalidad? En la tasa de mortalidad, el denominador representa a toda la población con riesgo de morir de la enfermedad, incluyendo tanto a los que presentan la enfermedad como a los que no la presentan (pero que tienen *riesgo* de padecerla). Sin embargo, en la tasa de letalidad, el denominador se limita a los que ya padecen la enfermedad. Por tanto, la tasa de letalidad es una medida de la gravedad de la enfermedad. También sirve para medir los beneficios de un nuevo tratamiento: a medida que el tratamiento mejora, es de esperar que la tasa de letalidad disminuya.

Lo ideal es que el numerador de la tasa de letalidad se limite a las muertes *debidas a esa enfermedad*. Sin embargo, no siempre es fácil diferenciar las muertes debidas a esa enfermedad de las muertes debidas a otras causas. Por ejemplo, una persona alcohólica puede morir en un accidente de tráfico; sin embargo, la muerte puede estar relacionada o no con el consumo de alcohol.

Fijémonos en un ejemplo hipotético para aclarar la diferencia entre mortalidad y letalidad (cuadro 4.1).

Cuadro 4.1 Comparación de la tasa de mortalidad y la tasa de letalidad en el mismo año

Suponga una población de 100.000 habitantes, de los cuales 20 están enfermos con la enfermedad X y, en un año, 18 de los 20 mueren de dicha enfermedad

$$\text{La tasa de mortalidad de la enfermedad } X = \frac{18}{100.000} = 0,00018 \text{ o } 0,018\%$$

$$\text{La tasa de letalidad de la enfermedad X} = \frac{18}{20} = 0,9 \text{ o } 90\%$$

Asumamos que en una población de 100.000 personas, 20 tienen la enfermedad X. En 1 año, 18 personas fallecen por dicha enfermedad. La mortalidad es muy baja (0,018%) porque la enfermedad es rara; sin embargo, cuando una persona presenta la enfermedad, la probabilidad de morir es alta (90%).

Mortalidad proporcional

Otra medida de la mortalidad es la mortalidad proporcional, que no es una tasa. La mortalidad proporcional por enfermedades cardiovasculares en Estados Unidos en el año 2015 se define así:

$$\text{Mortalidad proporcional por enfermedades cardiovasculares en EE. UU. en el año 2015 (porcentaje)} = \frac{\text{N.º de muertes por enfermedades cardiovasculares en EE. UU. en 2015}}{\text{Total de muertes en EE. UU. en 2015}} \times 100$$

En otras palabras, de todas las muertes en Estados Unidos, ¿qué *proporción* se debió a enfermedades cardiovasculares? En la figura 4.6 se muestra la mortalidad proporcional por cardiopatías por grupo de edad. En cada grupo de edad, la barra completa representa todas las muertes (100%) y las muertes debidas a cardiopatía se indican con la porción azul oscuro. Observamos que la *proporción* de muertes debidas a cardiopatía aumenta con la edad. Sin embargo, esto no indica que el *riesgo* de muerte por cardiopatía también esté aumentando. Esto se demuestra en los siguientes ejemplos.

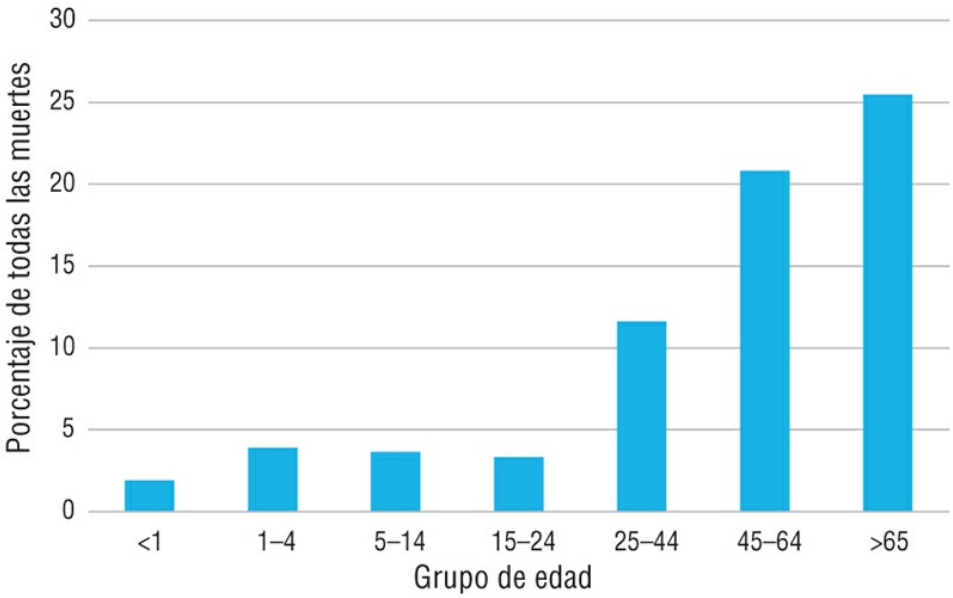


FIG. 4.6 Muertes por cardiopatía como porcentaje de las muertes por todas las causas, por grupo de edad, Estados Unidos, 2014. (De National Center for Health Statistics [NCHS]. *Datos de Health, United States, 2015, With Special Feature on Racial and Ethnic Health Disparities.* Hyattsville, MD: NCHS; 2016.)

La tabla 4.1 muestra todas las muertes y las muertes debidas a cardiopatía en dos comunidades, A y B. La mortalidad por todas las causas en la comunidad A es el doble que en la comunidad B. Cuando nos fijamos en la mortalidad proporcional, observamos que el 10% de las muertes en la comunidad A y el 20% de las muertes en la comunidad B se deben a cardiopatías. ¿Significa esto que el riesgo de morir de una cardiopatía es dos veces más alto en la comunidad B que en la A? La respuesta es no. Cuando se calculan las tasas de mortalidad por cardiopatías en las dos comunidades (10% de 30/1.000 y 20% de 15/1.000), vemos que las tasas de mortalidad son idénticas.

Tabla 4.1

Comparación entre tasa de mortalidad y la mortalidad proporcional: I. Muertes por cardiopatías en dos comunidades

	Comunidad A	Comunidad B
Tasa de mortalidad por todas las causas	30/1.000	15/1.000
Mortalidad proporcional por cardiopatía	10%	20%
Tasa de mortalidad por cardiopatía	3/1.000	3/1.000

Si observamos un cambio en la mortalidad proporcional por cierta enfermedad a lo largo del tiempo, el cambio puede deberse no a cambios en la mortalidad por dicha enfermedad, sino a cambios en la mortalidad por alguna otra enfermedad. Consideremos un ejemplo hipotético: en la tabla 4.2 observamos las tasas de mortalidad por cardiopatía, cáncer y otras causas en una población en un periodo temprano y en un periodo tardío. En primer lugar, comparemos las tasas de mortalidad en dos periodos temporales: la mortalidad por cardiopatías se duplicó con el paso del tiempo (de 40/1.000 a

80/1.000), pero las tasas de mortalidad por cáncer y por el resto de las causas no cambiaron (20/1.000). Sin embargo, si nos fijamos ahora en la mortalidad proporcional por cada causa, observamos que la mortalidad por cáncer y por otras causas ha disminuido en la población, pero solo porque la mortalidad proporcional por cardiopatías ha aumentado. Por tanto, si la proporción de un segmento de la «tarta» de la mortalidad aumenta, necesariamente se producirá una disminución en la proporción de algún otro segmento (fig. 4.7). En la figura 4.8 se expone otra forma de ver este aspecto.

Tabla 4.2

Ejemplo hipotético de las tasas de mortalidad y mortalidad proporcional en dos periodos

	Periodo temprano		Periodo tardío	
Causa de muerte	Tasa de mortalidad	Mortalidad proporcional	Tasa de mortalidad	Mortalidad proporcional
Cardiopatía	40/1.000	50%	80/1.000	66,7%
Cáncer	20/1.000	25%	20/1.000	16,7%
Todas las demás causas	20/1.000	25%	20/1.000	16,7%
Todas la muertes	80/1.000	100%	120/1.000	100,0%

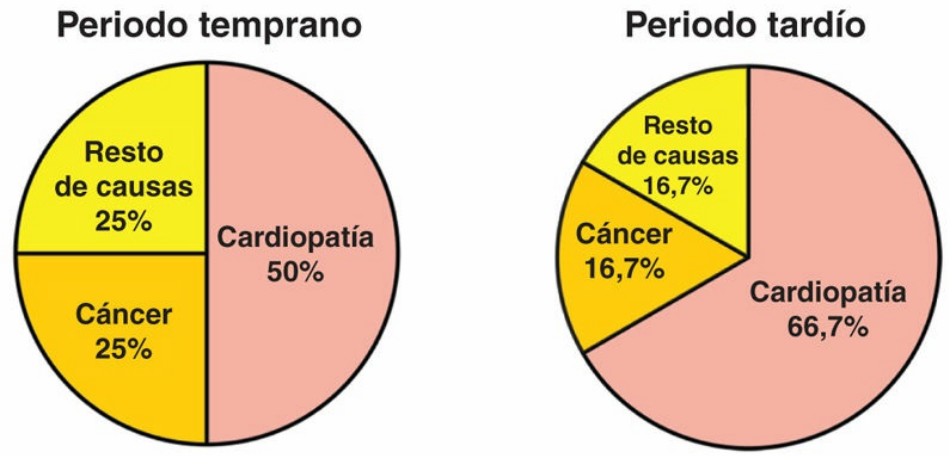


FIG. 4.7 Ejemplo hipotético de mortalidad proporcional: cambios en la mortalidad proporcional por cardiopatía, cáncer y otras causas desde el periodo temprano al periodo tardío.



FIG. 4.8 «¿Sabes qué? Los días se alargan al mismo tiempo que las noches se acortan.» Comprensión de la mortalidad proporcional. (Family Circus © 2002 Bill Keane, Inc. Distribuido por King Features Syndicate, Inc.)

Como se observa en el ejemplo de la tabla 4.3, si las tasas de mortalidad por todas las causas difieren, las tasas de mortalidad específica de causa pueden diferir significativamente, incluso cuando la mortalidad proporcional es la misma. Estos ejemplos nos muestran que, aunque la mortalidad proporcional puede darnos una visión rápida de las principales causas de muerte, no puede decirnos el riesgo de morir de una enfermedad. Es por este motivo por lo que necesitamos una tasa de mortalidad.

Tabla 4.3

Comparación entre tasa de mortalidad y mortalidad proporcional: II. Muertes por cardiopatías en dos comunidades

	Comunidad A	Comunidad B
Tasa de mortalidad por todas las causas	20/1.000	10/1.000
Mortalidad proporcional por cardiopatía	30%	30%
Tasa de mortalidad por cardiopatía	6/1.000	3/1.000

Años potenciales de vida perdidos

En los últimos años se ha estado utilizando cada vez más otro índice de mortalidad, los años potenciales de vida perdidos (APVP), con el fin de establecer prioridades sanitarias. Los APVP son una medida de la mortalidad prematura, o muerte precoz. Los APVP reconocen que la muerte que se produce en una persona a una edad más temprana implica claramente una mayor pérdida de años productivos futuros que la muerte que se produce a edades más avanzadas. Este índice se calcula en dos pasos. En un primer paso, para cada causa se resta la edad de la persona fallecida en el momento de morir de una edad predeterminada (o «promedio») en el momento de la muerte. En Estados Unidos, esta edad «estándar» predeterminada suele ser 75 años. Por tanto, un lactante que muera con 1 año ha perdido 74 años de vida (75 menos 1) y una persona que fallezca con 50 años ha perdido 25 años de vida (75 menos 50). Por tanto, cuanto más joven muera una persona, más años potenciales de vida se pierden. En el segundo paso, los «años potenciales de vida perdidos» de cada sujeto se suman para obtener los APVP totales para una causa de muerte específica. Cuando leemos publicaciones que utilizan los APVP, es importante tener en cuenta qué suposiciones ha realizado el autor, así como qué edad estándar predeterminada se ha elegido.

En la figura 4.9 se muestran los APVP en Estados Unidos antes de los 75 años en 2015. La barra superior muestra el total de APVP por todas las causas (100%) y las barras inferiores representan los APVP individuales para cada causa principal de muerte, con el porcentaje responsable del total de APVP. Observamos que, individualmente, la causa más importante de APVP son las neoplasias malignas, que en el mismo año fueron la segunda causa principal de muerte por su tasa de mortalidad (v. fig. 1.2 y tabla 1.1). En 2015, las lesiones no intencionadas ocupaban el cuarto puesto de la clasificación por su tasa de mortalidad, mientras que en la clasificación de APVP ocupaban el segundo puesto. Esta discrepancia se debe al hecho de que las lesiones son la principal causa de muerte hasta los 34 años de edad, por lo que suponen una proporción elevada de APVP.

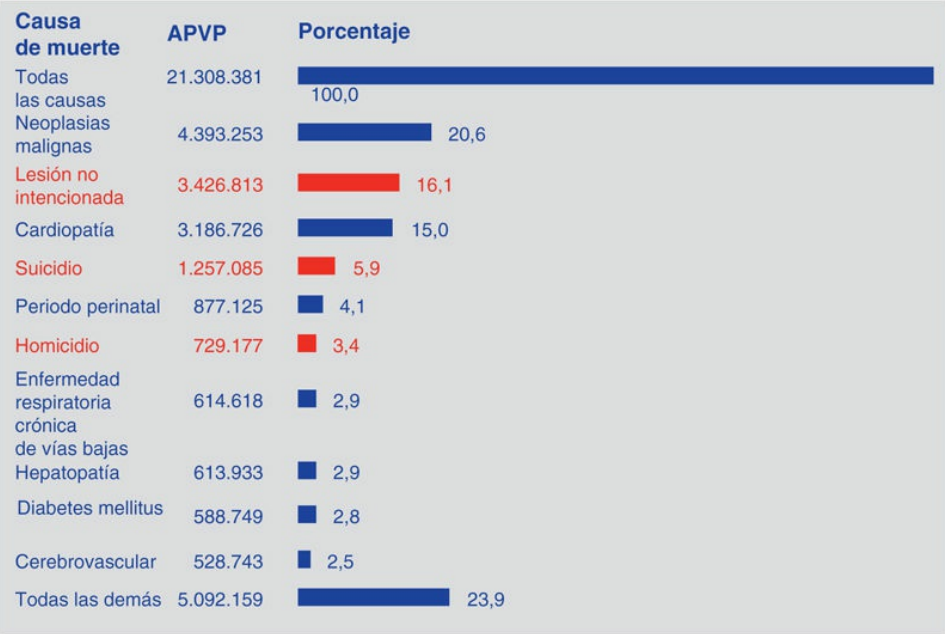
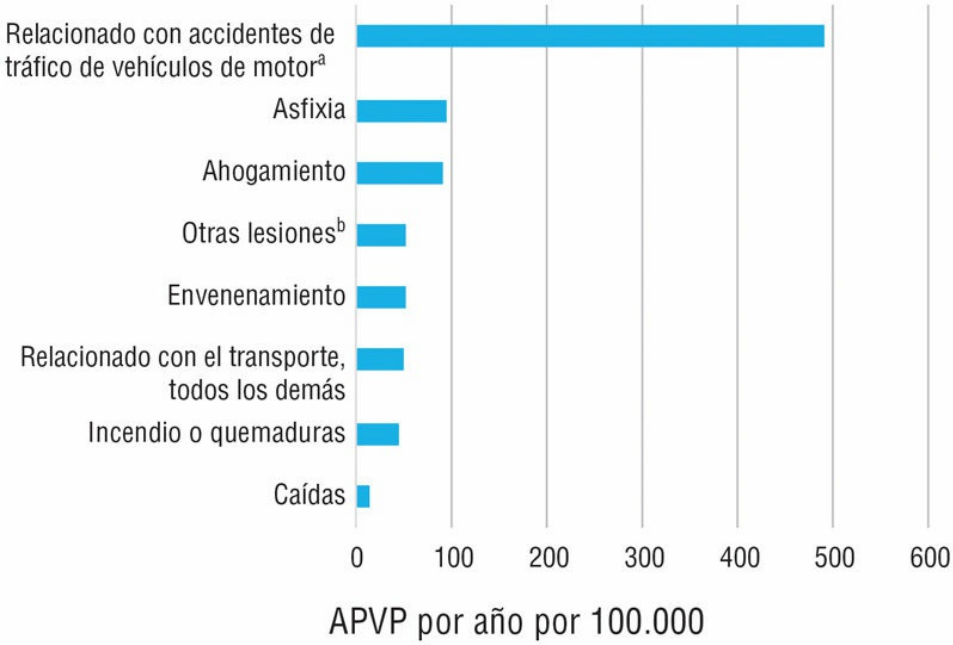


FIG. 4.9 Años potenciales de vida perdidos *(APVP)* antes de los 75 años de edad, en todas las razas en ambos sexos, todas las muertes, Estados Unidos, 2015. Las barras rojas representan causas de muerte no relacionadas con la enfermedad. (Modificado de Centers for Disease Control and Prevention. National Center for Injury Prevention and Control. Years of Potential Life Lost [YPLL] reports, 1999–2015, https://webappa.cdc.gov/sasweb/ncipc/ypll10.html. Consultado el 6 de junio de 2017.)

En la figura 4.10 se muestran los APVP por lesiones no intencionales antes de los 75 años en población con edades entre los 0 y los 19 años. Observamos que los APVP por accidentes de vehículos motorizados representan más de la mitad de los APVP en este grupo. Por tanto, si queremos actuar sobre los APVP en los niños y los adultos jóvenes, debemos abordar esta causa específica de lesiones relacionadas con vehículos motorizados.



[a]Categorizado por persona lesionada e incluye ocupantes de vehículos motorizados, motoci-clistas, ciclistas a pedal, peatones, ocupantes o pasajeros de otros modos de transporte en un accidente automovilístico y accidentes de vehículos motorizados para los cuales no se especifica la persona lesionada.

[b]Cortes o perforaciones, lesiones no intencionales relacionadas con armas de fuego, lesiones relacionadas con maquinaria, lesiones por causas naturales y ambientales, esfuerzo excesivo, golpe por o contra un objeto, y otros especificados y no especificados.

FIG. 4.10 Años potenciales de vida perdidos *(APVP)* anualizados, por 100.000 personas de 0 a 19 años, por lesiones no intencionales, Estados Unidos, 2000–09. (Modificado de Centers for Disease Control and Prevention. Years of potential life lost from unintentional injuries among persons aged 0– 19 years—United States, 2000–2009. *MMWR Morb Mortal Wkly Rep.* 2012;61:830–833. https://www.cdc.gov/mmwr/preview/mmwrhtml/mm6141a2.htm. Consultado el 6 de junio de 2017.)

En la tabla 4.4 se expone una clasificación de las causas de muerte en Estados Unidos en 2014 por APVP, junto con las tasas de mortalidad específicas de causa, ajustadas por edad. Según la tasa de mortalidad específica de causa, el suicidio ocupa el séptimo lugar, pero según los APVP ocupa la cuarta posición. Esto refleja el hecho de que una gran proporción de muertes por suicidio se producen en personas jóvenes.

Tabla 4.4

Estimación de los años potenciales de vida perdidos (APVP) antes de los 75 años y tasas de mortalidad ajustadas por edad por 100.000 habitantes, por causa de muerte, Estados Unidos, 2014

Enfermedad	Tasa de mortalidad ajustada por edad	APVP en 2014
Cardiopatía	167	3.130.959
Cáncer	161,2	4.416.968

Enfermedades crónicas del tracto respiratorio inferior	40,5	596.470
Lesiones por accidentes involuntarios	40,5	3.146.798
Diabetes mellitus	20,9	562.659
Gripe y neumonía	15,1	293.372
Suicidio	13	1.206.515
Septicemia	12,2	263.766
Hepatopatía crónica y cirrosis	12	581.980
Hipertensión	9,5	130.533

APVP, años potenciales de vida perdidos.

Datos de Centers for Disease Control and Prevention. Deaths: final data for 2014. *Natl Vital Stat Rep.* 2016; 65(4):1–122.

Los APVP pueden ser útiles para tres importantes funciones de salud pública: el establecimiento de prioridades de recursos e investigación, la vigilancia de tendencias temporales en la mortalidad prematura y la evaluación de la eficacia de programas de intervención[2].

¿Por qué nos interesa la mortalidad?

La mortalidad claramente sirve de índice de la gravedad de una enfermedad desde los puntos de vista clínico y de salud pública, pero también puede utilizarse como índice del riesgo de una enfermedad, como se muestra en las figuras 4.2 y 4.3. En general, para una enfermedad dada, los datos de mortalidad son más fáciles de obtener que los datos de incidencia, por lo que puede ser más factible emplear los datos de mortalidad como indicador indirecto de la incidencia. Sin embargo, cuando una enfermedad es leve y no mortal, la mortalidad no es un buen índice de incidencia. La tasa de mortalidad es un buen reflejo de la tasa de incidencia en dos situaciones: en primer lugar, cuando la tasa de letalidad es alta (como en la rabia no tratada), y, en segundo lugar, cuando la duración de la enfermedad (supervivencia) es corta. Bajo estas condiciones, la mortalidad es una buena medida de la incidencia y, por tanto, una medida del riesgo de la enfermedad. Por ejemplo, el cáncer de páncreas es una enfermedad muy letal: la muerte generalmente ocurre a los pocos meses del diagnóstico y la supervivencia a largo plazo es rara. Desafortunadamente, la mortalidad por cáncer de páncreas es, por tanto, un buen sustituto de la incidencia de la enfermedad.

La figura 4.11 muestra las tendencias de mortalidad en Estados Unidos desde 1980 hasta 2014 por raza. Resulta evidente que las tasas de mortalidad de los individuos blancos y negros han disminuido, pero aun así existe una clara disparidad entre las dos razas, como lo demuestra la separación constante entre las dos curvas. La figura 4.12 muestra las tendencias de mortalidad en Estados Unidos de 1955 a 2014 por género y grupo de edad. En ambos gráficos para hombres y mujeres, podemos observar que hay una disminución constante en la tasa de mortalidad a lo largo de los años, particularmente en los grupos de edad menores de 14 años. Esto podría

atribuirse potencialmente a la cobertura generalizada de las vacunas infantiles. Por otro lado, la disminución fue modesta en los grupos de edad de 45 a 64 años debido a las mejoras en la detección temprana de enfermedades cardiovasculares y cáncer, y la evolución de nuevos tratamientos efectivos. Si nos fijamos en el gráfico izquierdo para hombres, vemos un aumento en la tasa de mortalidad para los grupos de edad de 25 a 44 años en la década de 1980, seguido de un fuerte descenso a principios de la década de 1990. Esto puede explicarse por los casos de enfermedad entonces emergente por el virus de inmunodeficiencia humana (VIH), seguido por los tratamientos antirretrovirales de gran actividad recién introducidos, así como por cambios en el estilo de vida resultantes de medidas de educación en salud pública.

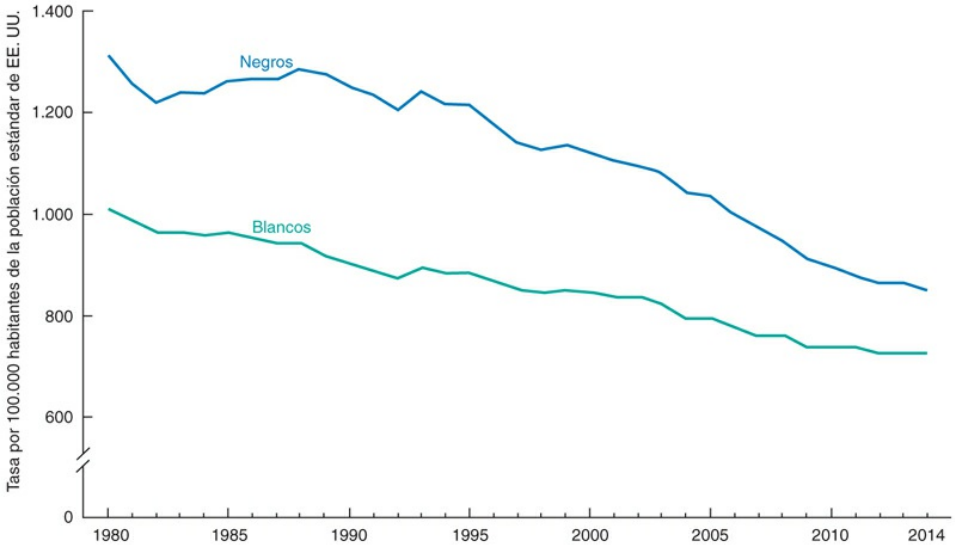
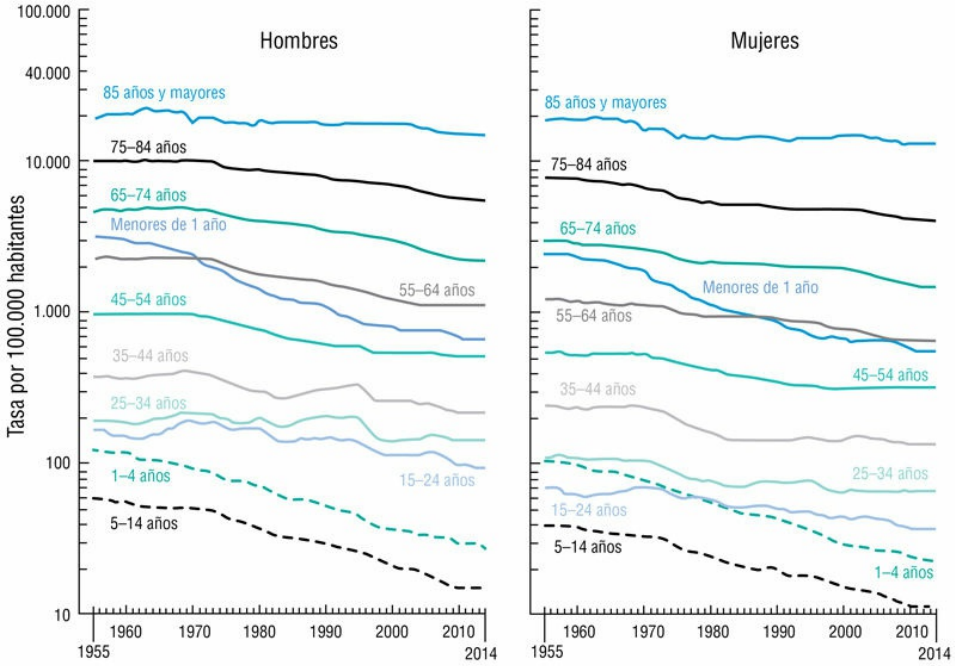


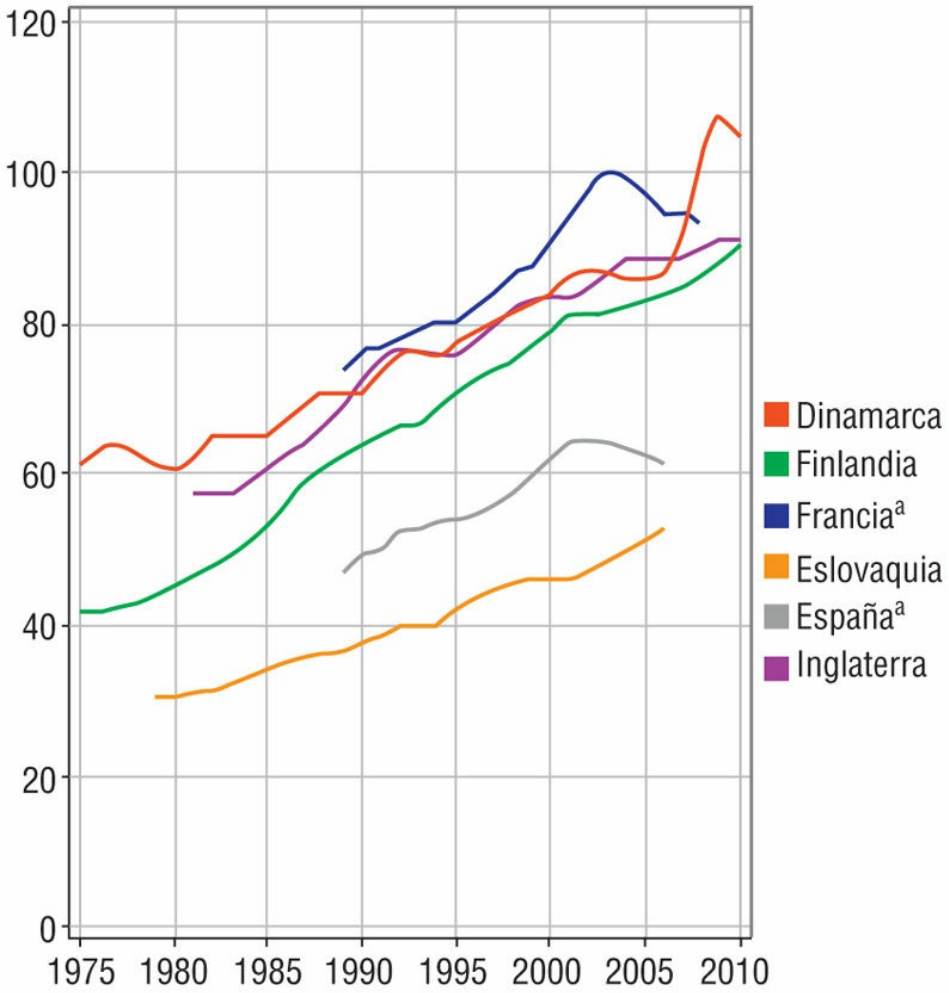
FIG. 4.11 Tasas de mortalidad ajustadas por edad, por raza: Estados Unidos, 1980–2014. (De Kochanek KD, Murphy SL, Xu J, Tejada-Vera B. Deaths: final data for 2014. *Natl Vital Stat Rep.* 2016;65:1–122.)



[a]Las tasas se basan en estimaciones de la población, que difieren de las tasas de mortalidad infantil (basadas en los nacidos vivos).

FIG. 4.12 Tasas de mortalidad, por edad y género: Estados Unidos, 1995-2014. (De Kochanek KD, Murphy SL, Xu J, Tejada-Vera B. Deaths: final data for 2014. *Natl Vital Stat Rep.* 2016;65:1–122.)

En la figuras 4.13 y 4.14 se muestra una comparación de la mortalidad y la incidencia. La figura 4.13 muestra las tasas de cáncer de mama por año en determinados países europeos desde 1975 hasta 2010. Durante este periodo, las tasas estandarizadas por edad por 100.000 habitantes aumentaron en todos los países que se muestran en la figura. Este aumento se ha atribuido a la detección temprana y a las mejoras diagnósticas. Sin embargo, como se observa en la figura 4.14, las tasas de mortalidad por cáncer de mama en países seleccionados disminuyeron notablemente a partir de la década de 1990, quizás como resultado de una detección más temprana y una intervención médica y quirúrgica cada vez más rápida.



[a]Datos regionales

NORDCAN (www.ancr.nu)
ECO (eco.iarc.fr)
England: www.ons.gov.uk

FIG. 4.13 Tendencias en la incidencia del cáncer de mama femenino en países seleccionados: tasa estandarizada por edad (W) por 100.000 en países europeos seleccionados, 1975–2010. (De International Agency for Research on Cancer, GLOBOCAN; 2012. http://globocan.iarc.fr/old/FactSheets/cancers/breast-new.asp. Consultado el 7 de junio de 2017.)

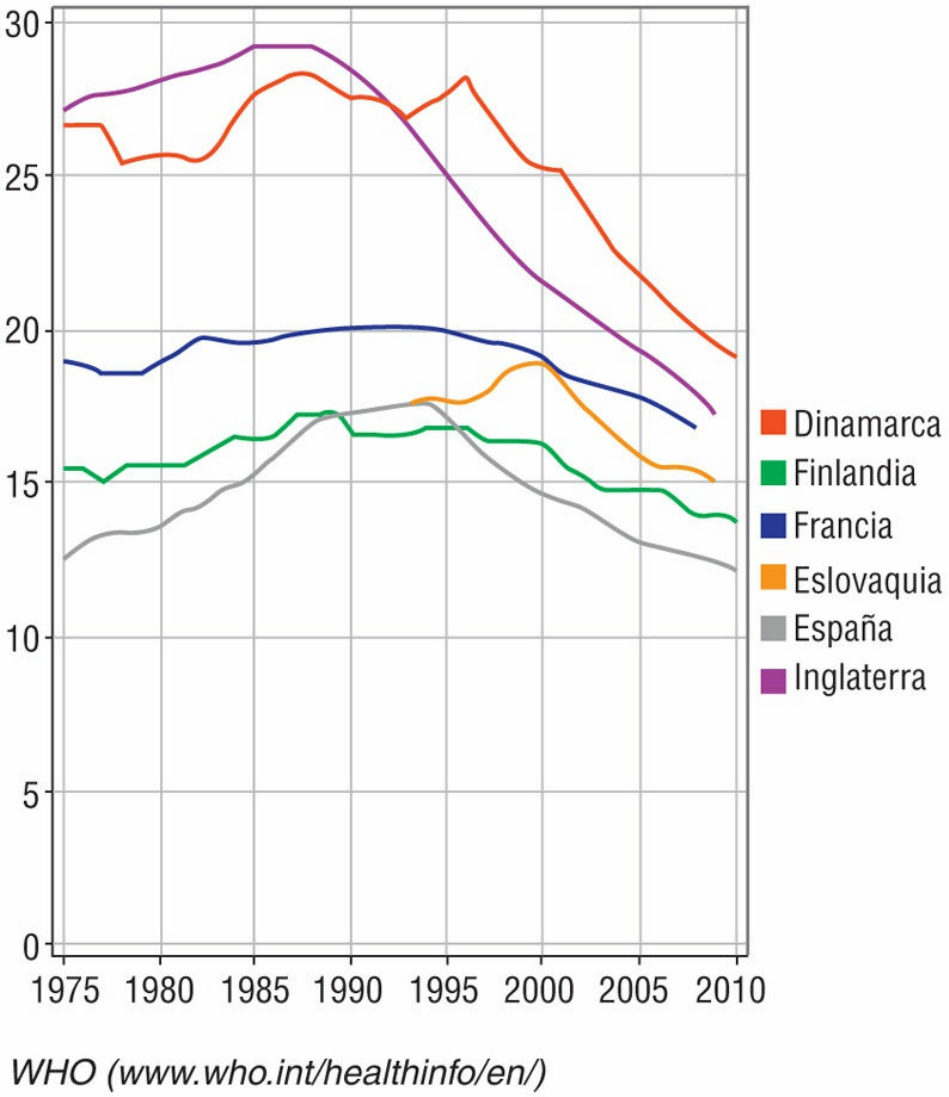


FIG. 4.14 Tendencias en la mortalidad del cáncer de mama femenino en países seleccionados: tasa estandarizada por edad por 100.000 en países europeos seleccionados, 1975–2010. (De International Agency for Research on Cancer, GLOBOCAN; 2012. http://globocan.iarc.fr/old/FactSheets/cancers/breast-new.asp. Consultado el 7 de junio de 2017.)

La figura 4.15 presenta datos recientes sobre las tendencias temporales de la incidencia y la mortalidad del cáncer de mama en las mujeres negras y en las mujeres blancas en Estados Unidos. Comparemos las tendencias temporales de la incidencia y la mortalidad. ¿Qué nos dicen estas curvas acerca de los nuevos casos de cáncer de mama a lo largo del tiempo y de la supervivencia del cáncer de mama? Comparemos los resultados de incidencia y mortalidad en las mujeres negras y en las blancas. ¿Cómo podemos describir las diferencias y cuáles pueden ser algunas de las posibles explicaciones?

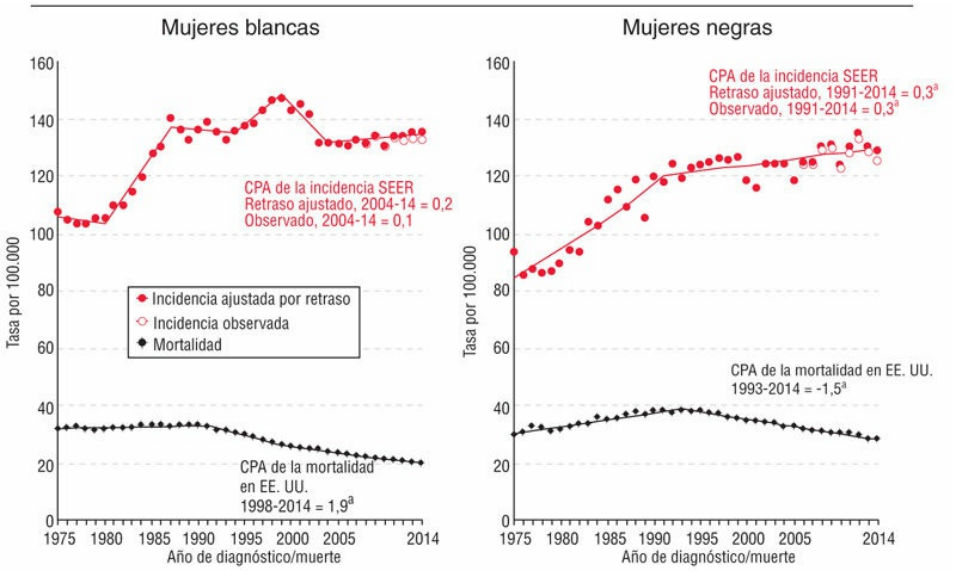


FIG. 4.15 Incidencia y mortalidad del cáncer de mama: mujeres blancas frente a mujeres negras. *CPA,* cambio porcentual anual; *SEER,* Surveillance, Epidemiology, and End Results. (De SEER Cancer Statistics Review; 1975–2014. National Cancer Institute, Bethesda, MD. https://seer.cancer.gov/csr/1975_2014/browse_csr.php?sectionSEL=4&pageSEL=sect_04_zfig.01.html. Consultado el 7 de junio de 2017.)

Un último ejemplo hace referencia a informes de años recientes que indican que la incidencia de cáncer de tiroides en Estados Unidos ha estado aumentando. Una de las dos posibles explicaciones es probable. La primera explicación postula que estos informes reflejan un aumento verdadero en la incidencia, que se debe al aumento de la prevalencia de factores de riesgo de la enfermedad. La segunda explicación dice que el aumento comunicado de la incidencia es solo un aumento de la incidencia *aparente*. No refleja un aumento verdadero de casos nuevos sino un aumento en la detección precoz y el diagnóstico de casos subclínicos, porque los nuevos métodos diagnósticos nos permiten identificar cánceres de tiroides pequeños y asintomáticos que no podían detectarse previamente.

Con el fin de diferenciar entre estas dos posibles explicaciones, Lim y cols.[3] estudiaron los cambios en la incidencia y la mortalidad por cáncer de tiroides en Estados Unidos desde 1974 hasta 2013. En la figura 4.16 se muestra cómo durante el periodo del estudio la *tasa de incidencia* de cáncer de tiroides (A) aumentó más del doble, pero durante el mismo periodo la *mortalidad* debida al cáncer de tiroides (B) permaneció prácticamente sin cambios.

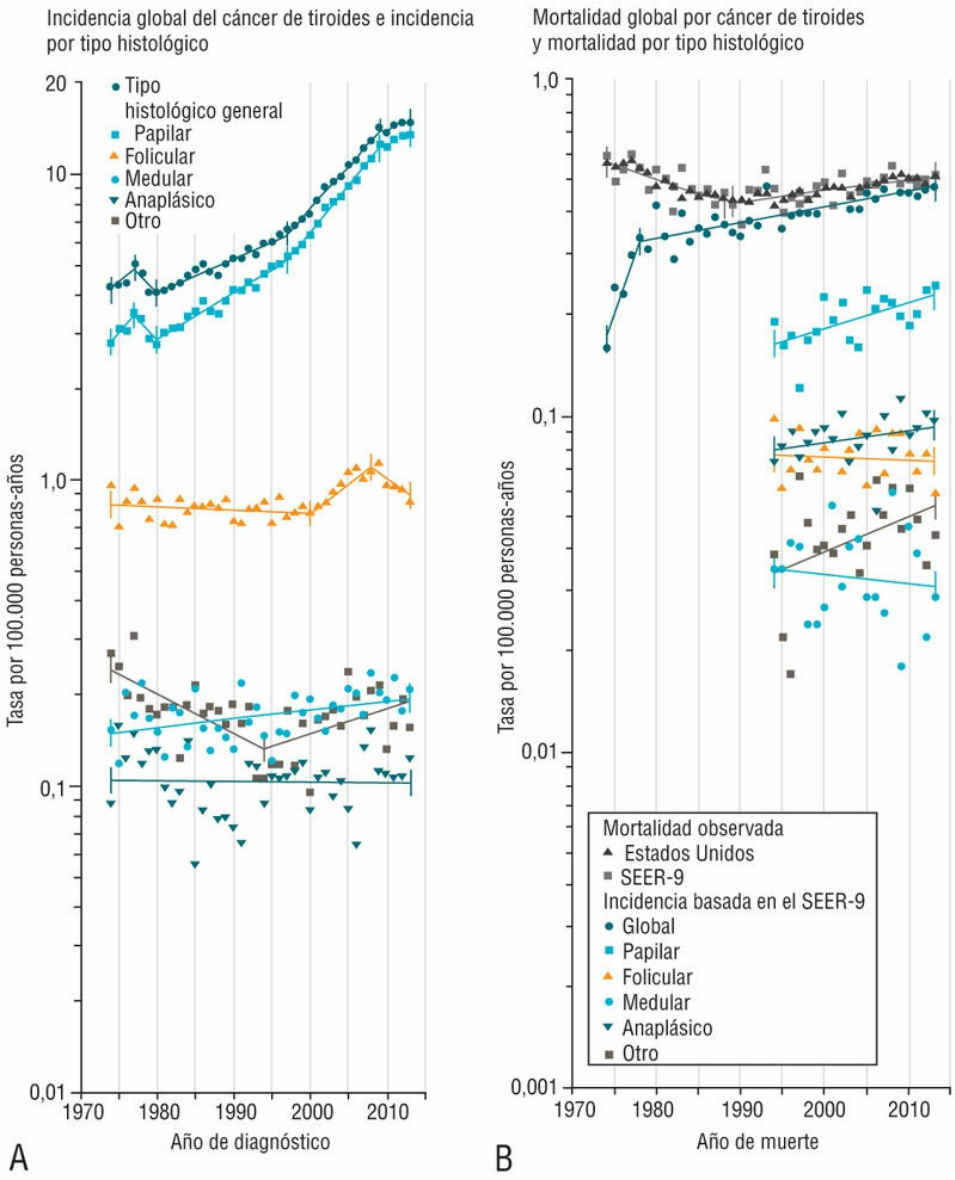


FIG. 4.16 Incidencia (A) y mortalidad (B) por cáncer de tiroides, 1974-2013. *SEER,* Surveillance, Epidemiology, and End Results. (De Lim H, Devesa SS, Sosa JA, et al. Trends in thyroid cancer incidence and mortality in the United States, 1974–2013. *JAMA.* 2017;317:1338–1348.)

El cáncer de tiroides se caracteriza por diferentes tipos histológicos, como se observa en la figura 4.17; en un extremo, el carcinoma papilar posee el mejor pronóstico y, en el extremo opuesto, los tipos mal diferenciados (el cáncer medular y el anaplásico) son generalmente los más agresivos y de peor pronóstico. Los autores encontraron que el aumento en la incidencia del cáncer de tiroides se debía casi por completo a un aumento en la incidencia del cáncer papilar (fig. 4.18). Dentro de los cánceres papilares, la mayor parte del aumento de su incidencia se debía a tumores de pequeño tamaño (fig. 4.19). Así, los autores encontraron que el 87% del aumento en la incidencia del cáncer de tiroides a lo largo de un periodo de 30 años se debía al aumento de los cánceres papilares de menor tamaño, los de mejor pronóstico. Diversos trabajos anteriores demostraron en estudios de autopsias una prevalencia elevada de cánceres papilares pequeños

asintomáticos, no diagnosticados previamente. Si el aumento de la incidencia se debiese a la disponibilidad de métodos diagnósticos más precisos, sería esperable observar un aumento en la incidencia de tumores pequeños, que es exactamente lo que los autores descubrieron en su estudio.

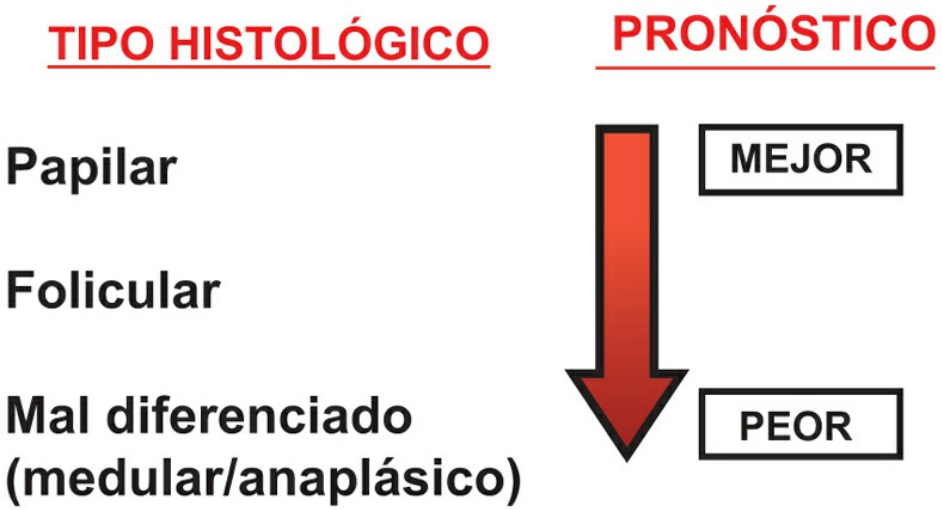


FIG. 4.17 Tipos histológicos de cáncer de tiroides y sus pronósticos.

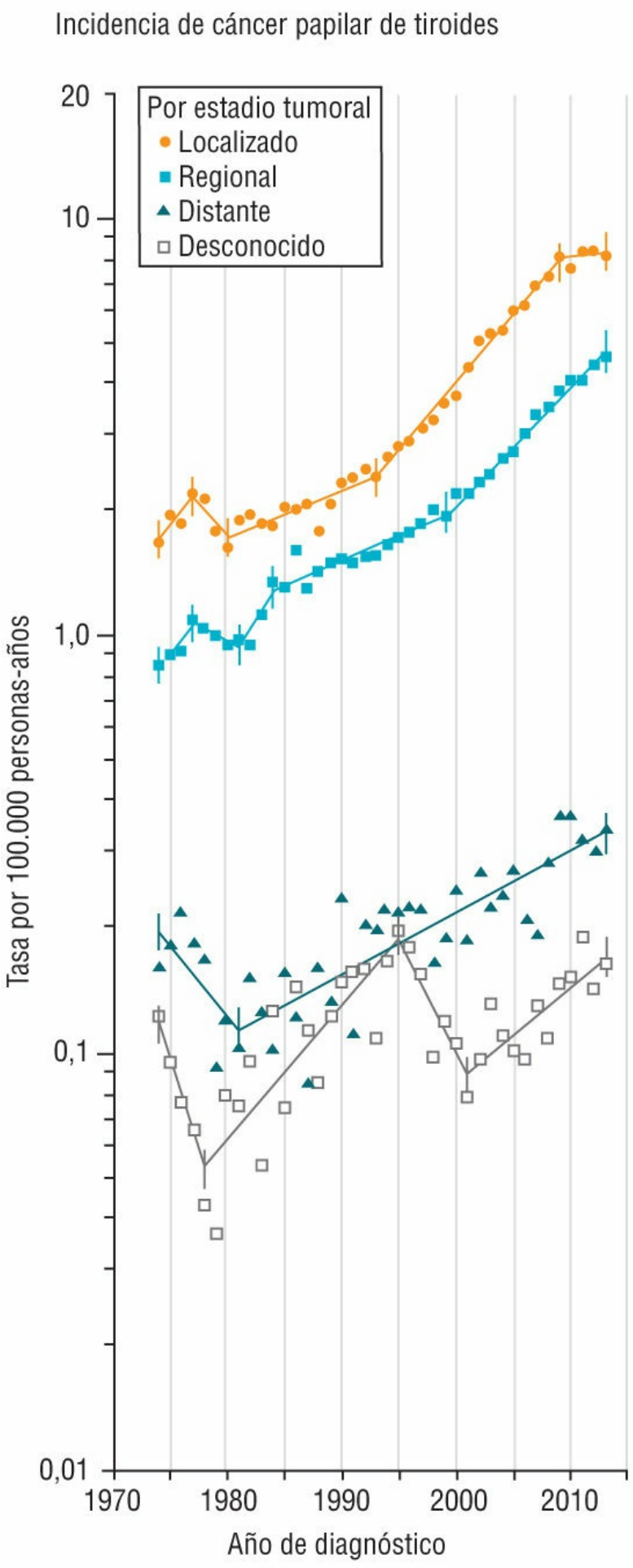

Incidencia de cáncer papilar de tiroides
Por estadio tumoral
Localizado
Regional
Distante
Desconocido
20
10
1,0
0,1
0,01
Tasa por 100.000 personas-años
1970
1980
1990
2000
2010
Año de diagnóstico

FIG. 4.18 Tendencias en la incidencia del cáncer de tiroides por estadio tumoral (1974-2013) en Estados Unidos. (De Lim H, Devesa SS, Sosa JA, et al. Trends in thyroid cancer incidence and mortality in the United States, 1974–2013. *JAMA.* 2017;317:1338–1348.)

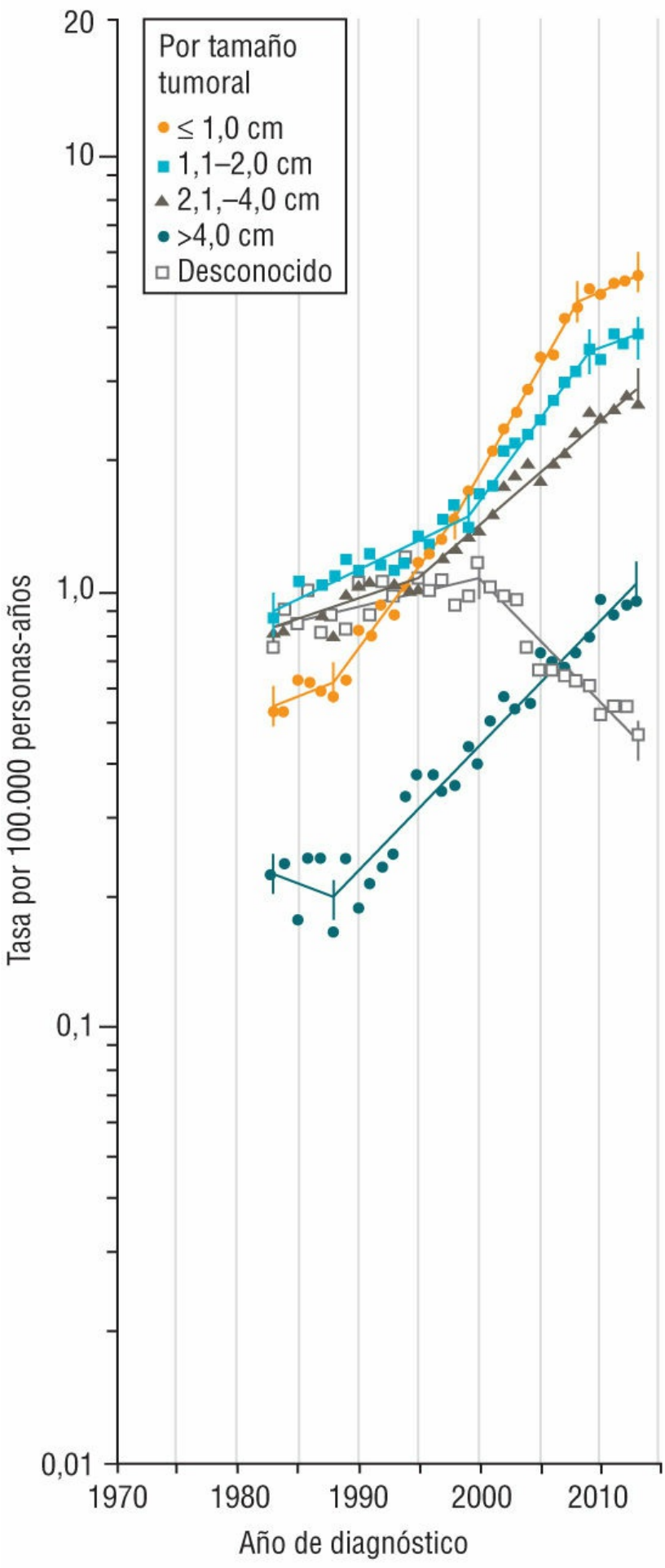

Por tamaño tumoral
≤ 1,0 cm
1,1–2,0 cm
2,1,–4,0 cm
>4,0 cm
Desconocido
20
10
1,0
0,1
0,01
Tasa por 100.000 personas-años
1970
1980
1990
2000
2010
Año de diagnóstico

FIG. 4.19 Tendencias en la incidencia de tumores papilares del tiroides, por tamaño, Estados Unidos, 1983-2013. (De Lim H, Devesa SS, Sosa JA, et al. Trends in thyroid cancer incidence and mortality in the United States, 1974–2013. *JAMA*. 2017;317:1338–1348.)

Problemas con los datos de mortalidad

La mayor parte de nuestra información sobre las muertes proviene de los certificados de defunción. En la figura 4.20 se muestra un certificado de defunción. Por acuerdo internacional, las muertes se clasifican según su *causa subyacente*. La causa subyacente de la muerte se define como «la enfermedad o la lesión que inició la cascada de acontecimientos mórbidos que condujeron directa o indirectamente a la muerte o las circunstancias del accidente o del acto violento que produjeron la lesión mortal»[4]. Así, el certificado de defunción mostrado en la figura 4.21 sería codificado como una muerte debida a cardiopatía isquémica crónica, la causa subyacente, que siempre se encuentra en la última línea del apartado 32 de la parte I del certificado. La causa subyacente de muerte, por tanto, «excluye información relacionada con la causa inmediata de muerte, causas contribuyentes y aquellas causas que intervienen entre la causa subyacente y la causa inmediata de la muerte»[5]. Como señalaron Savage y cols.[6], la contribución total de una causa determinada de muerte puede no verse reflejada en los datos de mortalidad tal y como se comunican generalmente; esto puede aplicarse más a unas enfermedades que a otras.



U.S. STANDARD CERTIFICATE OF DEATH

LOCAL FILE NO. STATE FILE NO.

NAME OF DECEDENT
For use by physician or institution
To Be Completed/ Verified By: FUNERAL DIRECTOR

1. DECEDENT'S LEGAL NAME (Include AKA's if any) (First, Middle, Last) | 2. SEX | 3. SOCIAL SECURITY NUMBER

4a. AGE-Last Birthday (Years) | 4b. UNDER 1 YEAR Months Days | 4c. UNDER 1 DAY Hours Minutes | 5. DATE OF BIRTH (Mo/Day/Yr) | 6. BIRTHPLACE (City and State or Foreign Country)

7a. RESIDENCE-STATE | 7b. COUNTY | 7c. CITY OR TOWN

7d. STREET AND NUMBER | 7e. APT. NO. | 7f. ZIP CODE | 7g. INSIDE CITY LIMITS? ☐ Yes ☐ No

8. EVER IN US ARMED FORCES? ☐ Yes ☐ No | 9. MARITAL STATUS AT TIME OF DEATH ☐ Married ☐ Married, but separated ☐ Widowed ☐ Divorced ☐ Never Married ☐ Unknown | 10. SURVIVING SPOUSE'S NAME (If wife, give name prior to first marriage)

11. FATHER'S NAME (First, Middle, Last) | 12. MOTHER'S NAME PRIOR TO FIRST MARRIAGE (First, Middle, Last)

13a. INFORMANT'S NAME | 13b. RELATIONSHIP TO DECEDENT | 13c. MAILING ADDRESS (Street and Number, City, State, Zip Code)

14. PLACE OF DEATH (Check only one: see instructions)

IF DEATH OCCURRED IN A HOSPITAL: ☐ Inpatient ☐ Emergency Room/Outpatient ☐ Dead on Arrival

IF DEATH OCCURRED SOMEWHERE OTHER THAN A HOSPITAL: ☐ Hospice facility ☐ Nursing home/Long term care facility ☐ Decedent's home ☐ Other (Specify):

15. FACILITY NAME (If not institution, give street & number) | 16. CITY OR TOWN , STATE, AND ZIP CODE | 17. COUNTY OF DEATH

18. METHOD OF DISPOSITION: ☐ Burial ☐ Cremation ☐ Donation ☐ Entombment ☐ Removal from State ☐ Other (Specify): | 19. PLACE OF DISPOSITION (Name of cemetery, crematory, other place)

20. LOCATION-CITY, TOWN, AND STATE | 21. NAME AND COMPLETE ADDRESS OF FUNERAL FACILITY

22. SIGNATURE OF FUNERAL SERVICE LICENSEE OR OTHER AGENT | 23. LICENSE NUMBER (Of Licensee)

To Be Completed By: MEDICAL CERTIFIER

ITEMS 24-28 MUST BE COMPLETED BY PERSON WHO PRONOUNCES OR CERTIFIES DEATH | 24. DATE PRONOUNCED DEAD (Mo/Day/Yr) | 25. TIME PRONOUNCED DEAD

26. SIGNATURE OF PERSON PRONOUNCING DEATH (Only when applicable) | 27. LICENSE NUMBER | 28. DATE SIGNED (Mo/Day/Yr)

29. ACTUAL OR PRESUMED DATE OF DEATH (Mo/Day/Yr) (Spell Month) | 30. ACTUAL OR PRESUMED TIME OF DEATH | 31. WAS MEDICAL EXAMINER OR CORONER CONTACTED? ☐ Yes ☐ No

CAUSE OF DEATH (See instructions and examples)

32. **PART I.** Enter the chain of events--diseases, injuries, or complications--that directly caused the death. DO NOT enter terminal events such as cardiac arrest, respiratory arrest, or ventricular fibrillation without showing the etiology. DO NOT ABBREVIATE. Enter only one cause on a line. Add additional lines if necessary. | Approximate interval: Onset to death

IMMEDIATE CAUSE (Final disease or condition ——> resulting in death) a. ____ Due to (or as a consequence of):

Sequentially list conditions, if any, leading to the cause listed on line a. Enter the **UNDERLYING CAUSE** (disease or injury that initiated the events resulting in death) **LAST**
b. ____ Due to (or as a consequence of):
c. ____ Due to (or as a consequence of):
d. ____

PART II. Enter other significant conditions contributing to death but not resulting in the underlying cause given in PART I | 33. WAS AN AUTOPSY PERFORMED? ☐ Yes ☐ No | 34. WERE AUTOPSY FINDINGS AVAILABLE TO COMPLETE THE CAUSE OF DEATH? ☐ Yes ☐ No

35. DID TOBACCO USE CONTRIBUTE TO DEATH? ☐ Yes ☐ Probably ☐ No ☐ Unknown

36. IF FEMALE: ☐ Not pregnant within past year ☐ Pregnant at time of death ☐ Not pregnant, but pregnant within 42 days of death ☐ Not pregnant, but pregnant 43 days to 1 year before death ☐ Unknown if pregnant within the past year

37. MANNER OF DEATH ☐ Natural ☐ Homicide ☐ Accident ☐ Pending Investigation ☐ Suicide ☐ Could not be determined

38. DATE OF INJURY (Mo/Day/Yr) (Spell Month) | 39. TIME OF INJURY | 40. PLACE OF INJURY (e.g., Decedent's home; construction site; restaurant; wooded area) | 41. INJURY AT WORK? ☐ Yes ☐ No

42. LOCATION OF INJURY: State: City or Town: Street & Number: Apartment No.: Zip Code:

43. DESCRIBE HOW INJURY OCCURRED: | 44. IF TRANSPORTATION INJURY, SPECIFY: ☐ Driver/Operator ☐ Passenger ☐ Pedestrian ☐ Other (Specify)

45. CERTIFIER (Check only one):
☐ Certifying physician-To the best of my knowledge, death occurred due to the cause(s) and manner stated.
☐ Pronouncing & Certifying physician-To the best of my knowledge, death occurred at the time, date, and place, and due to the cause(s) and manner stated.
☐ Medical Examiner/Coroner-On the basis of examination, and/or investigation, in my opinion, death occurred at the time, date, a nd place, and due to the cause(s) and manner stated.
Signature of certifier:

46. NAME, ADDRESS, AND ZIP CODE OF PERSON COMPLETING CAUSE OF DEATH (Item 32)

47. TITLE OF CERTIFIER | 48. LICENSE NUMBER | 49. DATE CERTIFIED (Mo/Day/Yr) | 50. **FOR REGISTRAR ONLY**- DATE FILED (Mo/Day/Yr)

To Be Completed By: FUNERAL DIRECTOR

51. DECEDENT'S EDUCATION-Check the box that best describes the highest degree or level of school completed at the time of death.
☐ 8th grade or less
☐ 9th - 12th grade; no diploma
☐ High school graduate or GED completed
☐ Some college credit, but no degree
☐ Associate degree (e.g., AA, AS)
☐ Bachelor's degree (e.g., BA, AB, BS)
☐ Master's degree (e.g., MA, MS, MEng, MEd, MSW, MBA)
☐ Doctorate (e.g., PhD, EdD) or Professional degree (e.g., MD, DDS, DVM, LLB, JD)

52. DECEDENT OF HISPANIC ORIGIN? Check the box that best describes whether the decedent is Spanish/Hispanic/Latino. Check the "No" box if decedent is not Spanish/Hispanic/Latino.
☐ No, not Spanish/Hispanic/Latino
☐ Yes, Mexican, Mexican American, Chicano
☐ Yes, Puerto Rican
☐ Yes, Cuban
☐ Yes, other Spanish/Hispanic/Latino (Specify)

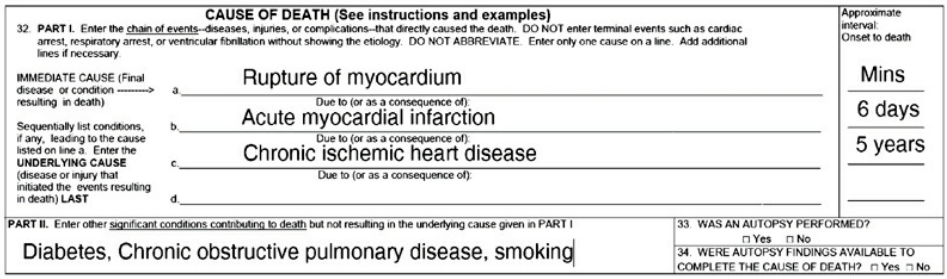
53. DECEDENT'S RACE (Check one or more races to indicate what the decedent considered himself or herself to be)
☐ White
☐ Black or African American
☐ American Indian or Alaska Native (Name of the enrolled or principal tribe)
☐ Asian Indian
☐ Chinese
☐ Filipino
☐ Japanese
☐ Korean
☐ Vietnamese
☐ Other Asian (Specify)
☐ Native Hawaiian
☐ Guamanian or Chamorro
☐ Samoan
☐ Other Pacific Islander (Specify)
☐ Other (Specify)

54. DECEDENT'S USUAL OCCUPATION (Indicate type of work done during most of working life. DO NOT USE RETIRED).

55. KIND OF BUSINESS/INDUSTRY

REV. 11/2003

FIG. 4.20 Certificado de defunción estándar en EE.UU. (De Centers for Disease Control and Prevention. https://www.cdc.gov/nchs/data/dvs/death11-03final-acc.pdf. Consultado el 7 de junio de 2017.)


CAUSE OF DEATH (See instructions and examples)

32. PART I. Enter the chain of events--diseases, injuries, or complications--that directly caused the death. DO NOT enter terminal events such as cardiac arrest, respiratory arrest, or ventricular fibrillation without showing the etiology. DO NOT ABBREVIATE. Enter only one cause on a line. Add additional lines if necessary.

Approximate interval: Onset to death

IMMEDIATE CAUSE (Final disease or condition ---------> resulting in death)
a. Rupture of myocardium — Mins

Due to (or as a consequence of):

Sequentially list conditions, if any, leading to the cause listed on line a. Enter the UNDERLYING CAUSE (disease or injury that initiated the events resulting in death) LAST
b. Acute myocardial infarction — 6 days

Due to (or as a consequence of):
c. Chronic ischemic heart disease — 5 years

Due to (or as a consequence of):
d.

PART II. Enter other significant conditions contributing to death but not resulting in the underlying cause given in PART I

Diabetes, Chronic obstructive pulmonary disease, smoking

33. WAS AN AUTOPSY PERFORMED? ☐ Yes ☐ No

34. WERE AUTOPSY FINDINGS AVAILABLE TO COMPLETE THE CAUSE OF DEATH? ☐ Yes ☐ No

FIG. 4.21 Ejemplo de la sección completada de causa de muerte de un certificado de defunción; se incluyen las causas inmediata y subyacente y otras patologías importantes.

La calidad de los datos proporcionados en los certificados de defunción varía ampliamente según los países y las regiones. Los estudios de validez de los certificados de defunción comparados con los registros hospitalarios y de autopsias generalmente encuentran mayor validez para ciertas enfermedades, como cánceres, que para otras.

Las muertes se codifican siguiendo la Clasificación Internacional de Enfermedades (CIE), ahora en su décima revisión. Como las categorías de codificación y las regulaciones cambian de una revisión a otra, ante cualquier estudio sobre tendencias temporales de mortalidad que abarque más de una revisión se debe valorar la posibilidad de que los cambios observados se deban en su totalidad o en parte a cambios en la CIE. En 1949 se observó una reducción espectacular de las tasas de mortalidad por diabetes en varones y mujeres (fig. 4.22). Sin embargo, la euforia que desataron estos datos fue de corta duración, ya que el análisis de esta disminución indicó que se produjo en un momento de cambio de la séptima a la octava revisión de la CIE. Antes de 1949, la política era que todo certificado de defunción que mencionase la diabetes en cualquier apartado se codificase como una muerte debida a la diabetes. A partir de 1949, solo los certificados de defunción en los que la causa subyacente de muerte era la diabetes eran codificados como muertes por diabetes. Así pues, la reducción observada en la figura 4.22 era un artefacto del cambio en la codificación. Siempre que se observe una tendencia temporal de aumento o reducción de la mortalidad, la primera pregunta que debemos plantearnos es: «¿Es real?». Específicamente, cuando analizamos las tendencias en la mortalidad a lo largo del tiempo, debemos preguntarnos si se produjeron cambios en la forma de codificar los certificados de defunción durante el periodo analizado y si estos cambios podrían haber contribuido a los cambios observados en la mortalidad durante el mismo periodo.

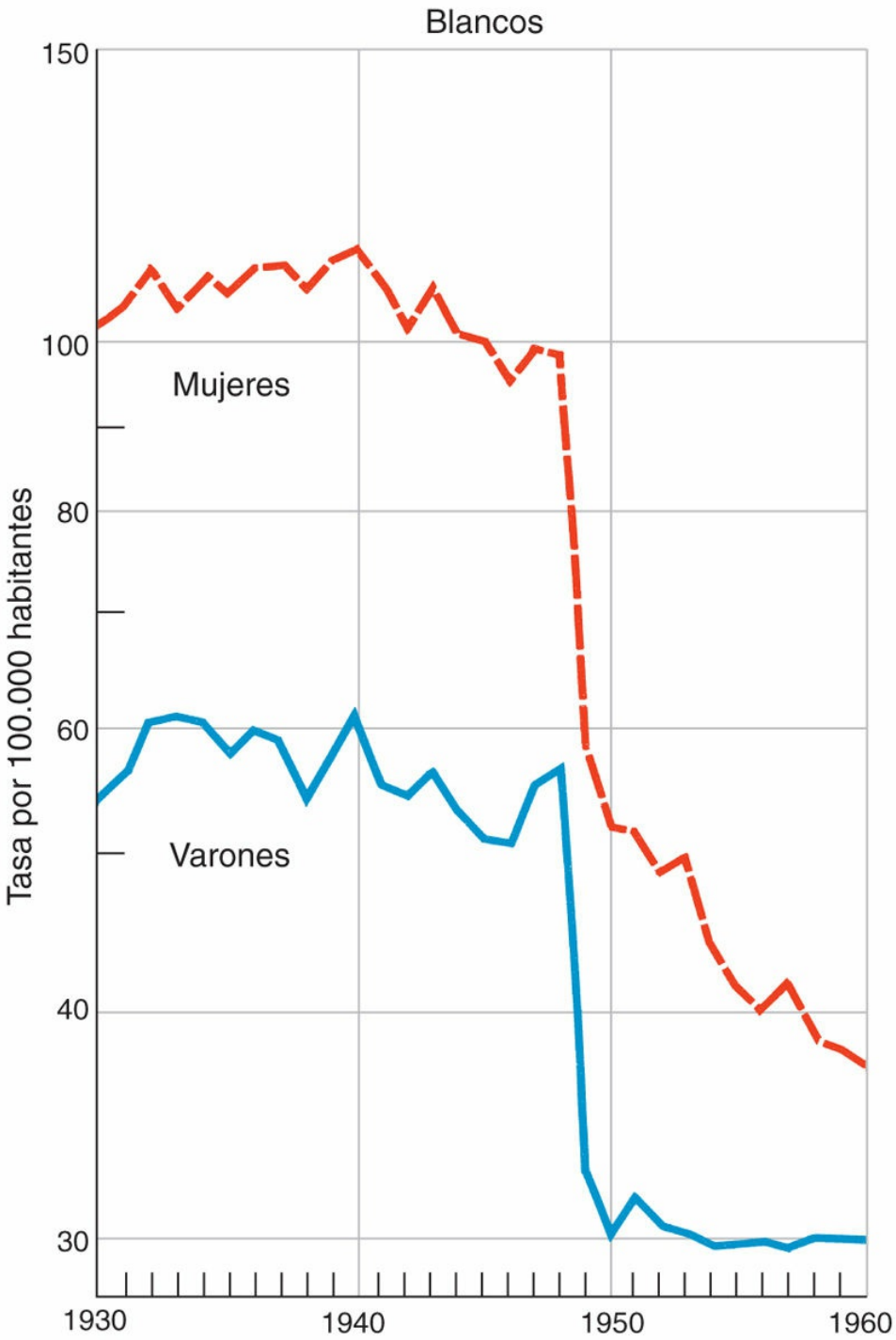


FIG. 4.22 Disminución de las tasas de mortalidad por diabetes en varones y mujeres de 55-64 años, Estados Unidos, 1930-60, debido a cambios en los códigos de la clasificación internacional de enfermedades. (De *US Public Health Service Publication No. 1000, series 3, No. 1*. Washington, DC: U.S. Government Printing Office, 1964.)

Los cambios en la definición de una enfermedad también pueden ejercer un efecto importante sobre el número de casos de la enfermedad que son comunicados o que son comunicados y posteriormente clasificados como que reúnen los criterios diagnósticos de la enfermedad. A principios de 1993 se estableció una nueva definición del síndrome de inmunodeficiencia adquirida (SIDA); como se muestra en la figura 4.23, este cambio resultó en un rápido

aumento del número de casos comunicados. Con la nueva definición, incluso tras el repunte inicial, el número de casos comunicados permaneció siendo más elevado que el comunicado durante varios años.

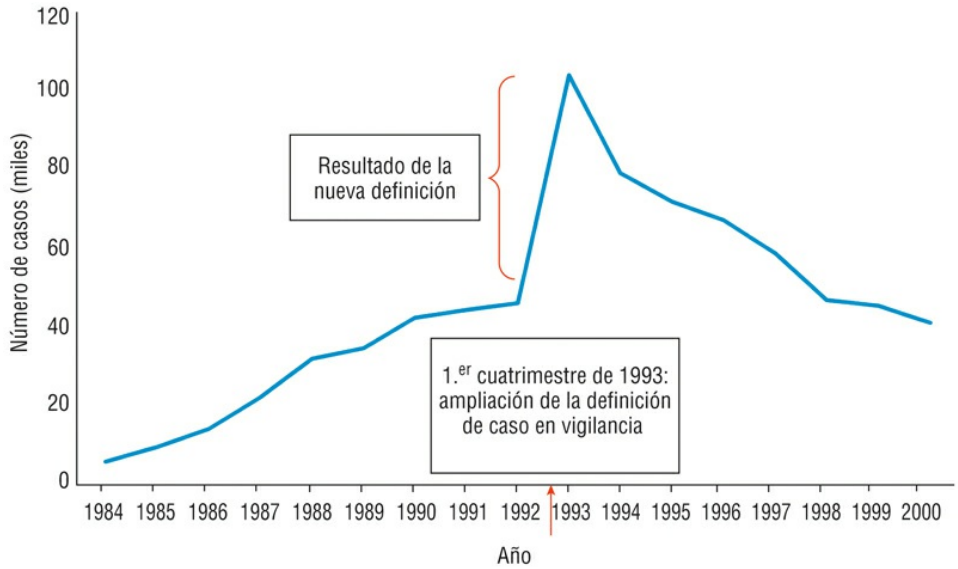


FIG. 4.23 Casos de síndrome de inmunodeficiencia adquirida por cuatrimestre, Estados Unidos, 1984-2000. (De Centers for Disease Control and Prevention. Summary of notifiable diseases, United States, 2000. *MMWR.* 2000;49:86; y Centers for Disease Control and Prevention: Summary of notifiable diseases, United States, 1993. *MMWR.* 1993;45:68.)

Al analizar la morbilidad en el capítulo 3, dijimos que todos los sujetos del grupo representado por el denominador deben tener riesgo de entrar en el grupo representado en el numerador, y expusimos las tasas de incidencia de cáncer cervical uterino como ejemplo. El mismo principio respecto al numerador y al denominador es de aplicación para las tasas de mortalidad. En la figura 4.24 se muestra un grupo similar de observaciones para las tasas de mortalidad por cáncer cervical uterino. De nuevo, la corrección en función de la presencia de una histerectomía reduce el número de mujeres en el denominador y, por tanto, aumenta la tasa de mortalidad. En un tono más ligero, en el cuadro 4.2 se enumeran algunas causas de muerte que fueron reflejadas en certificados de defunción a principios del siglo xx.

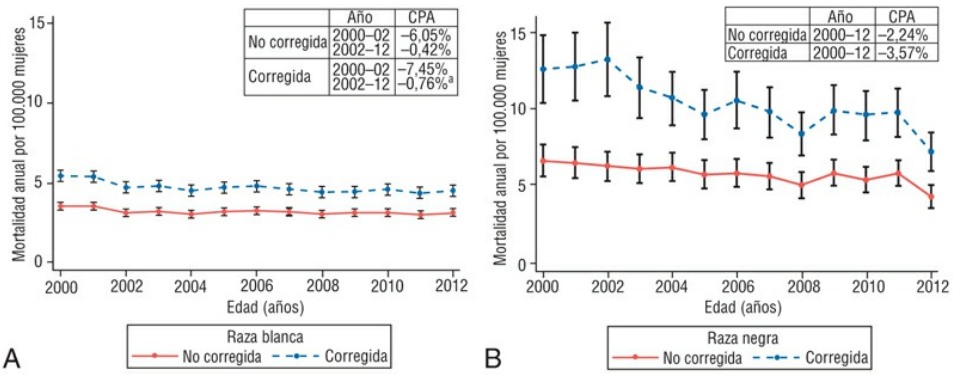


FIG. 4.24 Tendencias en las tasas de mortalidad por cáncer de cuello uterino estandarizadas por edad, no corregidas y corregidas por la prevalencia de histerectomía, desde 2000 a 2012 para (A) mujeres blancas y (B) mujeres negras. *CPA,* cambio porcentual anual. (De Beavis AL, Gravitt PE, Rositch AF. Hysterectomy-corrected cervical cancer mortality rates reveal a larger racial disparity in the United States. *Cancer.* 2017;123:1044–1050.)

Cuadro 4.2 Algunas causas de muerte comunicadas en certificados de defunción a principios del siglo xx

«Murió de repente sin la ayuda de un médico»
«Una madre murió en la infancia»
«El fallecido nunca había tenido una enfermedad mortal»
«Murió de repente, nada grave»
«Se acostó bien pero se despertó muerto»

Comparación de la mortalidad en diferentes poblaciones

Un uso importante de los datos de mortalidad es la comparación de dos o más poblaciones, o de una población en diferentes periodos de tiempo. Dichas poblaciones pueden diferir con respecto a muchas características que afectan a la mortalidad, de las que la distribución por edades es la más importante. De hecho, la edad es el factor predictivo aislado más importante de mortalidad. Así, se han desarrollado métodos para comparar la mortalidad en dichas poblaciones, a la vez que de un modo efectivo se mantienen características constantes como la edad.

En la tabla 4.5 se muestran datos que ejemplifican este problema. Se exponen las tasas de mortalidad de los residentes blancos y negros del estado de Maryland en 2015. Los datos pueden parecer sorprendentes porque sería de esperar que las tasas fuesen más elevadas en la población negra debido a los problemas asociados con las peores condiciones de vida y el peor acceso a la asistencia médica, en especial en dicha época. En la tabla 4.6 vuelven a aparecer los datos de la tabla 4.5 a la izquierda, pero ahora se han añadido datos para cada grupo (estrato) de edad específico de la población. Resulta interesante que, aunque en cada grupo de edad específico la mortalidad es más elevada en la población negra que en la blanca, la mortalidad global (también denominada *mortalidad bruta o no ajustada*) es más elevada en la población blanca que en la negra. ¿Cuál es la explicación? Esto es un reflejo del hecho de que tanto en la población blanca como en la negra la mortalidad aumenta de manera importante en los grupos de edad más avanzada, que son los que más contribuyen a la mortalidad. Sin embargo, la población blanca de este ejemplo es de mayor edad que la población negra, y en 2015 existían menos negros en los grupos de edad más avanzada. Por tanto, en la población blanca, la mortalidad global está claramente sesgada por las elevadas tasas en los grupos de edad más avanzada. La tasa de mortalidad global (o bruta) de la población blanca está aumentada por el mayor número de muertes en el numeroso subgrupo de sujetos blancos de edad avanzada, pero la tasa de mortalidad global en la población negra no es tan elevada porque se producen muchos menos fallecimientos en el menor número de sujetos negros pertenecientes a grupos de edad más avanzada. Claramente, la mortalidad bruta refleja diferencias en la fuerza de la mortalidad y diferencias en la composición de la población por grupos etarios. Fijémonos en dos abordajes para tratar este problema: el ajuste por edad directo e indirecto.

Tabla 4.5

Tasas de mortalidad bruta por raza, estado de Maryland, Estados Unidos 2015

Raza	Mortalidad por 1.000 habitantes
Blanca	9,95
Negra	7,35

Tabla 4.6

Tasas de mortalidad por edad y raza, estado de Maryland, 2015

TASAS DE MORTALIDAD POR EDAD POR 1.000 HABITANTES[A]												
Raza	Todas las edades	<1 año	1-4 años	5-14 años	15-24 años	25-34 años	35-44 años	45-54 años	55-64 años	65-74 años	75-84 años	>85 años
Blanca	9,95	4,06	0,21	0,11	0,64	1,29	1,73	3,62	7,68	16,45	45,39	138,7
Negra	7,35	11,25	0,43	0,18	1,14	1,74	2,23	5,09	11,14	21,55	49,49	124,45

[a] Ajustada por edad según la población de Estados Unidos en el año 2000.

De Maryland Vital Statistics Annual Report; 2015. https://health.maryland.gov/vsa/Documents/15annual.pdf. Consultado el 8 de junio de 2017. Algunos datos fueron proporcionados por la Vital Statistics Administration, Maryland Department of Health, Baltimore, Maryland. El Departamento no se responsabiliza de ningún análisis, interpretación o conclusión.

Ajuste por edad directo

Las tablas 4.7 a 4.9 muestran un ejemplo hipotético de ajuste por edad directo. En la tabla 4.7 se muestra la mortalidad en una población en dos periodos de tiempo diferentes. La tasa de mortalidad es considerablemente superior en el periodo tardío. Estos datos son complementados con datos específicos de edad en la tabla 4.8. En esta tabla observamos tres grupos de edad, y la mortalidad específica de edad es menor en cada grupo en el periodo tardío. ¿Cómo es posible entonces justificar la mortalidad global más alta en el periodo tardío en este ejemplo?

Tabla 4.7

Ejemplo hipotético de ajuste directo por edad: I. Comparación de las tasas de mortalidad totales en una población en dos periodos diferentes

PERIODO TEMPRANO			PERIODO TARDÍO		
Población	Número de muertes	Tasa de mortalidad por 100.000	Población	Número de muertes	Tasa de mortalidad por 100.000
900.000	862	96	900.000	1.130	126

Tabla 4.8

Ejemplo hipotético de ajuste directo por edad: II. Comparación de las tasas de mortalidad específicas de edad en dos periodos diferentes

Grupo de edad (años)	PERIODO TEMPRANO			PERIODO TARDÍO		
	Población	Número de muertes	Tasas de mortalidad por 100.000	Población	Número de muertes	Tasas de mortalidad por 100.000
Todas las edades	900.000	862	96	900.000	1.130	126
30-49	500.000	60	12	300.000	30	10
50-69	300.000	396	132	400.000	400	100
+70	100.000	406	406	200.000	700	350

Tabla 4.9

Ejemplo hipotético de ajuste directo por edad: III. Práctica de un ajuste por edad empleando e total de las dos poblaciones como el estándar

Grupo de edad (años)	Población estándar	Tasas de mortalidad específicas de edad «temprana» por 100.000	Número esperado de muertes empleando tasas «tempranas»	Tasas de mortalidad específicas de edad «tardía» por 100.000	Número esperado de muert empleanc tasas «tardías»
Todas las edades	1.800.000				
30-49	800.000	12	96	10	80
50-69	700.000	132	924	100	700
+70	300.000	406	1.218	350	1.050
Número total de muertes esperadas en la población estándar:			2.238		1.830
Tasas ajustadas por edad:		$\text{«Temprana»} = \frac{2.238}{1.800.000} = 124,3$		$\text{«Tardía»} = \frac{1.830}{1.800.000} = 101,7$	

La respuesta se encuentra en la estructura cambiante de la edad de la población. La mortalidad es más elevada en los grupos de edad más avanzada, y durante el periodo tardío, el tamaño del grupo de mayor edad se duplicó de 100.000 a 200.000, mientras que el número de personas jóvenes disminuyó de manera importante, de 500.000 a 300.000. Nos gustaría eliminar esta diferencia de edad, y de hecho preguntar: ¿si la composición de edades de la población fuese la misma, habría alguna diferencia en la mortalidad entre el periodo temprano y el periodo tardío?

En el *ajuste por edad directo* se emplea una población estándar para eliminar los efectos de cualquier diferencia de edad entre las dos o más poblaciones

que se comparan (v. tabla 4.9). Se crea una población «estándar» hipotética a la que aplicamos la tasa de mortalidad específica de edad en el periodo temprano y en el periodo tardío. Al aplicar las tasas de mortalidad de ambos periodos a una sola población estándar, eliminamos toda posibilidad de que las diferencias observadas puedan ser el resultado de diferencias en la edad de la población. (En este ejemplo hemos creado una población estándar sumando las poblaciones de los periodos temprano y tardío, pero podría haberse empleado cualquier población.)

Al aplicar cada tasa de mortalidad específica de edad a la población en cada grupo de edad de la población estándar, deducimos el número esperado de muertes que habrían ocurrido si se hubiesen aplicado esas tasas. A continuación, podemos calcular el número total de muertes esperadas en la población estándar si se hubiesen aplicado las tasas específicas de edad del periodo temprano y el número total de muertes esperadas en la población estándar si se hubiesen aplicado las tasas específicas de edad del periodo tardío. Dividiendo cada uno de estos dos números de muertes esperadas totales por la población estándar total, podemos calcular una tasa de mortalidad esperada en la población estándar si hubiese tenido la experiencia de mortalidad del periodo temprano y la tasa de mortalidad esperada para la población estándar si hubiese tenido la experiencia de mortalidad para el periodo tardío. Estas son las denominadas *tasas ajustadas por edad,* que reflejan apropiadamente la disminución observada de las tasas específicas de edad. Las diferencias en la composición de edades de la población ya no representan un factor.

En este ejemplo, las tasas han sido ajustadas por edad, pero el ajuste puede llevarse a cabo para cualquier característica, como el sexo, la situación socioeconómica o la raza, y existen técnicas disponibles para ajustar múltiples variables simultáneamente.

Aunque las tasas ajustadas por edad pueden resultar muy útiles para realizar comparaciones, el primer paso al examinar y analizar datos comparativos de mortalidad debe ser siempre examinar cuidadosamente las tasas *específicas de edad* de cualquier diferencia o cambio interesante. Estas diferencias pueden verse ocultas en las tasas *ajustadas por edad* y pueden perderse si procedemos de inmediato al ajuste por edad sin examinar primero las tasas específicas de edad.

Las tasas ajustadas por edad son hipotéticas porque implican aplicar tasas específicas de edad reales a una población estándar hipotética. No reflejan el riesgo verdadero de mortalidad de una población «real» porque el valor numérico de una tasa de mortalidad ajustada por edad depende de la población estándar empleada. La selección de dicha población es algo arbitraria porque no existe una población estándar «correcta», aunque por lo general se acepta que la «estándar» no debería ser muy diferente de las poblaciones que están siendo comparadas con respecto a la edad o a la variable empleada para realizar el ajuste. En Estados Unidos, durante más de 50 años se utilizó habitualmente la población estadounidense de 1940 como la

población estándar para el ajuste por edad para la mayoría de los fines, pero, en los últimos años, esta población se ha considerado cada vez más obsoleta e incompatible con la estructura de edad más avanzada de la población estadounidense. Comenzando con las estadísticas sobre mortalidad de 1999, la población estadounidense del año 2000 sustituyó a la población de 1940 como la población estándar para realizar ajustes.

El cambio de la población estándar a la población estadounidense del año 2000 ha tenido algunos efectos importantes, como se ilustra en una comparación de tasas de mortalidad específicas de causa utilizando datos de 1995[7]. Entre estos se observaron aumentos en las tasas de mortalidad ajustadas por edad para causas en las que el riesgo aumenta significativamente con la edad. Por ejemplo, la mortalidad ajustada por edad de enfermedades cerebrovasculares (ictus) es de 26,7 muertes por 100.000 habitantes utilizando la población estándar de 1940, pero si se emplea la población estándar del año 2000 es de 63,9 por 100.000. La mortalidad por cáncer aumentó empleando la población estándar del año 2000 respecto a cuándo se utilizaba una población anterior como estándar porque más personas sobreviven hasta edades más avanzadas, cuando muchos de los principales tipos de cáncer son más comunes. Las tasas para las cardiopatías, la enfermedad pulmonar obstructiva crónica, la diabetes, las nefropatías y la enfermedad de Alzheimer se vieron afectadas de un modo similar porque las tasas de mortalidad específicas de edad para todos estos trastornos son más elevadas en los grupos de edad más avanzada.

Las tasas de cáncer ajustadas por edad son más elevadas en la población negra que en la población blanca en Estados Unidos, pero la diferencia entre negros y blancos es menor con la población estándar del año 2000 que con la población estándar anterior. Por tanto, el cambio a la población estadounidense del año 2000 como la población estándar complica las comparaciones de las tasas ajustadas por edad antes y después de 1999, porque muchas de las tasas anteriores a 1999 se calcularon empleando la población estándar de 1940. Sin embargo, las tasas de 1999 y posteriores se calculan utilizando la población del año 2000 como el nuevo estándar.

En resumen, el objetivo del ajuste directo es comparar tasas en al menos dos poblaciones diferentes cuando deseamos eliminar el efecto posible de un factor determinado, como la edad, sobre las tasas que estamos comparando. Es importante tener en cuenta que las tasas ajustadas no son tasas «reales» en las poblaciones que se están comparando, porque dependen de la elección de la población estándar utilizada para realizar el ajuste. Sin embargo, el ajuste directo es una herramienta muy útil para realizar dichas comparaciones y, de hecho, la comparación de las tasas en diferentes poblaciones casi siempre utiliza el ajuste directo, como el ajuste por edad. Tenga en cuenta que el ajuste se basa en reemplazar cada población con un conjunto común de ponderaciones (la población estándar) para estimar los promedios ponderados, es decir, las tasas ajustadas.

Ajuste por edad indirecto (razón de mortalidad estandarizada)

El *ajuste por edad indirecto* a menudo se utiliza cuando las cifras de muertes para cada estrato específico de edad no se encuentran disponibles. También se utiliza para estudiar la mortalidad en una población con una exposición ocupacional. ¿La mortalidad de las personas que trabajan en una cierta industria, como la minería o la construcción, es superior a la de las personas de la misma edad de la población general? ¿Dicha ocupación supone un riesgo adicional?

Para responder a la pregunta de si la población de trabajadores posee una mortalidad más elevada que la esperada en una población similar que no trabaja en la ocupación que se está estudiando, a cada grupo de edad de la población de interés se aplican tasas específicas de edad para una población conocida, como la de todos los varones de la misma edad. Esto proporcionará el número de muertes esperadas en cada grupo de edad en la población de interés si esta población hubiese tenido la experiencia de mortalidad de la población conocida. Por tanto, se calcula el número de muertes *esperadas* para cada grupo de edad, y estas cifras se suman. También se calcula y se suma el número de muertes que fueron realmente *observadas* en dicha población. A continuación, se calcula la proporción entre el número total de muertes observadas realmente y el número total de muertes esperadas, si la población de interés hubiera tenido la experiencia de mortalidad de la población conocida. Esta proporción se denomina *razón de mortalidad estandarizada (RME)*.

La RME se define como sigue:

$$\text{RME} = \frac{\text{N.º observado de muertes por año}}{\text{N.º esperado de muertes por año}}$$

Fijémonos en el ejemplo de la tabla 4.10. En una población hipotética de 460.463 varones blancos trabajadores, 406 murieron de enfermedad X en 2016. El tema en el que estamos interesados es si esta experiencia de mortalidad por enfermedad X es mayor, menor o similar a la esperada en los varones blancos de edades parecidas de la población general (la mayoría de los cuales no están incluidos en esta clasificación de trabajadores). Para ayudar al estudio de este aspecto, podemos calcular el número esperado de muertes en cada grupo de edad de los trabajadores blancos aplicando la tasa de mortalidad específica de edad conocida en la población general al número de trabajadores de cada grupo de edad. Haciendo esto, nos estamos preguntando: «¿Cuántas muertes cabría esperar en estos trabajadores blancos si tuviesen la misma experiencia de mortalidad que los varones blancos del

mismo grupo de edad de la población general?». Estos datos se exponen en la columna 3. En la columna 4 se muestra el número real de muertes observado en los trabajadores.

Tabla 4.10

Cálculo hipotético de la razón de mortalidad estandarizada (RME) para la enfermedad X en trabajadores blancos de 20 a 59 años, 2016

	Población estimada de trabajadores blancos	Tasa de mortalidad (por 100.000) de la enfermedad X en varones en la población general	Muertes esperadas por la enfermedad X en trabajadores blancos si tuviesen los mismos riesgos que la población general	Muertes observadas por la enfermedad en trabajadores blancos
Edad (años)	1	2	3 = 1 × 2	4
20-24	62.253	8,9	5,5	5
25-29	72.732	12,7	9,3	15
30-34	68.500	18,1	12,4	17
35-44	136.525	30,6	41,7	93
45-54	90.304	53,4	48,2	169
55-59	30.149	71,8	21,7	107
Totales	460.463		138.8	406

$$\text{RME (en población de 20 a 59 años)} = \frac{406}{138,8} \times 100 = 292$$

La RME se calcula dividiendo el número total de muertes observadas (406) entre el número esperado de muertes (138,8), lo que resulta en 2,92. A menudo se multiplica por 100 para obtener resultados sin decimales. Si hiciéramos esto en nuestro ejemplo el resultado sería 292. Una RME de 100 indica que el número observado de muertes y el número esperado de muertes son iguales. Una RME mayor de 100 indica que el número observado de muertes es superior al número esperado, y una RME menor de 100 indica que el número observado de muertes es inferior al número esperado.

El efecto cohorte

La tabla 4.11 muestra la prevalencia de la obesidad específica por edad (%) de 1971 a 2006 en Estados Unidos, utilizando datos del National Center for Health Statistics. (Para esta discusión, ignoraremos los datos para los grupos de edad de 2 a 19 años, ya que la obesidad infantil es un fenómeno un tanto diferente.) Si, por ejemplo, leemos la columna de la tabla (los datos para un ciclo dado del National Health and Nutrition Examination Survey [NHANES]) para 1971–75, parece que la prevalencia de la obesidad alcanza su nivel máximo en el grupo de edad de 55 a 59 años y luego disminuye en la población de edad avanzada. Esta forma de ver los datos, por año, se denomina *visión transversal*.

Tabla 4.11

Tabla de contingencia por grupo de edad para la prevalencia de obesidad por edad (filas) y periodo (columnas) en Estados Unidos, 1971–2006 (n = 91.755)

Edad (años)	NHANES I 1971-75	NHANES II 1976-80	NHANES III, fase 1 1988- 91	NHANES III, fase 2 1991- 94	NHANES 99-00 1999-2000	NHANES 01-02 2001-02	NHANES 03-04 2003-2004	NHANES 05-06 2005-06
2-4	3,1	3,25	3,48	4,34	7,02	6,29	8,72	8,54
5-9	5,48	7,17	8,75	13,12	17,45	16,92	20,22	16,25
10-14	6,88	7,9	8,93	13,57	18,97	18,72	22,85	21,81
15-19	6,64	5,5	8,31	13,55	18,03	17,8	19,94	18,43
20-24	6,08	7,14	9,87	13,81	20,59	26,67	26,59	24,98
25-29	10,34	10,49	11,97	18,74	27,69	26,55	26,47	35,9
30-34	13,64	13,49	18,02	20,07	31,64	24,82	30,19	36,6
35-39	14,34	14,73	17,24	23,3	29,08	30,19	36,54	33,3
40-44	16,76	15,82	19,27	24,63	32,68	32,85	39,68	42,69
45-49	15,26	18,05	18,85	30,75	31,93	35,83	35,79	38,5
50-54	17,18	17,46	22,37	35,42	40,55	31,69	39,32	38,73
55-59	19,5	19,62	26,55	32,46	35,7	38	38,62	46,9
60-64	18,68	17,57	20,82	30,67	41,37	44,28	34,49	42,67
65-69	16,83	18,51	21,26	27,79	41,23	35,43	38	40,64
70-74	17,15	16,31	18,68	25,03	29,34	34,87	32,48	31,45

NHANES, National Health and Nutrition Examination Survey.

De Keyes KM, Utz RL, Robinson W, Li G. What is a cohort effect? Comparison of three statistical methods for modeling cohort effects in obesity prevalence in the United States, 1971– 2006. *Soc Sci Med*. 2010;70(7):1100–1108.

Sin embargo, en realidad, la cuestión de la prevalencia de la obesidad es algo diferente (tabla 4.12). Una persona que tenía entre 20 y 24 años de edad en 1971 tenía entre 25 y 29 años de edad en 1976. En otras palabras, las personas nacidas en cierto año se desplazan a la vez a lo largo del tiempo. Ahora podemos examinar la prevalencia de la obesidad a lo largo del tiempo en la misma cohorte (es decir, un grupo de personas que comparten la misma experiencia), nacidos en el mismo periodo de 5 años. Observando a las personas que tenían entre 20 y 24 años de edad en el ciclo 1971-75 y siguiéndolas a lo largo del tiempo, según indican las casillas en negrita de la tabla, es evidente que la prevalencia de obesidad para esta *cohorte* ha aumentado a lo largo de los años y no declinó más adelante, como hemos observado en la vista transversal de los datos. Cuando examinamos los cambios en la prevalencia a lo largo del tiempo, siempre nos debemos preguntar si cualquier cambio aparente observado puede haber sido resultado del efecto cohorte.

Tabla 4.12

Tabla de contingencia por grupo de edad para la prevalencia de obesidad por edad (filas) y periodo (columnas) en Estados Unidos, 1971–2006 (*n* = 91.755)

Edad (años)	NHANES I 1971-75	NHANES II 1976-80	NHANES III, fase 1 1988-91	NHANES III, fase 2 1991-94	NHANES 99-00 1999-2000	NHANES 01-02 2001-02	NHANES 03-04 2003-2004	NHANES 05-06 2005-06
2–4	3,1	3,25	3,48	4,34	7,02	6,29	8,72	8,54
5–9	5,48	7,17	8,75	13,12	17,45	16,92	20,22	16,25
10–14	6,88	7,9	8,93	13,57	18,97	18,72	22,85	21,81
15–19	6,64	5,5	8,31	13,55	18,03	17,8	19,94	18,43
20–24	**6,08**	7,14	9,87	13,81	20,59	26,67	26,59	24,98
25–29	10,34	**10,49**	11,97	18,74	27,69	26,55	26,47	35,9
30–34	13,64	13,49	**18,02**	20,07	31,64	24,82	30,19	36,6
35–39	14,34	14,73	17,24	**23,3**	29,08	30,19	36,54	33,3
40–44	16,76	15,82	19,27	24,63	**32,68**	32,85	39,68	42,69
45–49	15,26	18,05	18,85	30,75	31,93	**35,83**	35,79	38,5
50–54	17,18	17,46	22,37	35,42	40,55	31,69	**39,32**	38,73
55–59	19,5	19,62	26,55	32,46	35,7	38	38,62	**46,9**
60–64	18,68	17,57	20,82	30,67	41,37	44,28	34,49	42,67
65–69	16,83	18,51	21,26	27,79	41,23	35,43	38	40,64
70–74	17,15	16,31	18,68	25,03	29,34	34,87	32,48	31,45

Las casillas en negrita denotan las personas que tenían de 20 a 24 años de edad durante el ciclo 1971-1975 y fueron seguidas a lo largo del tiempo, formando una cohorte.*NHANES*, National Health and Nutrition Examination Survey.

De Keyes KM, Utz RL, Robinson W, Li G. What is a cohort effect? Comparison of three statistical methods for modeling cohort effects in obesity prevalence in the United States, 1971– 2006. *Soc Sci Med*. 2010;70(7): 1100–1108.

Interpretación de los cambios observados en la mortalidad

Si encontramos una diferencia en la mortalidad a lo largo del tiempo o entre poblaciones, ya sea un aumento o una disminución, puede tratarse de un artefacto o ser real. Si se trata de un artefacto, el artefacto puede ser el resultado de problemas con el numerador o el denominador (tabla 4.13). Sin embargo, si concluimos que el cambio es real, ¿cuál puede ser la explicación? En el cuadro 4.3 se muestran algunas posibilidades.

Tabla 4.13

Posibles explicaciones de las tendencias o diferencias en la mortalidad: I. Artefactos

1. Numerador	Errores diagnósticos Errores en la edad Cambios en las normas de codificación Cambios en la clasificación
2. Denominador	Errores en el recuento de la población Errores en la clasificación por características demográficas (p. ej., edad, raza, sexo) Diferencias en los porcentajes de las poblaciones en riesgo

Cuadro 4.3 Posibles explicaciones de las tendencias o diferencias en la mortalidad: II. Reales

Cambios en la supervivencia sin cambios en la incidencia
Cambios en la incidencia
Cambios en la composición de la edad de la o las poblaciones
Una combinación de los factores anteriores

Otras medidas del impacto de la enfermedad

Calidad de vida

La mayoría de las enfermedades producen un impacto importante sobre los pacientes por encima y más allá de la mortalidad. Las enfermedades que no son mortales pueden asociarse con un sufrimiento físico y emocional considerable debido a la discapacidad asociada con la enfermedad. Por tanto, es importante considerar el impacto total de la enfermedad midiendo sus efectos sobre la calidad de vida de las personas, incluso aunque dichas medidas no sean de hecho medidas de la ocurrencia de la enfermedad. Por ejemplo, en los pacientes con artritis es posible examinar el grado de impedimento que supone la enfermedad para realizar sus actividades cotidianas. Aunque existe gran controversia acerca de qué medidas de calidad de vida son las más apropiadas y válidas, por lo general se admite que dichas medidas pueden utilizarse de modo razonable para planificar programas terapéuticos a corto plazo para grupos de pacientes. Dichos pacientes pueden ser evaluados a lo largo de un periodo de meses para determinar los efectos del tratamiento en su calidad de vida, valorada por ellos mismos. Las medidas de calidad de vida también se han empleado para establecer prioridades en situaciones en las que los recursos de asistencia sanitaria son escasos. Aunque la priorización de los recursos de asistencia sanitaria a menudo se basa principalmente en los datos de mortalidad, la calidad de vida también debe tenerse en cuenta para este fin, porque muchas enfermedades son crónicas y no suponen una amenaza para la vida, pero pueden asociarse con muchos años de discapacidad. Los pacientes pueden dar más o menos importancia a diferentes medidas de calidad de vida en función de su ocupación y otras actividades, personalidad, antecedentes culturales, educación y valores éticos y morales. Como resultado, la medición de la calidad de vida y el desarrollo de índices válidos que sean útiles para obtener datos comparativos en diferentes pacientes y en diferentes poblaciones siguen siendo un desafío importante.

Proyección de la carga futura de la enfermedad

Un uso interesante y valioso de los datos actuales para predecir el impacto futuro de la enfermedad fue la valoración exhaustiva de la mortalidad y discapacidad actual asociada con enfermedades, lesiones y factores de riesgo en todas las regiones del mundo en 1990, proyectado hasta el año 2020. El estudio, titulado Global Burden of Disease, intenta cuantificar no solo las muertes sino también el impacto de la muerte prematura y la discapacidad en una población, así como combinar estos aspectos en un solo índice para

expresar la «carga de la enfermedad» en conjunto. El índice desarrollado para este estudio es el de años de vida ajustados por discapacidad (AVAD o DALY, *disability-adjusted life year*), es decir, los años de vida perdidos por una muerte prematura y los años vividos con una discapacidad de gravedad y duración específica. Así, un AVAD es 1 año perdido de vida saludable.

Los resultados mostraron que 5 de las 10 principales causas de discapacidad en 1990 fueron trastornos psiquiátricos; los trastornos psiquiátricos y neurológicos representaban el 28% de todos los años vividos con discapacidad de gravedad y duración conocidas, en comparación con el 1,4% de todas las muertes y el 1,1% de años de vida perdidos. En la figura 4.25 se muestran algunas de las principales causas de carga por enfermedad globalmente, en países con ingresos altos y con ingresos bajos en 2015[9]. De nuevo, resulta muy evidente la importancia de la cardiopatía isquémica en los países con ingresos altos y las infecciones del tracto respiratorio inferior en los países de ingresos bajos.

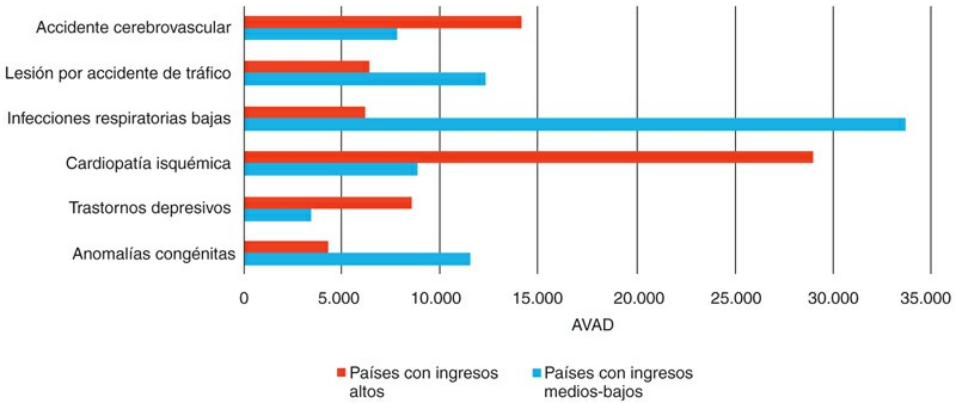


FIG. 4.25 Causas seleccionadas de carga de enfermedad en países de ingresos bajos y altos, 2015. *AVAD,* años de vida ajustados por discapacidad. (De Global Health Estimates 2015. *Disease Burden by Cause, Age, Sex, by Country and by Region, 2000–2015.* Geneva, Switzerland: World Health Organization; 2016.)

En 2015, la carga por enfermedad no se encontraba distribuida equitativamente. Como se observa en la tabla 4.14, las 20 causas principales de carga por enfermedad fueron responsables del 55,7% de los AVAD. Cinco de ellas afectan sobre todo a niños menores de 5 años. Tres de las 10 causas principales (cardiopatía isquémica, enfermedad cerebrovascular y depresión) son trastornos crónicos. Esta tabla muestra el valor de usar una medida como los AVAD para valorar la carga de la enfermedad, una medida que no se limita a la morbilidad o a la mortalidad, sino que valora ambas.

Tabla 4.14

Estimaciones de salud global 2015: 20 causas principales de AVAD globalmente, 2015

Puesto	Causa	AVAD (miles)	% AVAD	AVAD por 100.000 habitantes
0	Todas las causas	2.668.296	100,0	36.331
1	Cardiopatía isquémica	192.056	7,2	2.615
2	Infecciones respiratorias bajas	142.384	5,3	1.939
3	Enfermedad cerebrovascular	139.874	5,2	1.905
4	Complicaciones del nacimiento pretérmino	102.297	3,8	1.393
5	Enfermedades diarreicas	84.928	3,2	1.156
6	Accidentes de tráfico en carretera	76.020	2,8	1.035
7	Enfermedad pulmonar obstructiva crónica	72.815	2,7	991
8	Diabetes mellitus	70.667	2,6	962
9	Asfixia y traumatismo del parto	67.266	2,5	916
10	Anomalías congénitas	64.825	2,4	883
11	VIH/SIDA	62.759	2,4	855
12	Tuberculosis	56.037	2,1	763
13	Trastornos depresivos	54.215	2,0	738
14	Anemia ferropénica	52.080	2,0	709
15	Lumbalgia y cervicalgia	52.016	1,9	708
16	Cirrosis hepática	41.486	1,6	565
17	Cáncer traqueal, bronquial, pulmonar	41.129	1,5	560
18	Paludismo	38.520	1,4	524
19	Nefropatías	38.104	1,4	519
20	Autolesiones	37.672	1,4	513

AVAD, años de vida ajustados por discapacidad; *SIDA*, síndrome de inmunodeficiencia adquirida; *VIH*, virus de la inmunodeficiencia humana.

Datos de World Health Organization. Health statistics and information systems. http://www.who.int/healthinfo/global_burden_disease/estimates/en/index2.ht Consultado el 9 de junio de 2017. De *The Global Burden of Disease: 2004 Update.* Geneva, Switzerland: World Health Organization; 2004.

Con el envejecimiento de la población mundial y los avances en el desarrollo económico, particularmente en los países de ingresos bajos y medios, se está produciendo una «transición epidemiológica», de modo que en 2020 es probable que las enfermedades no contagiosas representen el 70% de todas las muertes en los países desarrollados. Como se expone en la figura 4.26, en 2020 es de esperar que disminuya espectacularmente la carga por enfermedad debida a enfermedades transmisibles, trastornos maternos y perinatales y deficiencias nutricionales (grupo I). Se espera que la carga debida a enfermedades no transmisibles (grupo II) aumente de manera importante, al igual que la carga debida a lesiones (grupo III). Del mismo modo, es esperable que en 2020 la carga por enfermedad atribuible al tabaco

supere a la causada por cualquier enfermedad aislada, claramente una llamada urgente a la acción de los servicios de salud pública. Aunque no existe consenso universal sobre la metodología o la aplicabilidad de una sola medida de carga por enfermedad como los AVAD, este estudio es una demostración excelente de un intento de vigilancia mundial diseñado para desarrollar las medidas que permitan comparaciones regionales válidas y proyecciones futuras de modo que puedan desarrollarse las intervenciones apropiadas.

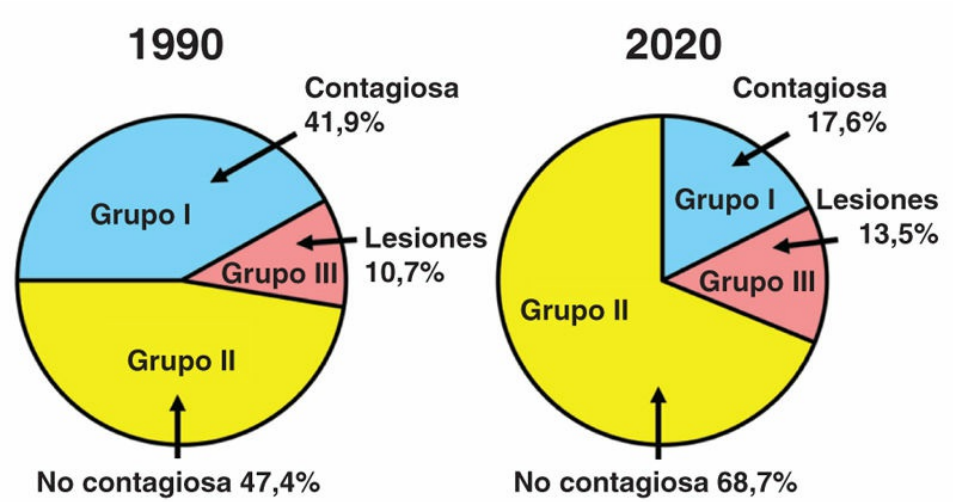


FIG. 4.26 La «transición epidemiológica»: distribución de muertes por causas contagiosas y no contagiosas en países en vías de desarrollo en 1990 y proyección en 2020. (De Murray CJL, Lopez AD. *The Global Burden of Disease: A Comprehensive Assessment of Mortality and Disability from Diseases, Injuries, and Risk Factors in 1990 and Projected to 2020.* Cambridge, Harvard University Press on behalf of the World Health Organization and the World Bank, 1996.)

Conclusión

En los capítulos 3 y 4 hemos revisado abordajes importantes para expresar y medir cuantitativamente la morbilidad y la mortalidad humana. Los conceptos revisados en estos capítulos pueden parecer a primera vista abrumadores (fig. 4.27), pero, como veremos en siguientes capítulos, resultan fundamentales para comprender cómo nos ayuda la epidemiología a esclarecer las medidas del riesgo de la enfermedad, la determinación de la causa de la enfermedad y la evaluación de la eficacia de intervenir para modificar el proceso de la enfermedad.

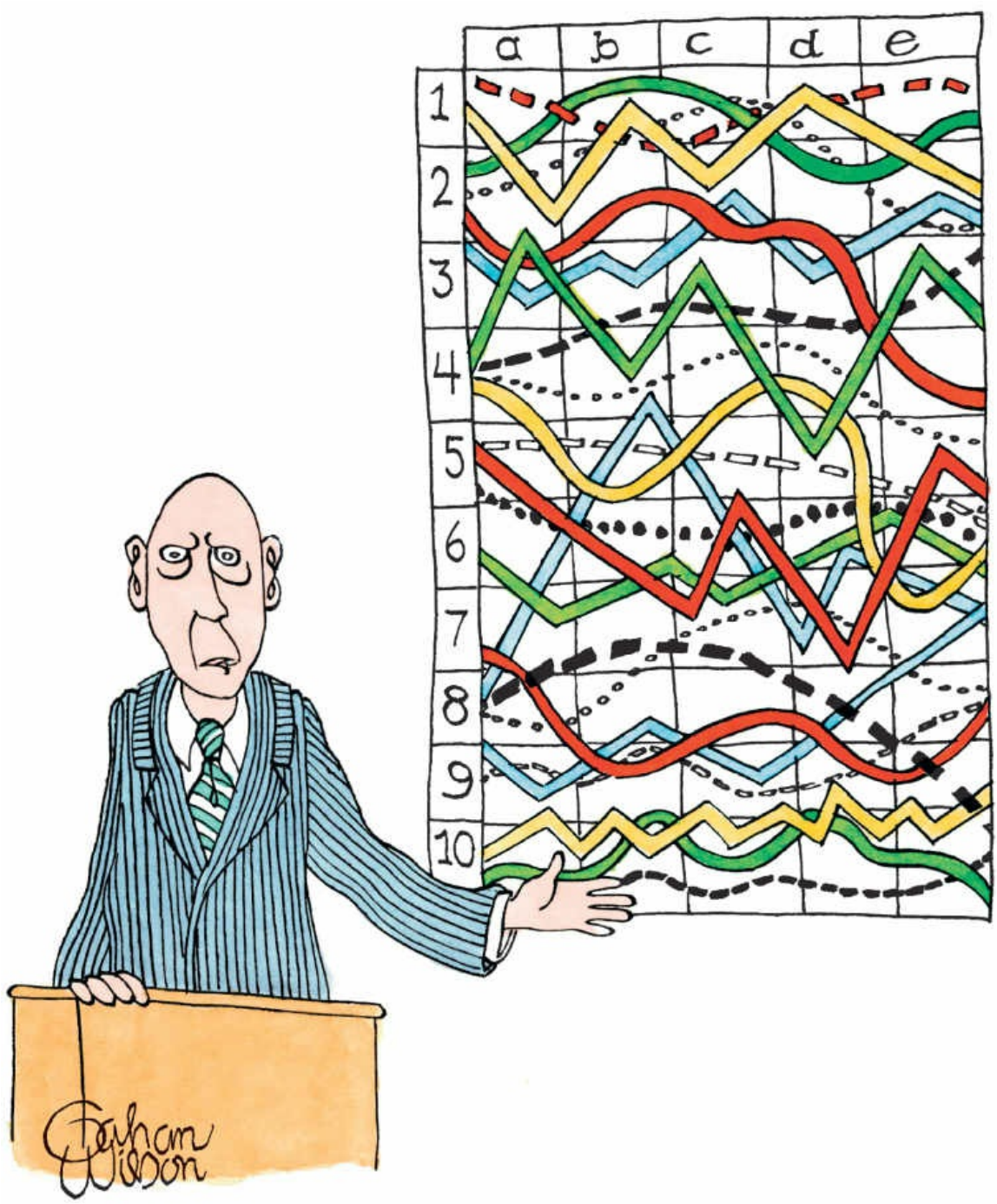


FIG. 4.27 «Haré una pequeña pausa para que puedan asimilar esta información.» (Gahan Wilson/The New Yorker Collection/The Cartoon Bank.)

En el capítulo 5 abordaremos aspectos sobre los numeradores de las tasas de morbilidad: ¿cómo identificamos a aquellos sujetos que padecen una enfermedad y los diferenciamos de los que no la tienen? y ¿cómo evaluamos la calidad de las pruebas diagnósticas y de cribado que se utilizan para separar a estos sujetos y a las poblaciones? En el capítulo 18 se analiza el uso

de las pruebas de cribado en los programas de salud pública.

Bibliografía

1. De Beauvoir S. *A Very Easy Death*. Translated by Patrick O'Brian. New York: Pantheon Books; 1965.
2. Centers for Disease Control and Prevention Premature mortality in the United States: public health issues in the use of years of potential life lost. *MMWR Suppl*. 1986;35:1S–11S.
3. Lim H, Devesa SS, Sosa JA, et al. Trends in thyroid cancer incidence and mortality in the United States, 1974-2013. *JAMA*. 2017;317(13):1338–1348.
4. National Center for Health Statistics (NCHS).Instructions for Classifying the Underlying Cause of Death. Hyattsville, MD: NCHS; 1983.
5. Chamblee RF, Evans MC. *TRANSAX: The NCHS System for Producing Multiple Cause-of-Death Statistics, 1968–1978*. Vital and Health Statistics, Series 1, No. 20, DHHS Publication No. (PHS) 86–1322. Washington, DC: Bureau of Vital and Health Statistics; June 1986.
6. Savage G, Rohde FC, Grant B, et al. *Liver Cirrhosis Mortality in the United States, 1970–90: Surveillance Report No. 29*. Bethesda, MD: Department of Health and Human Services; December 1993.
7. Anderson RN, Rosenberg HM. *Age Standardization of Death Rates: Implementation of the Year 2000 Standard*. National Vital Statistics Reports, Vol. 47, No. 3, pp. 1–16. Hyattsville, MD: National Center for Health Statistics; October 7, 1998.
8. Murray CJL, Lopez AD. The Global Burden of Disease. Cambridge, MA: Harvard University Press; 1996.
9. World Health Organization (WHO). The Global Burden of Disease: 2015 Update. Geneva, Switzerland: WHO; 2015.

Preguntas de repaso del capítulo 4

Las preguntas 1 y 2 se basan en la siguiente información:

En un país asiático con una población de 6 millones de personas, se produjeron 60.000 muertes durante el año que acabó el 31 de diciembre de 2010. Entre las mismas se encuentran 30.000 muertes por cólera de 100.000 personas que estaban enfermas de cólera.

1. *¿Cuál fue la tasa de mortalidad específica de causa del cólera en 2010?*

2. *¿Cuál fue la tasa de letalidad del cólera en 2010?*

3. *Las tasas de mortalidad ajustadas por edad se usan para:*

a. Corregir los errores en las tasas de mortalidad debidos a la edad.

b. Determinar el número real de muertes que se produjeron en grupos de edad específicos en una población.

c. Corregir tasas de mortalidad cuando falta información sobre la edad.

d. Comparar las muertes en personas del mismo grupo de edad.

e. Eliminar los efectos de las diferencias en las distribuciones de edad de poblaciones cuando se comparan las tasas de mortalidad.

4. *La tasa de mortalidad de la enfermedad X en la ciudad A es de 75/100.000 en personas de 65-69 años. La tasa de mortalidad por la misma enfermedad en la ciudad B es de 150/100.000 en personas de 65 a 69 años. La conclusión de que la enfermedad X es dos veces más prevalente en las personas de 65 a 69 años en la ciudad B que en las personas de igual edad en la ciudad A es:*

a. Correcta.

b. Incorrecta, por no distinguir entre prevalencia y mortalidad.

c. Incorrecta, por no ajustar por diferencias en la distribución por edades.

d. Incorrecta, por no distinguir entre prevalencia puntual y de periodo.

e. Incorrecta, porque utiliza una proporción cuando se necesita una tasa para apoyar la conclusión.

5. *La tasa de incidencia de una enfermedad es cinco veces mayor en mujeres que en varones, pero las tasas de prevalencia no muestran diferencias por sexo. La mejor explicación es que:*

a. La tasa de mortalidad bruta por todas las causas es mayor en las mujeres.

b. La tasa de letalidad de esta enfermedad es mayor en las mujeres.

c. La tasa de letalidad de esta enfermedad es menor en las mujeres.

d. La duración de esta enfermedad es más corta en los varones.

e. Los factores de riesgo de la enfermedad son más frecuentes en las

mujeres.

6. *Para una enfermedad como el cáncer de páncreas, que es sumamente letal y de corta duración:*
   a. Las tasas de incidencia y de mortalidad son similares.
   b. Las tasas de mortalidad son mucho más elevadas que las tasas de incidencia.
   c. Las tasas de incidencia son mucho más elevadas que las tasas de mortalidad.
   d. Las tasas de incidencia no se relacionan con las tasas de mortalidad.
   e. Ninguna de las anteriores.

7. *En 1990 se produjeron 4.500 muertes en mineros de 20 a 64 años debidas a enfermedades pulmonares. El número esperado de muertes por enfermedades pulmonares en este grupo ocupacional en 1990, basado en las tasas de mortalidad específicas de edad para todos los varones de 20 a 64 años, era de 1.800. ¿Cuál fue la razón de mortalidad estandarizada (RME) para las enfermedades pulmonares en los mineros?*

La pregunta 8 se basa en la siguiente información:

Muertes anuales por cáncer en trabajadores varones blancos en dos industrias

	INDUSTRIA A		INDUSTRIA B	
Localización del cáncer	N.° de muertes	% de todas las muertes por cáncer	N.° de muertes	% de todas las muertes por cáncer
Aparato respiratorio	180	33	248	45
Aparato digestivo	160	29	160	29
Genitourinario	80	15	82	15
Resto de localizaciones	130	23	60	11
Total	550	100	550	100

Teniendo en cuenta la información anterior, se concluyó que los trabajadores de la industria B tenían un mayor riesgo de muerte por cáncer del aparato respiratorio que los trabajadores de la industria A. (Suponga que las distribuciones por edades de los trabajadores de las dos industrias son casi idénticas.)

8. *¿Cuál de las siguientes afirmaciones es correcta?*

a. La conclusión alcanzada es correcta.

b. La conclusión alcanzada puede ser incorrecta porque se usaron tasas de mortalidad proporcional cuando era necesario emplear tasas de mortalidad específicas de edad.

c. La conclusión alcanzada puede ser incorrecta porque no hubo un grupo para comparar.

d. La conclusión alcanzada puede ser incorrecta porque se empleó la mortalidad proporcional cuando se necesitaban tasas de mortalidad específicas de causa.

e. Ninguna de las anteriores.

9. *El director de un programa de un organismo internacional de financiación de proyectos sanitarios necesita identificar las regiones que se beneficiarían de una intervención dirigida a reducir la discapacidad prematura. El director del programa pide a un consultor en asistencia sanitaria que desarrolle una propuesta utilizando un índice que le ayude a tomar esta decisión. ¿Cuál de los siguientes le serviría mejor para este fin?*

a. Tasa de letalidad.

b. Tasa de mortalidad bruta.

c. Años de vida ajustados por discapacidad.

d. Razón de mortalidad estandarizada.

10. *Los siguientes datos son razones de mortalidad estandarizadas (RME) del cáncer de pulmón en Inglaterra:*

	RAZONES DE MORTALIDAD ESTANDARIZADAS	
Ocupación	**1949-60**	**1968-79**
Carpinteros	209	135
Albañiles	142	118

Basándonos únicamente *en estas RME, es posible concluir que:*

a. El número de muertes por cáncer de pulmón en los carpinteros de 1949 a 1960 fue superior al número de muertes por cáncer de pulmón en albañiles durante el mismo periodo.

b. La mortalidad proporcional por cáncer de pulmón en los albañiles entre 1949 y 1960 fue mayor que la mortalidad proporcional debida a cáncer de pulmón en el mismo grupo ocupacional de 1968 a 1979.

c. La tasa de mortalidad por cáncer de pulmón ajustada por edad en

los albañiles fue mayor de 1949 a 1960 que de 1968 a 1979.

d. La tasa de mortalidad por cáncer de pulmón en los carpinteros de 1968 a 1979 fue mayor que la que cabría esperar en un grupo de varones de edades similares de todas las ocupaciones.

e. La tasa de mortalidad proporcional por cáncer de pulmón en los carpinteros de 1968 a 1979 fue 1,35 veces mayor que la que cabría esperar en un grupo de varones de edades similares de todas las ocupaciones.

Las preguntas 11, 12 y 13 se basan en la siguiente información:

Número de personas y muertes por la enfermedad Z por grupo de edad en las comunidades X e Y

	COMUNIDAD X		COMUNIDAD Y	
Grupo de edad	N.° de personas	N.° de muertes por la enfermedad Z	N.° de personas	N.° de muertes por la enfermedad Z
Jóvenes	8.000	69	5.000	48
Ancianos	11.000	115	3.000	60

Calcule la tasa de mortalidad ajustada por edad para la enfermedad Z en las comunidades X e Y por el método directo, empleando como población estándar el total de ambas comunidades.

11. *La tasa de mortalidad ajustada por edad de la enfermedad Z en la comunidad X es:*

12. *La mortalidad proporcional de la enfermedad Z en la comunidad Y es:*

a. 9,6/1.000.

b. 13,5/1.000.

c. 20,0/1.000.

d. 10,8/1.000.

e. Ninguna de las anteriores.

13. *¿Cuál de las siguientes afirmaciones con respecto al ajuste directo es VERDADERA?*

a. La tasa de mortalidad ajustada por edad de la comunidad X es aún más alta que la tasa de mortalidad de la comunidad Y, en comparación con la tasa de mortalidad bruta.

b. Las tasas de mortalidad ajustadas por edad para la comunidad X deben usarse para tomar decisiones relacionadas con la

asignación de fondos para la atención hospitalaria de los moribundos en la comunidad X.

c. Para el ajuste directo por edad, el peso para una categoría de edad dada es el porcentaje de muertes para ese grupo de edad.

d. Para el ajuste directo por edad, el peso para una categoría de edad dada es el número de individuos en la población estándar para ese grupo de edad.

e. La diferencia en las tasas de mortalidad ajustadas entre la comunidad X y la comunidad Y siempre es atribuible a las diferencias en la composición por edad entre las dos poblaciones.

14. *Los datos de vigilancia indican que la prevalencia de enfermedad hepática crónica en Estados Unidos aumentó un 104% entre los años 1990 y 2008. Mientras que la enfermedad hepática crónica ocurre en personas de todas las edades, la tasa de mortalidad más elevada se observa en personas de 65 años o más. Estados Unidos tiene proporcionalmente más personas de 65 años o más que el país X. ¿Qué pasaría si las tasas brutas de mortalidad en Estados Unidos estuvieran estandarizadas por edad a la población del país X con el fin de comparar el riesgo de morir de enfermedad hepática crónica en las dos poblaciones?*

a. La tasa de mortalidad estandarizada por edad para Estados Unidos sería menor que la tasa bruta de mortalidad para Estados Unidos.

b. La tasa de mortalidad estandarizada por edad para Estados Unidos sería mayor que la tasa bruta de mortalidad para Estados Unidos.

c. La tasa de mortalidad estandarizada por edad para Estados Unidos sería la misma que la tasa bruta de mortalidad para Estados Unidos.

d. La tasa de mortalidad estandarizada por edad para Estados Unidos no se puede utilizar para esta comparación.

e. La tasa de mortalidad estandarizada por edad para Estados Unidos sería la misma que la tasa de mortalidad proporcional.

15. *Entre los trabajadores de una planta procesadora de pescado, el 30% de todas las muertes se debieron a infarto de miocardio. Entre los trabajadores de una fábrica de cerveza, el 10% de todos los fallecimientos se debieron a infarto de miocardio. Los investigadores concluyeron que los trabajadores de la planta procesadora de pescado tenían un mayor riesgo de muerte por*

infarto de miocardio que los trabajadores de la fábrica de cerveza. Esta conclusión:

a. Es correcta.

b. Puede ser incorrecta porque se basa en la mortalidad proporcional.

c. Puede ser incorrecta porque asume la misma tasa de letalidad por infarto de miocardio en ambos sitios de trabajo.

d. Puede ser incorrecta porque el aceite de pescado consumido protege contra la muerte debido a un infarto de miocardio.

e. Puede ser incorrecta porque se desconoce la prevalencia de infarto de miocardio en los dos grupos.

CAPÍTULO 5

Evaluación de la validez y fiabilidad de las pruebas diagnósticas y de cribado

Palabras clave

sensibilidad
especificidad
pruebas secuenciales (dos etapas) y simultáneas. valor predictivo
fiabilidad
porcentaje de concordancia y estadístico kappa
validez

Un sujeto normal es aquel que no ha sido estudiado suficientemente.
—Anónimo.

Objetivos de aprendizaje

- Definir la validez y fiabilidad de las pruebas diagnósticas y de cribado.
- Comparar las medidas de validez, incluidas la sensibilidad y la especificidad.
- Ilustrar el uso de pruebas múltiples (pruebas secuenciales y simultáneas).
- Introducir el valor predictivo positivo y negativo.
- Utilizar medidas de fiabilidad, incluidas kappa y el porcentaje de concordancia.

Para comprender cómo se transmite y se desarrolla una enfermedad y para proporcionar una asistencia sanitaria efectiva y apropiada, es necesario diferenciar las personas de la población que padecen una enfermedad de las que no la tienen. Esto es un desafío importante, tanto en el ámbito clínico centrado en la asistencia al paciente como en el ámbito de la salud pública, en

el que se consideran programas de prevención secundaria dirigidos a la detección precoz de la enfermedad mediante cribado e intervenciones y en el que se realizan estudios etiológicos que proporcionan la base para la prevención primaria, si es posible. Por tanto, la calidad de las pruebas diagnósticas y de cribado es un tema fundamental. Con independencia de que la prueba se trate de una exploración física, una radiografía de tórax, un electrocardiograma o un análisis de sangre o de orina, la pregunta común es: ¿cómo de buena es la prueba para identificar grupos de personas con y sin la enfermedad en cuestión? Este capítulo aborda el tema de la valoración de la calidad de las pruebas diagnósticas y de cribado de reciente disponibilidad para adoptar decisiones razonables sobre su uso e interpretación.

Variación biológica de las poblaciones humanas

Cuando se utiliza una prueba para diferenciar entre individuos con resultados normales y anormales, es importante comprender cómo se distribuyen las características en las poblaciones humanas.

La figura 5.1 muestra la distribución de los casos confirmados de infección por el virus de la hepatitis C notificados recientemente en Massachusetts en 2009. Podemos observar que hay dos picos de casos de infección por el virus de la hepatitis C, correspondientes a adultos jóvenes y a personas de mediana edad. Este tipo de distribución, en el que existen dos picos, se denomina *curva bimodal*. La distribución bimodal permite la identificación de mayores tasas de *casos nuevos* entre estos dos grupos de edad distintos, lo que podría deberse a diversos motivos. En esta situación se ha producido un aumento importante en la hepatitis entre los usuarios de drogas por vía parenteral, una práctica asociada con compartir las jeringuillas, lo que llevó a esta distribución bimodal.

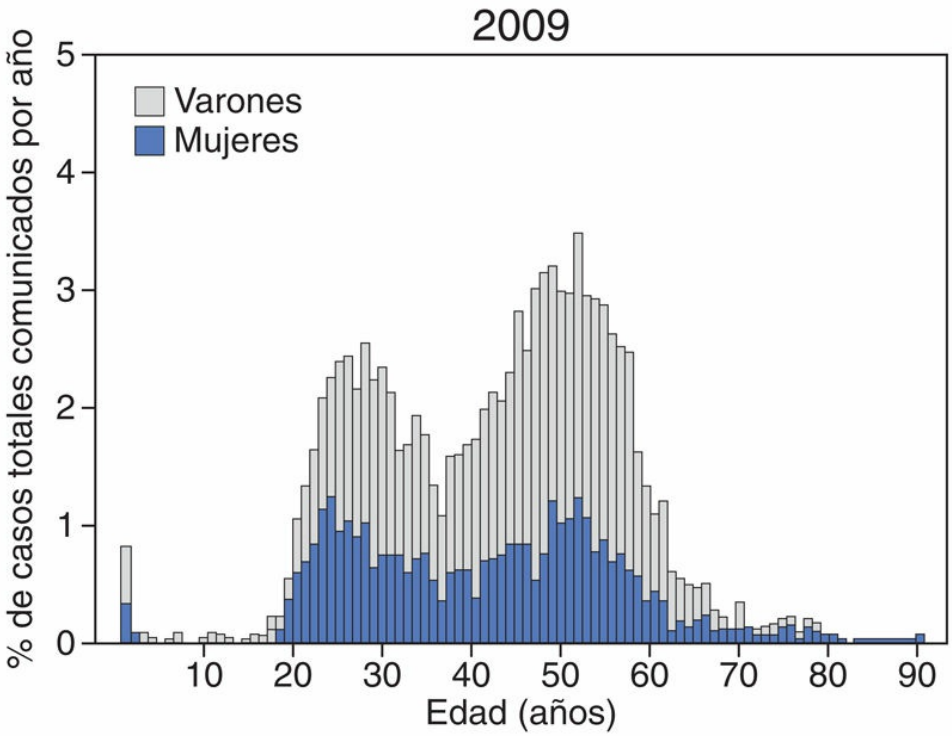
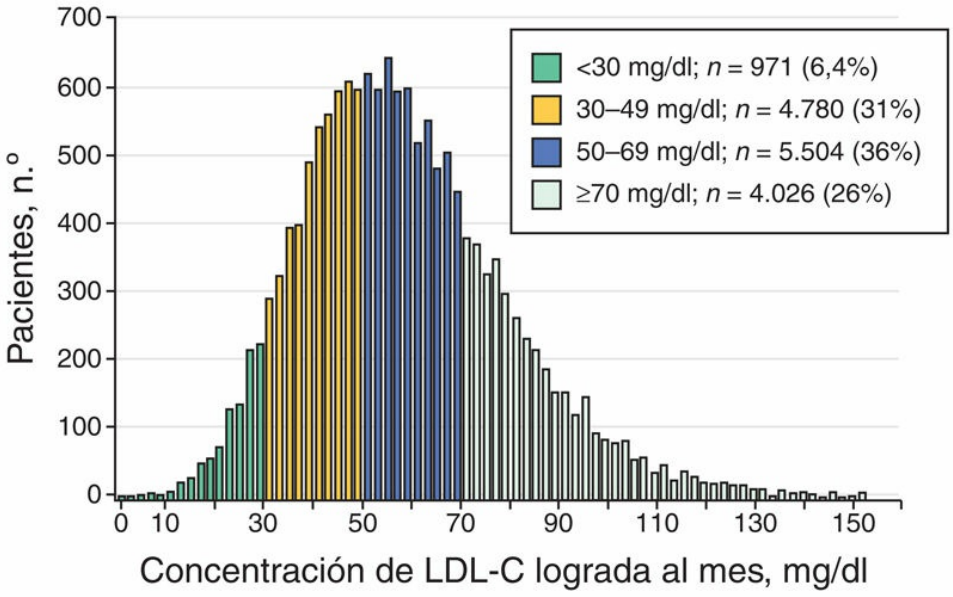


FIG. 5.1 Distribución de los casos confirmados comunicados recientemente de infección por el virus de la hepatitis C en Massachusetts, 2009. (Modificado de Centers for Disease Control and Prevention. Hepatitis C virus infection among adolescents and young adults: Massachusetts, 2002–2009. *MMWR Morb Mortal Wkly Rep.* 2011;60:537–541.)

Sin embargo, la mayoría de las características humanas no presentan

generalmente una distribución bimodal. La figura 5.2 muestra la distribución del colesterol de las lipoproteínas de baja densidad (LDL-C) lograda en los participantes de un ensayo clínico para estudiar la seguridad de la reducción intensiva del LDL-C en comparación con la disminución menos intensiva de este en pacientes que han sufrido un síndrome coronario agudo. En esta figura no existe una curva bimodal; lo que observamos en una *curva unimodal* (con un solo pico). Por tanto, si queremos separar a los sujetos del grupo que lograron un nivel bajo seguro de LDL-C, se debe establecer un punto de corte de LDL-C debajo del cual las personas están consideradas como pertenecientes al grupo de «nivel bajo seguro» y por encima del cual no están consideradas como tales. Este estudio muestra que no hay un nivel concreto de LDL-C que deba ser un objetivo de tratamiento. Aunque podemos elegir un punto de corte basándonos en consideraciones estadísticas, como han demostrado los autores del estudio, lo ideal es que quisiéramos elegir el punto de corte basado en información biológica; es decir, queremos saber si una estrategia para reducir de modo intensivo la concentración de LDL-C por debajo del valor de corte elegido se asocia con un mayor riesgo de sufrir posteriormente efectos adversos del tratamiento, eventos adversos musculares, hepatobiliares y neurocognitivos, o complicaciones de la enfermedad, accidente cerebrovascular hemorrágico, insuficiencia cardiaca, cáncer y muerte no cardiovascular. Desafortunadamente, para muchas características humanas no contamos con una información que sirva de guía para establecer ese valor.



La concentración media de LDL-C fue de 56 mg/dl (rango intercuartil, 43-70 mg/dl).

Para convertir LDL-C a milimoles por litro, multiplique por 0,0259.

FIG. 5.2 Distribución de la concentración lograda calculada de colesterol de las lipoproteínas de baja densidad *(LDL-C)* a 1 mes entre los pacientes que no tuvieron una eficacia primaria o un evento de seguridad preespecificado antes de

la muestra. (Datos de Giugliano RP, Wiviott SD, Blazing MA, et al. Long-term safety and efficacy of achieving very low levels of low-density lipoprotein cholesterol: a prespecified analysis of the IMPROVE-IT trial. *JAMA Cardiol.* 2017;2:547–555.)

En cualquiera de las distribuciones (unimodal o bimodal) por lo general resulta fácil diferenciar entre los valores extremos de lo normal y lo anormal. Sin embargo, en ambos tipos de curvas, existen casos dudosos que permanecen en una zona gris.

Validez de las pruebas de cribado

La *validez* de una prueba se define como su capacidad para diferenciar entre los que tienen una enfermedad y los que no la tienen. La validez posee dos componentes: sensibilidad y especificidad. La *sensibilidad* de la prueba se define como la capacidad de la prueba para identificar correctamente a aquellos que *tienen* la enfermedad. La *especificidad* de la prueba se define como la capacidad de la prueba para identificar correctamente a aquellos que *no tienen* la enfermedad.

Pruebas con resultados dicotómicos (positivo o negativo)

Suponga que tenemos una población hipotética de 1.000 personas, de las que 100 tienen cierta enfermedad y 900 no la tienen. Existe una prueba disponible que proporciona resultados positivos o negativos. Queremos utilizar esta prueba para diferenciar las personas que tienen la enfermedad de las que no la tienen. Los resultados obtenidos aplicando esta prueba a esta población de 1.000 personas se muestran en la tabla 5.1.

Tabla 5.1

Cálculo de la sensibilidad y la especificidad de las exploraciones de cribado

Ejemplo: asuma una población de 1.000 personas, de las cuales 100 tienen la enfermedad y 900 no la tienen. Se utiliza una prueba de cribado para identificar a las 100 personas que tienen la enfermedad.

Resultados del cribado	Características verdaderas de la población: Tienen la enfermedad	Características verdaderas de la población: No tienen la enfermedad	Totales
Positivo	80	100	180
Negativo	20	800	820
Totales	100	900	1.000

Sensibilidad: $\frac{80}{100} = 80\%$

Especificidad: $\frac{800}{900} = 89\%$

¿Cómo de buena era la prueba? En primer lugar, ¿cómo de buena fue la prueba para identificar correctamente a aquellos que padecían la enfermedad? La tabla 5.1 indica que de los 100 individuos que tenían la enfermedad, 80 fueron identificados correctamente como «positivos» por la

prueba y 20 no fueron identificados como positivos. Por tanto, la *sensibilidad* de la prueba, que se define como la proporción de individuos enfermos que fueron identificados correctamente como «positivos» por la prueba, fue de 80/100 o del 80%.

En segundo lugar, ¿cómo de buena fue la prueba para identificar correctamente a aquellos que no presentaban la enfermedad? Fijándonos de nuevo en la tabla 5.1, observamos que, de las 900 personas que no tenían la enfermedad, la prueba identificó correctamente a 800 como «negativos». La *especificidad* de la prueba, que se define como la proporción de personas no enfermas que son identificadas correctamente como «negativos» por la prueba, es por tanto de 800/900 o del 89%.

Para calcular la sensibilidad y la especificidad de una prueba, debemos saber quiénes tienen «realmente» la enfermedad y quiénes «no» basándonos en una fuente distinta a la prueba que estamos utilizando. De hecho, estamos comparando nuestros resultados de la prueba con algún «método de referencia» *(gold standard),* una fuente externa de «verdad» referente al estado de la enfermedad de cada individuo de la población. En ocasiones, esta verdad puede ser el resultado de otra prueba ya utilizada, y otras veces es el resultado de una prueba más definitiva y a menudo más invasiva (p. ej., una biopsia tumoral, un cateterismo cardiaco o una biopsia de tejido). Sin embargo, en una situación real, cuando utilizamos una prueba para identificar a personas enfermas y no enfermas de una población, claramente desconocemos quién tiene la enfermedad y quién no la tiene. (Si ya conociésemos esta información, realizar la prueba no tendría sentido.) No obstante, para valorar cuantitativamente la sensibilidad y la especificidad de una prueba, debemos tener otra fuente de verdad con la que comparar los resultados de la prueba a estudio.

En la tabla 5.2 se comparan los resultados de una prueba dicotómica (resultados inequívocamente positivos o negativos) con el estado real de la enfermedad. De manera ideal, querríamos que todos los individuos en los que realizamos la prueba se localicen en dos celdillas, en las esquinas superior izquierda e inferior derecha de la tabla: personas con la enfermedad identificadas correctamente mediante la prueba como «positivos» *(verdaderos positivos)* y personas sin la enfermedad identificadas correctamente mediante la prueba como «negativos» *(verdaderos negativos).* Desafortunadamente, esta situación suele ser una excepción. Algunas personas que no tienen la enfermedad son identificadas incorrectamente como «positivos» por la prueba *(falsos positivos)* y algunas personas que padecen la enfermedad son identificadas incorrectamente como «negativos» *(falsos negativos).*

Tabla 5.2

Comparación de los resultados de una prueba dicotómica respecto al estado de la enfermedad

	Características verdaderas de la población	
Resultados de la prueba	**Tienen la enfermedad**	**No tienen la enfermedad**
Positivo	**Verdadero positivo (VP):** tienen la enfermedad y la prueba es positiva	**Falso positivo (FP):** no tienen la enfermedad, pero la prueba es positiva
Negativo	**Falso negativo (FN):** tienen la enfermedad, pero la prueba es negativa	**Verdadero negativo (VN):** no tienen la enfermedad y la prueba es negativa

$$\text{Sensibilidad} = \frac{TP}{TP + FN} \qquad \text{Especificidad} = \frac{TN}{TN + FP}$$

¿Por qué son importantes estas cuestiones? Cuando realizamos un programa de cribado, a menudo encontramos un grupo numeroso de personas con resultados positivos en el cribado, en el que se incluyen personas que realmente tienen la enfermedad *(verdaderos positivos)* y personas que no tienen la enfermedad *(falsos positivos).* La cuestión de los *falsos positivos* es importante porque todas las personas con resultados positivos en el cribado son exploradas de nuevo con pruebas más sofisticadas y caras, o bien son sometidas a procedimientos invasivos que son innecesarios. De los diversos problemas que resultan, el primero es la carga sobre el sistema de asistencia sanitaria. Otro es la ansiedad y la preocupación inducidas en las personas a las que se informa que el resultado de la prueba ha sido positivo. Existen datos de peso que indican que muchas personas consideradas «positivas» por la prueba de cribado nunca se libran por completo de dicha etiqueta, incluso aunque los resultados de evaluaciones posteriores sean negativos. Por ejemplo, los niños identificados como «positivos» en un programa de cribado de cardiopatía pueden ser tratados como discapacitados por los padres y el personal de la escuela incluso después de haberles

informado de que pruebas posteriores más definitivas fueron negativas. Además, dichos individuos pueden encontrar limitaciones a nivel laboral y no ser asegurados por la interpretación errónea de los resultados positivos en una prueba de cribado, incluso aunque pruebas posteriores no logren confirmar los hallazgos positivos.

¿Por qué es importante el problema de los *falsos negativos*? Si una persona presenta la enfermedad pero es informada incorrectamente de que los resultados de la prueba son negativos, en caso de que se trate de una enfermedad grave para la que se dispone de intervenciones efectivas, el problema es realmente grave. Por ejemplo, si la enfermedad es un tipo de cáncer que es curable únicamente en sus etapas más iniciales, un resultado falso negativo puede representar una sentencia de muerte virtual. Por tanto, la importancia de los resultados falsos negativos depende de la naturaleza y la gravedad de la enfermedad para la que se realiza el cribado, de la eficacia de las medidas de intervención disponibles y de si la eficacia es mayor si la intervención se aplica en las etapas iniciales de la evolución natural de la enfermedad.

Pruebas de variables continuas

Hasta ahora hemos analizado una prueba con tan solo dos resultados posibles: positivo o negativo. Pero a menudo realizamos pruebas para una variable continua, como la presión arterial o la concentración de glucosa en sangre, para la que no existe un resultado «positivo» o «negativo». Por tanto, se debe adoptar una decisión para establecer un punto de corte por encima del cual el resultado de la prueba es considerado positivo y por debajo del cual el resultado es considerado negativo. Consideremos el diagrama que se muestra en la figura 5.3.

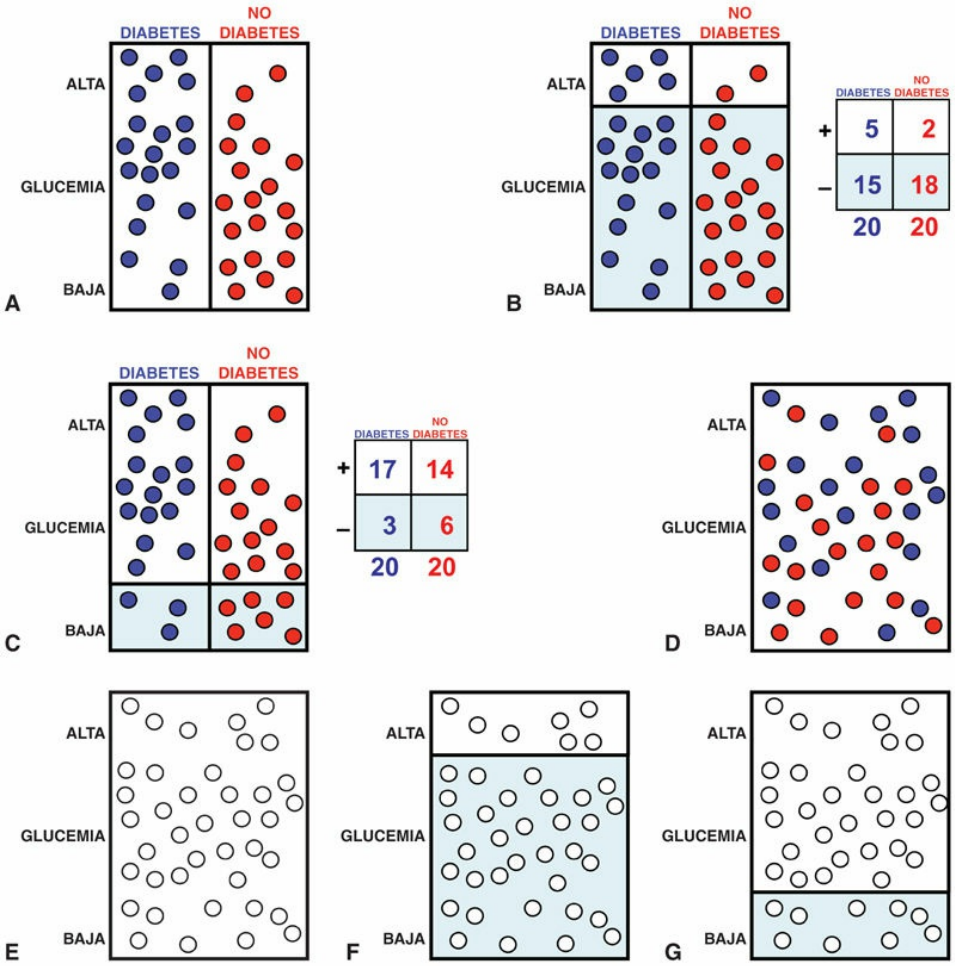


FIG. 5.3 (A-G) Los efectos de elegir diferentes niveles de corte para definir un resultado positivo en una prueba cuando se realiza un cribado de diabetes utilizando un marcador continuo, la glucemia, en una población hipotética. (V. exposición en el apartado «Pruebas de variables continuas», en la página 97.)

En la figura 5.3A se muestra una población de 20 diabéticos y 20 no diabéticos en los que se realizan pruebas de cribado utilizando una prueba de glucosa en sangre cuya escala se muestra en el eje vertical de valores máximos a mínimos. Los diabéticos son representados por círculos azules y los no diabéticos, por círculos rojos. Observamos que, aunque las concentraciones de glucosa en sangre suelen ser más elevadas en los diabéticos que en los no diabéticos, no existe una concentración que separe claramente los dos grupos; existe cierto solapamiento entre diabéticos y no diabéticos en cada concentración de glucosa en sangre. Sin embargo, debemos seleccionar un punto de corte de modo que aquellos cuyos resultados se encuentren por encima de dicho punto de corte puedan considerarse «positivos» y puedan ser vueltos a explorar con más pruebas, y aquellos cuyos resultados se encuentren por debajo de dicho punto se consideren «negativos» y no sean programados para realizar pruebas adicionales.

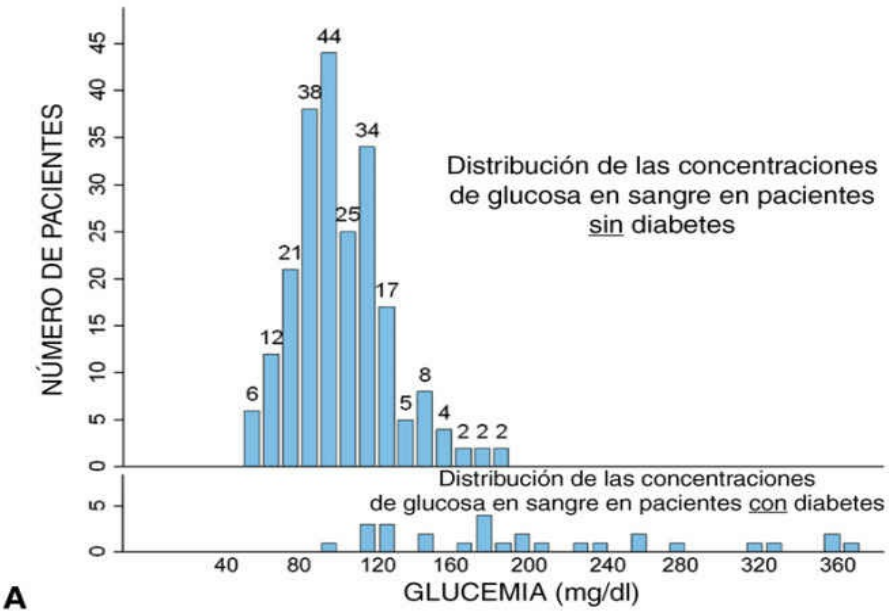
Supongamos que se elige una concentración de corte relativamente elevada (v. fig. 5.3B). Claramente, muchos diabéticos no serán identificados como positivos; por otro lado, la mayoría de los no diabéticos serán identificados

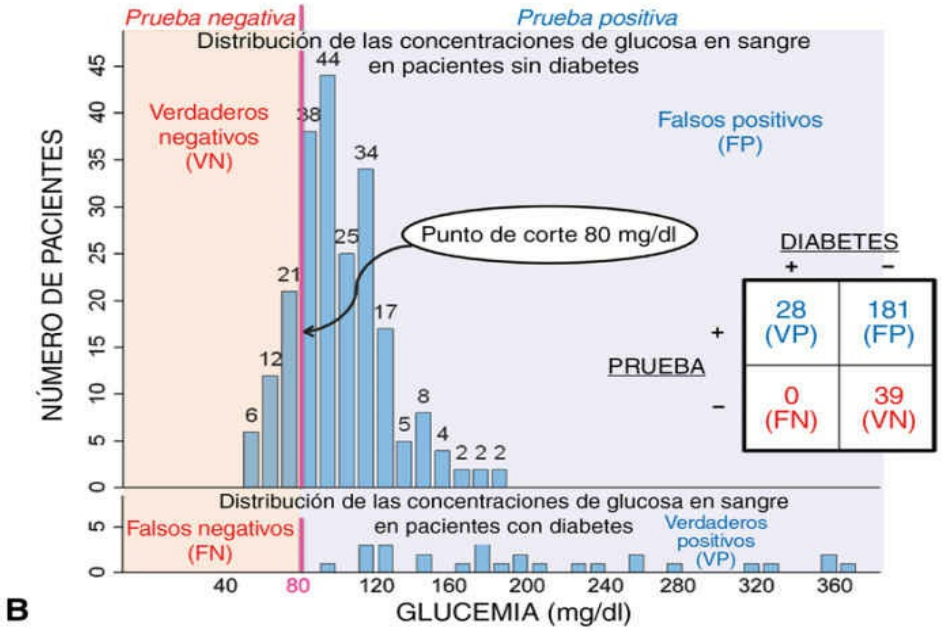
correctamente como negativos. Si representamos estos resultados en una tabla de 2 × 2, la sensibilidad de la prueba utilizando esta concentración de corte será del 25% (5/20) y la especificidad, del 90% (18/20). Entonces, la mayoría de los diabéticos no serán detectados, pero la mayoría de los no diabéticos se clasificarán correctamente.

¿Qué ocurre si se elige una concentración de corte baja (v. fig. 5.3C)? Muy pocos diabéticos serían mal diagnosticados. ¿Cuál es entonces el problema? Una gran proporción de los no diabéticos son identificados ahora como positivos por la prueba. Como se observa en la tabla de 2 × 2, la sensibilidad es ahora del 85% (17/20), pero la especificidad es únicamente del 30% (6/20).

La dificultad estriba en que en el mundo real no existe una línea vertical que separe los diabéticos de los no diabéticos, y, de hecho, se encuentran mezclados (v. fig. 5.3D); no son distinguibles ni con círculos rojos o azules (v. fig. 5.3E). Por tanto, si se usa una concentración de corte elevada (v. fig. 5.3F), a todos aquellos con resultados por debajo de la línea se les podrá asegurar que no tienen la enfermedad y no necesitan más seguimiento; si se usa una concentración de corte baja (v. fig. 5.3G), todos aquellos con resultados por encima de la línea serán vueltos a explorar con nuevas pruebas.

En la figura 5.4A se muestran datos reales de un estudio antiguo sobre la distribución de las concentraciones de glucosa en sangre en diabéticos y en no diabéticos. Supongamos que quisiéramos realizar una prueba de cribado en esta población. Si decidimos establecer el punto de corte de modo que podamos identificar a todos los diabéticos (100% de sensibilidad), podríamos elegir una concentración de 80 mg/dl (v. fig. 5.4B). Sin embargo, el problema es que procediendo así también consideraremos positivos a muchos de los no diabéticos, es decir, la especificidad será muy baja. Por otro lado, si establecemos el punto de corte en 200 mg/dl (v. fig. 5.4C), todos los no diabéticos serán identificados como negativos (100% de especificidad), pero ahora podemos pasar por alto a muchos de los diabéticos verdaderos debido a que la sensibilidad será muy baja. Por tanto, entre sensibilidad y especificidad existe una compensación: si aumentamos la sensibilidad disminuyendo el punto de corte, disminuimos la especificidad; y si aumentamos la especificidad elevando el punto de corte, estamos reduciendo la sensibilidad. Como dijo un sabio: «Nadie da nada por nada.»


NÚMERO DE PACIENTES
Distribución de las concentraciones de glucosa en sangre en pacientes sin diabetes
Distribución de las concentraciones de glucosa en sangre en pacientes con diabetes
GLUCEMIA (mg/dl)
A


Prueba negativa
Prueba positiva
Distribución de las concentraciones de glucosa en sangre en pacientes sin diabetes
Verdaderos negativos (VN)
Falsos positivos (FP)
Punto de corte 80 mg/dl
DIABETES
PRUEBA
28 (VP)
181 (FP)
0 (FN)
39 (VN)
Distribución de las concentraciones de glucosa en sangre en pacientes con diabetes
Falsos negativos (FN)
Verdaderos positivos (VP)
NÚMERO DE PACIENTES
GLUCEMIA (mg/dl)
B

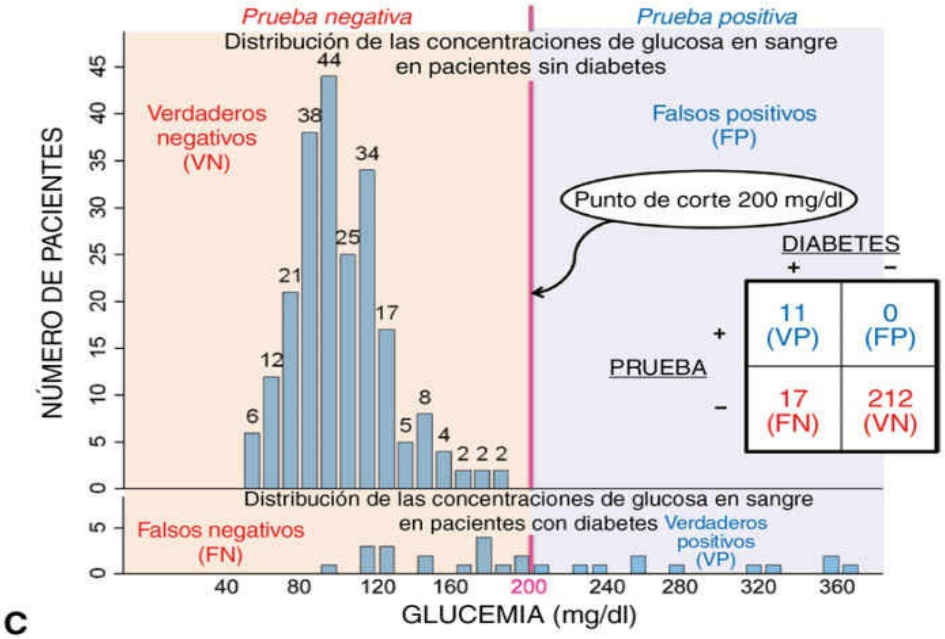

Prueba negativa
Prueba positiva
Distribución de las concentraciones de glucosa en sangre en pacientes sin diabetes
Verdaderos negativos (VN)
Falsos positivos (FP)
Punto de corte 200 mg/dl
DIABETES
PRUEBA
11 (VP)
0 (FP)
17 (FN)
212 (VN)
Distribución de las concentraciones de glucosa en sangre en pacientes con diabetes
Falsos negativos (FN)
Verdaderos positivos (VP)
NÚMERO DE PACIENTES
GLUCEMIA (mg/dl)
C

FIG. 5.4 (A) Distribución de las concentraciones de glucosa en sangre en pacientes hospitalizados con y sin diabetes. (El número de pacientes diabéticos se muestra para cada concentración específica de glucosa en sangre en la distribución *[superior]* de pacientes *sin* diabetes. Debido a las limitaciones de espacio, el número de personas para cada concentración específica de glucosa en sangre no se muestra en la distribución *[inferior]* de personas *con* diabetes.) (B y C) Se muestran dos puntos de corte diferente de glucemia utilizados en el estudio para definir la diabetes. Los datos de los gráficos se presentan a la derecha de cada gráfico en una tabla de 2 × 2. (B) Cuando se toma un punto de corte de glucemia >80 mg/dl para definir la diabetes en esta población, la sensibilidad de la prueba de cribado es del 100%, pero la especificidad es baja. (C) Cuando se utiliza un punto de corte de glucemia >200 mg/dl para definir la diabetes en esta población, la sensibilidad de la prueba de cribado es baja, pero la especificidad es del 100%. (V. explicación en el apartado «Pruebas de variables continuas», pág. 97). *FN*, falsos negativos; *FP*, falsos positivos; *VN*, verdaderos negativos; *VP*, verdaderos positivos. (Modificado de Blumberg M. Evaluating health screening procedures. *Oper Res.* 1957;5:351–360.)

El dilema de decidir si se elige un punto de corte alto o bajo reside en el problema de los falsos positivos y los falsos negativos que resultan de la prueba. Es importante recordar que al realizar pruebas de cribado obtenemos grupos clasificados únicamente según los resultados de las pruebas de cribado, como positivos o negativos. Carecemos de información acerca del verdadero estado de su enfermedad, que, por supuesto, es el motivo para realizar el cribado. De hecho, los resultados de la prueba de cribado no proporcionan cuatro grupos, como se observa en la figura 5.5, sino dos grupos: un grupo de personas con resultados positivos en la prueba y otro grupo con resultados negativos. A los que obtuvieron resultados positivos se les notificarán los resultados de la prueba y se les pedirá que vuelvan para realizar pruebas adicionales. A las personas del otro grupo, con resultados negativos, se les notificará dicho resultado y, por tanto, no se les pedirá que vuelvan para realizar nuevas pruebas (fig. 5.6).

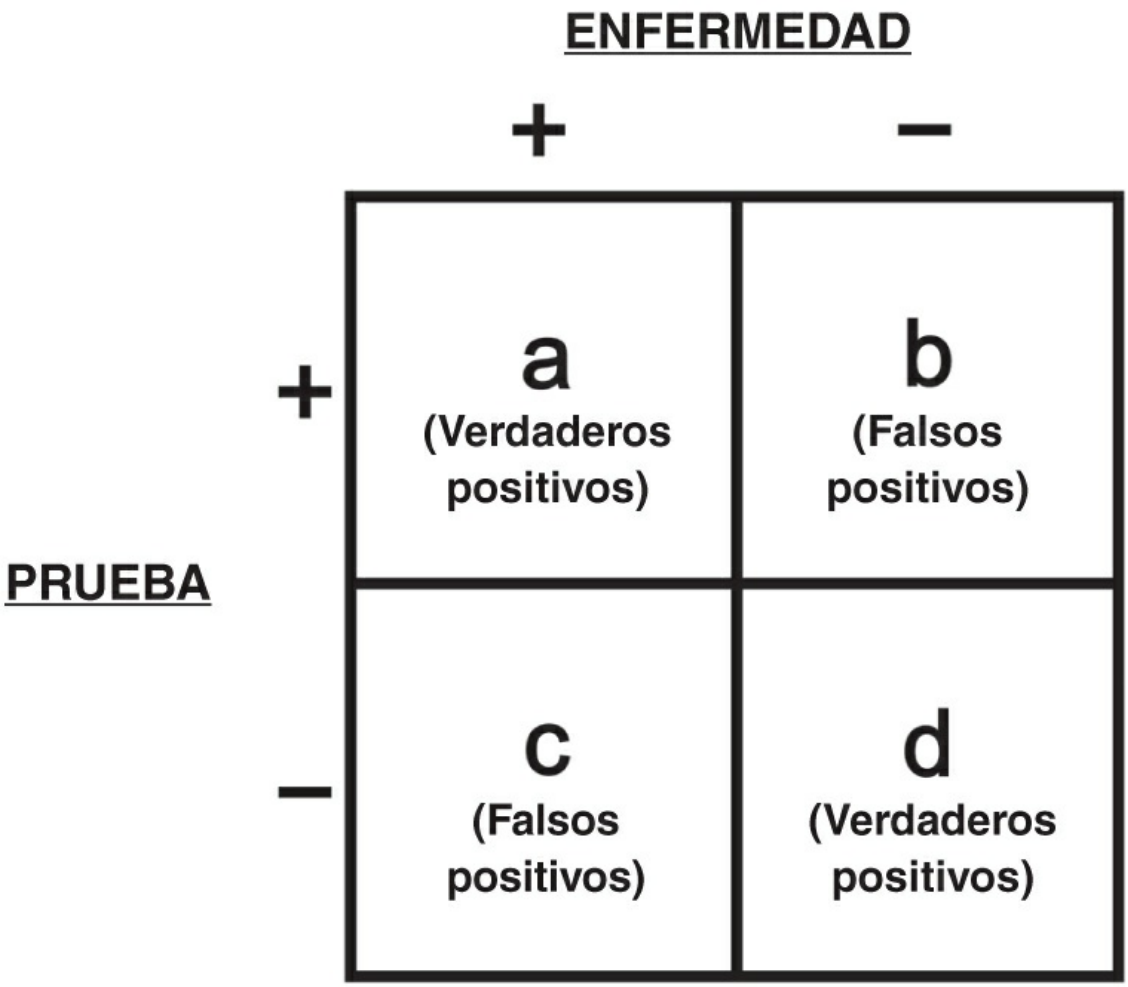


FIG. 5.5 Diagrama en el que se muestran cuatro grupos posibles tras una prueba de cribado con una prueba dicotómica.

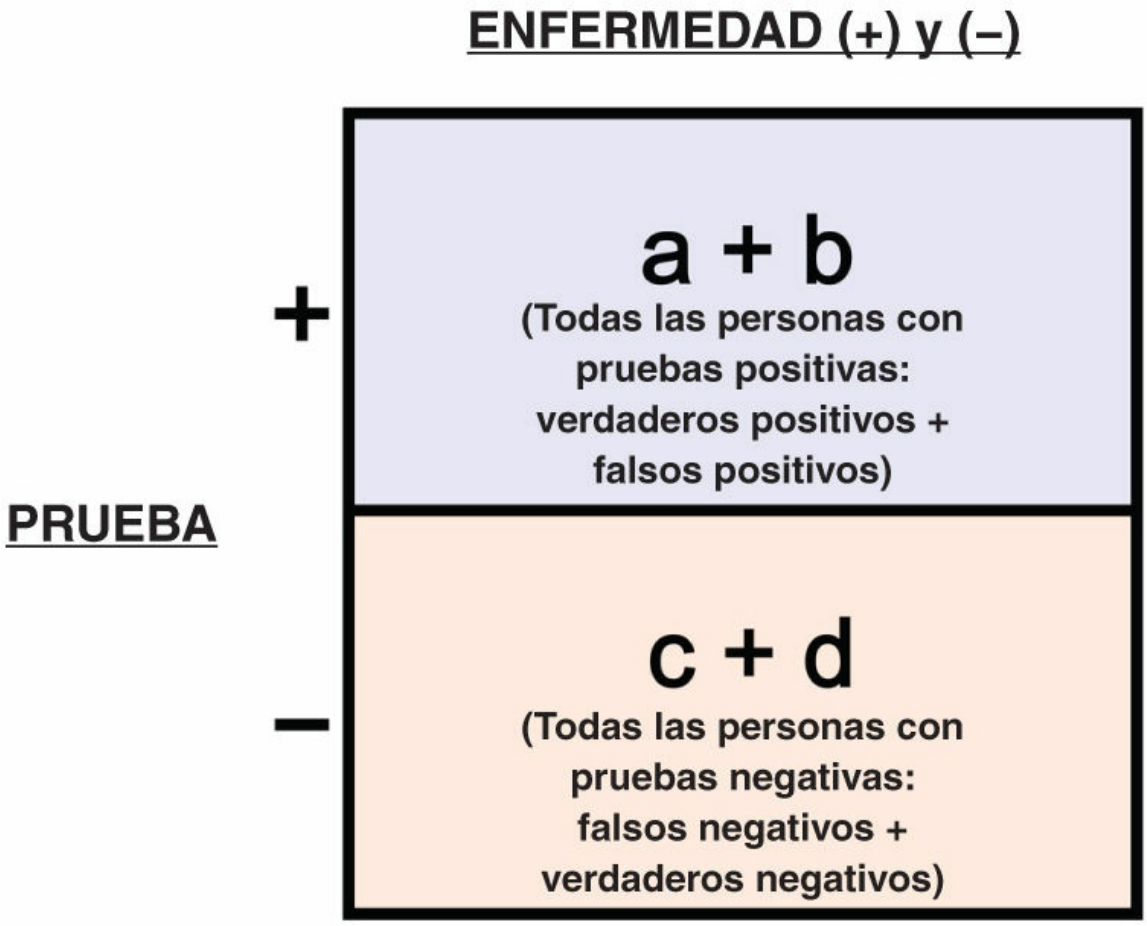


FIG. 5.6 Diagrama que muestra los dos grupos de personas resultantes de una prueba de cribado con una prueba dicotómica: todas las personas con resultados positivos en la prueba y todas las personas con resultados negativos en la prueba.

La elección de un punto de corte alto o bajo para realizar pruebas de cribado depende, por tanto, de la importancia que le otorguemos a los falsos positivos y los falsos negativos. Los falsos positivos se asocian con costes (emocionales y económicos), así como con la dificultad de «desetiquetar» a una persona que obtuvo resultados positivos y que posteriormente se concluyó que no presentaba la enfermedad. Además, los resultados falsos positivos pueden suponer una carga importante al sistema de asistencia sanitaria, ya que un grupo numeroso de personas debe ser citado de nuevo para repetir pruebas, cuando solo unas pocas presentarán la enfermedad. Por otro lado, los pacientes con resultados falsos negativos serán informados de que no tienen la enfermedad y no seguirán siendo revisados, por lo que posiblemente pueden pasarse por alto enfermedades graves en etapas tempranas tratables. Por tanto, la elección de los puntos de corte depende de la importancia relativa de la falsa positividad y la falsa negatividad para la enfermedad en cuestión.

Uso de pruebas múltiples

A menudo pueden realizarse varias pruebas de cribado en los mismos individuos para detectar una enfermedad, ya sea secuencialmente (una después de la otra) o simultáneamente (ambas a la vez). En esta sección se describen los resultados de estos abordajes.

Pruebas secuenciales (en dos etapas)

En las pruebas de cribado secuenciales o en dos etapas, por lo general se realiza primero una prueba menos cara, menos invasiva o menos incómoda, y aquellos en los que el resultado es positivo son vueltos a citar para realizar pruebas adicionales con una prueba más cara, más invasiva o más incómoda, que puede tener más sensibilidad y especificidad. Es de esperar que, citando únicamente a aquellos positivos en la primera prueba de cribado para realizar pruebas adicionales, se reduzca el problema de los falsos positivos.

Consideremos el ejemplo hipotético de la figura 5.7A, en el que se realizan pruebas de cribado de diabetes en una población empleando una prueba con una sensibilidad del 70% y una especificidad del 80%. ¿Cómo se obtienen los datos mostrados en esta tabla? La prevalencia de la enfermedad en esta población es del 5%, por lo que 500 de cada 10.000 habitantes poseen la enfermedad. Con una sensibilidad del 70%, la prueba identificará correctamente a 350 de las 500 personas que tienen la enfermedad. Con una especificidad del 80%, la prueba identificará correctamente como no diabéticas a 7.600 personas de las 9.500 que no son diabéticas; sin embargo, 1.900 de estos 9.500 presentarán resultados positivos. Por tanto, un total de 2.250 personas obtendrán resultados positivos y serán vueltas a citar para realizar una segunda prueba. (Recuérdese que en la vida real no contamos con una línea vertical que separe a los diabéticos de los no diabéticos y no sabemos que solo 350 de los 2.250 son diabéticos.)

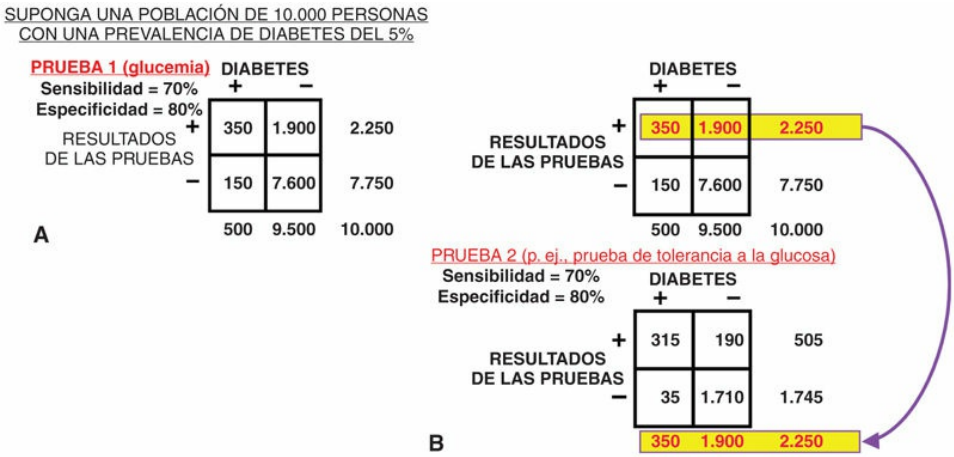


FIG. 5.7 (A-B) Ejemplo hipotético de un programa de cribado en dos etapas. (A)

Hallazgos de la prueba 1 en una población de 10.000 personas. (B) Hallazgos de la prueba 2 en los participantes con resultados positivos en la prueba 1. (V. explicación en el apartado «Pruebas secuenciales (en dos etapas)», en la página 99.)

Las 2.250 personas son vueltas a citar para realizar un cribado con una segunda prueba (como la prueba de tolerancia a la glucosa), que, para este ejemplo, asumimos que tiene una sensibilidad del 90% y una especificidad del 90%. En la figura 5.7B se muestra la prueba 1 conjuntamente con la prueba 2, que se realiza solo en las 2.250 personas con resultados positivos en la primera prueba de cribado y que han sido citados de nuevo para la segunda etapa del cribado.

Como 350 personas (de las 2.250) presentan la enfermedad y la prueba posee una sensibilidad del 90%, 315 de esas 350 serán identificadas correctamente como positivas. Como 1.900 (de las 2.250) no tienen diabetes y la especificidad de la prueba es del 90%, 1.710 de las 1.900 serán identificadas correctamente como negativas y 190 serán falsos positivos.

Ahora somos capaces de calcular la *sensibilidad neta* y la *especificidad neta* del uso de ambas pruebas secuencialmente. Tras completar ambas pruebas, 315 personas del total de 500 diabéticos en esta población de 10.000 habrán sido considerados correctamente positivos: 315/500 = *sensibilidad neta* del 63% (que también puede calcularse multiplicando la sensibilidad de la primera prueba por la sensibilidad de la segunda prueba; es decir, 0,70 × 0,90 = 0,63). Por tanto, empleando ambas pruebas secuencialmente se produce una pérdida de sensibilidad neta. Para calcular la *especificidad neta,* hay que tener en cuenta que 7.600 individuos de los 9.500 de esta población que no son diabéticos fueron considerados correctamente negativos en la primera etapa del cribado y no fueron sometidos a más pruebas; en la segunda etapa del cribado 1.710 individuos más de los 9.500 no diabéticos fueron considerados correctamente negativos. Así, un total de 7.600 + 1.710 de los 9.500 no diabéticos fueron considerados correctamente negativos: 9.310/9.500 = *especificidad neta* del 98%. Por tanto, el uso de ambas pruebas secuencialmente ha resultado en una ganancia de *especificidad neta*.

Pruebas simultáneas

Centrémonos ahora en el uso de pruebas simultáneas. Asumamos que en una población de 1.000 personas, la prevalencia de una enfermedad es del 20%. Por tanto, 200 personas padecen la enfermedad, pero no sabemos quiénes son. Para identificar a las 200 personas que tienen esta enfermedad, realizamos pruebas de cribado en esta población de 1.000 personas utilizando dos pruebas para esta enfermedad, la prueba A y la prueba B, al mismo tiempo. Asumamos que la sensibilidad y la especificidad de las dos pruebas son las siguientes:

Prueba A	**Prueba B**
Sensibilidad = 80%	Sensibilidad = 90%

Especificidad = 60%	Especificidad = 90%

Sensibilidad neta utilizando dos pruebas simultáneas

La primera pregunta que nos planteamos es: ¿cuál es la *sensibilidad neta* si se utilizan la prueba A y la prueba B *simultáneamente*? Para considerar a una persona positiva y, por tanto, poder incluirla en el numerador para calcular la sensibilidad neta de las dos pruebas utilizadas simultáneamente, dicha persona debe ser identificada como positiva por la prueba A, la prueba B o ambas.

Para calcular la sensibilidad neta, consideremos primero los resultados del cribado con la prueba A, cuya sensibilidad es del 80%: de las 200 personas que tienen la enfermedad, 160 son identificadas como positivas (tabla 5.3). En la figura 5.8A, la elipse representa a las 200 personas que tienen la enfermedad. En la figura 5.8B, el círculo rosa en el interior de la elipse representa a las 160 personas identificadas como positivas con la prueba A. Estas 160 personas son verdaderos positivos con la prueba A.

Tabla 5.3

Resultados del cribado con la prueba A

Resultados del cribado	Población	
	Enfermedad	**No enfermedad**
Positivo	**160**	320
Negativo	40	480
Total	200	800

Sensibilidad = 80% Especificidad = 60%

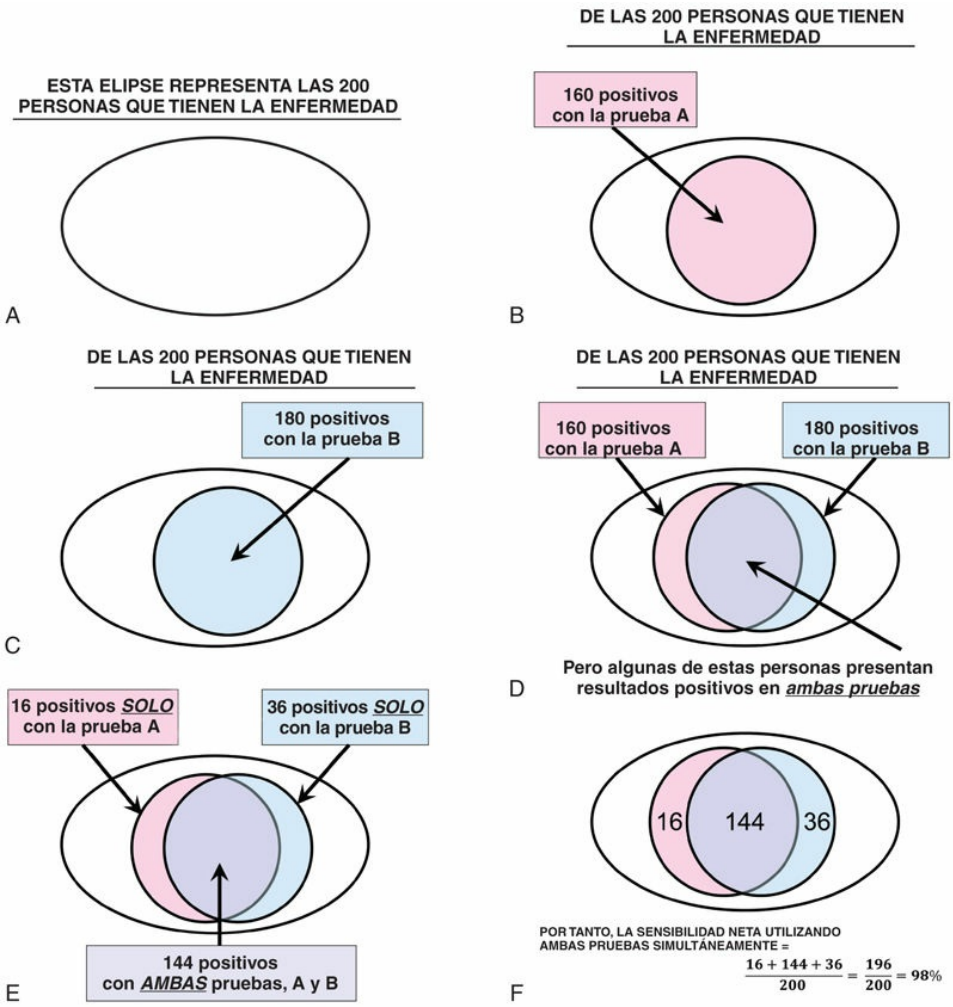
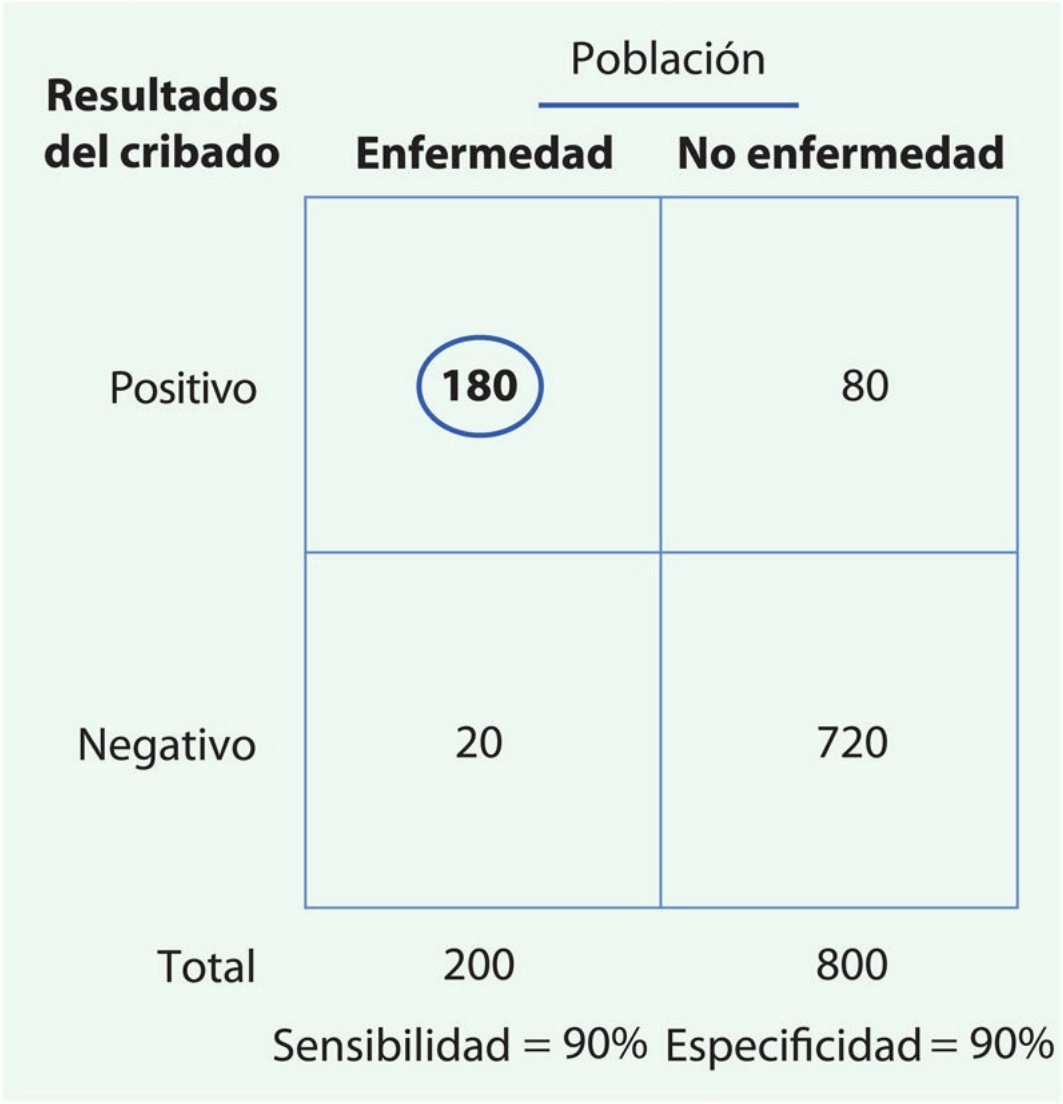


FIG. 5.8 (A-F) Sensibilidad neta: ejemplo hipotético de pruebas simultáneas. (V. explicación en el apartado «Sensibilidad neta utilizando dos pruebas simultáneas», pág. 102.)

Consideremos a continuación los resultados del cribado con la prueba B, cuya sensibilidad es del 90% (tabla 5.4). De las 200 personas que tienen la enfermedad, 180 son identificadas como positivas por la prueba B. En la figura 5.8C, la elipse representa de nuevo a las 200 personas que tienen la enfermedad. El círculo azul en el interior de la elipse representa a las 180 personas identificadas como positivas con la prueba B. Estas 180 personas son verdaderos positivos con la prueba B.

Tabla 5.4

Resultados del cribado con la prueba B



Con el fin de calcular el numerador para la sensibilidad neta, no podemos sumar simplemente el número de personas identificadas como positivas con la prueba A y el número de personas identificadas como positivas con la prueba B, pues algunas personas fueron identificadas como positivas con ambas pruebas. Estas personas se representan en lavanda en el área de solapamiento entre ambos círculos, y no queremos contarlas dos veces (v. fig. 5.8D). ¿Cómo determinamos cuántas personas fueron identificadas como positivas con ambas pruebas?

La prueba A posee una sensibilidad del 80% y, por tanto, identifica como positivas al 80% de las 200 personas que tienen la enfermedad (160 personas). La prueba B posee una sensibilidad del 90% y, por tanto, identifica como positivas al 90% de las mismas 160 personas que fueron identificadas por la prueba A (144 personas). Por tanto, cuando empleamos simultáneamente las pruebas A y B, 144 personas son identificadas como positivas con ambas pruebas (v. fig. 5.8E).

Recordemos que la prueba A identificó correctamente como positivas a 160 personas con la enfermedad. Como 144 de ellas fueron identificadas por ambas pruebas, 160 − 144 = 16 personas fueron identificadas correctamente

solo con la prueba A.

La prueba B identificó correctamente como positivas a 180 de las 200 personas con la enfermedad. Como 144 de ellas fueron identificadas por ambas pruebas, 180 − 144 = 36 personas fueron identificadas correctamente *solo* con la prueba B. Por tanto, como se observa en la figura 5.8F, cuando se emplean simultáneamente las pruebas A y B, la

$$\text{sensibilidad neta} = \frac{16+144+36}{200} = \frac{196}{200} = 98\%$$

Especificidad neta utilizando dos pruebas simultáneas

La siguiente pregunta que debemos plantearnos es: «¿cuál es la *especificidad neta* si se emplean las pruebas A y B *simultáneamente*?» Para poder incluir a una persona en el numerador para calcular la especificidad neta de las dos pruebas utilizadas simultáneamente, dicha persona debe ser identificada como *negativa* por *ambas pruebas*. Con el fin de calcular el numerador para la especificidad neta, necesitamos por tanto determinar cuántas personas presentaron resultados negativos en ambas pruebas. ¿Cómo hacemos esto?

La prueba A posee una especificidad del 60% y, por tanto, identifica correctamente al 60% de las 800 personas que no tienen la enfermedad (480 personas) (tabla 5.5). En la figura 5.9A, la elipse representa a las 800 personas que no tienen la enfermedad. El círculo verde en el interior de la elipse de la figura 5.9B representa a las 480 personas con resultados negativos en la prueba A. Estos son los verdaderos negativos empleando la prueba A.

Tabla 5.5

Resultados del cribado con la prueba A

Resultados del cribado	Población	
	Enfermedad	**No enfermedad**
Positivo	160	320
Negativo	40	**480**
Total	200 Sensibilidad = 80%	800 Especificidad = 60%

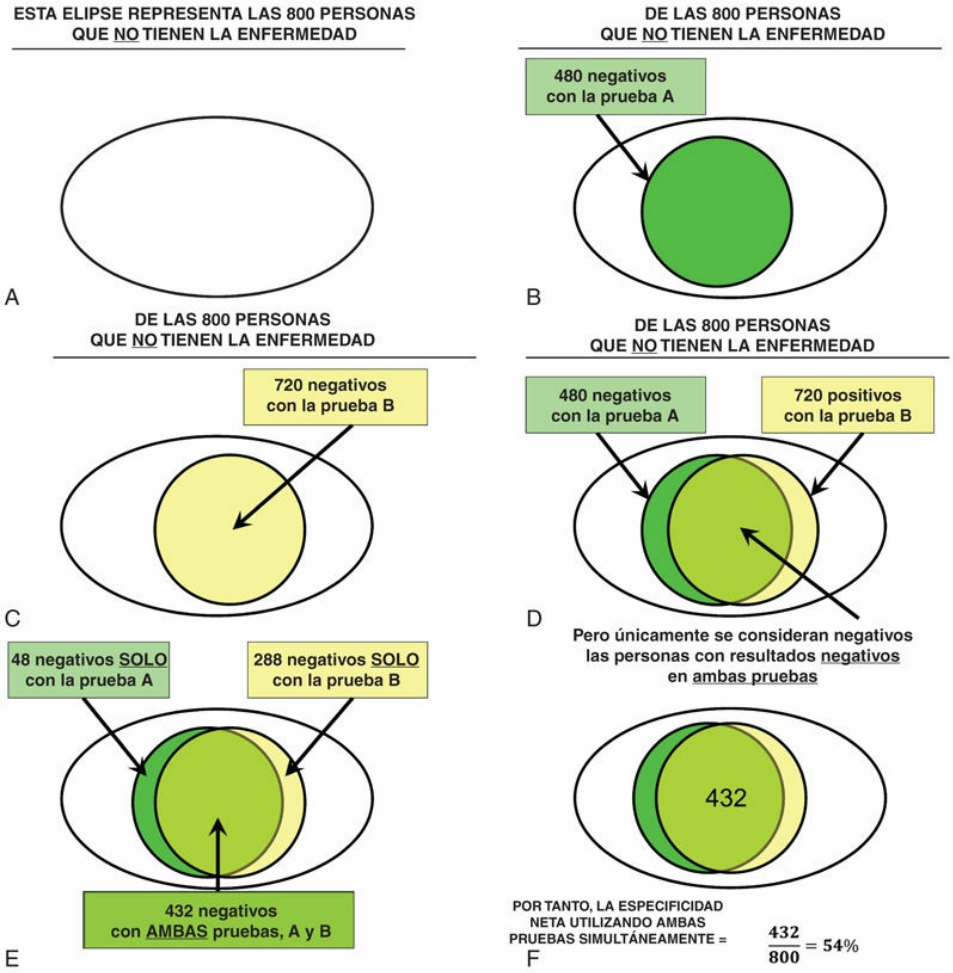


FIG. 5.9 (A-F) Especificidad neta: ejemplo hipotético de pruebas simultáneas. (V. explicación en el apartado «Sensibilidad neta utilizando dos pruebas simultáneas», pág. 104.)

La prueba B posee una especificidad del 90% y, por tanto, identifica como negativas al 90% de las 800 personas que no tienen la enfermedad (720 personas) (tabla 5.6 y círculo amarillo de la fig. 5.9C). Sin embargo, para ser identificadas como negativas en pruebas simultáneas, solo se considera que tienen resultados negativos las personas con resultados negativos en ambas pruebas (v. fig. 5.9D). Estas personas se muestran en verde claro en el área de solapamiento entre los dos círculos. La prueba B también identifica como negativas al 90% de las mismas 480 personas identificadas como negativas por la prueba A (432 personas). Por tanto, como se muestra por los círculos que se solapan, cuando se utilizan simultáneamente las pruebas A y B, 432 personas son identificadas como negativas por ambas pruebas (v. fig. 5.9E). Así, cuando se emplean simultáneamente las pruebas A y B (v. fig. 5.9F), la

$$\text{Especificidad neta} = \frac{432}{800} = 54\%$$

Tabla 5.6

Resultados del cribado con la prueba B

Resultados del cribado	Población: Enfermedad	Población: No enfermedad
Positivo	180	80
Negativo	20	**720**
Total	200 Sensibilidad = 90%	800 Especificidad = 90%

Por tanto, cuando se emplean dos pruebas simultáneas existe una ganancia neta de sensibilidad (del 80% utilizando la prueba A y el 90% utilizando la prueba B al 98% utilizando ambas pruebas simultáneamente). Sin embargo, existe una pérdida neta de especificidad (especificidad neta = 54%) respecto a cuándo se utiliza cada prueba aisladamente (especificidad del 60% con la prueba A y del 90% con la prueba B).

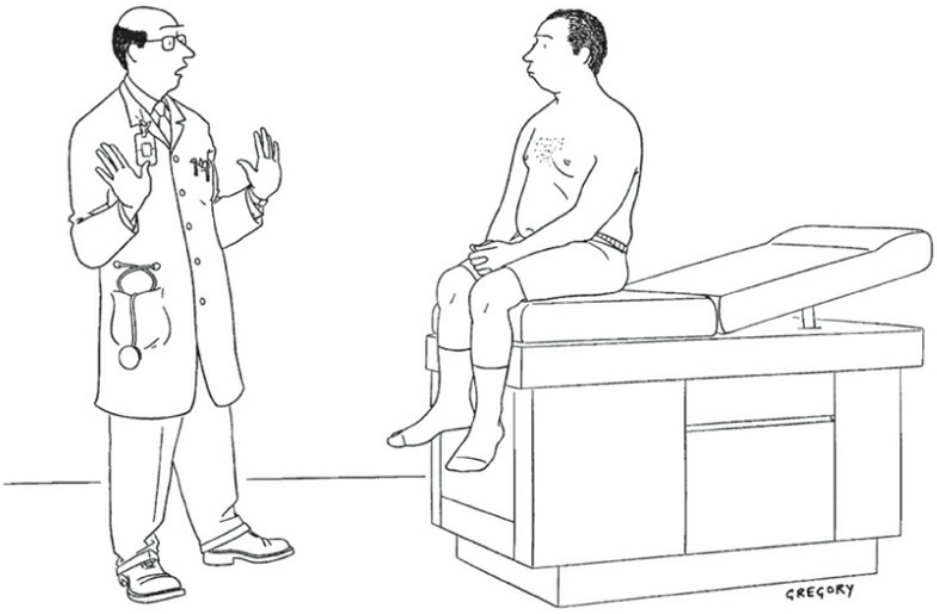
Comparación de las pruebas simultáneas y secuenciales

En un contexto clínico, a menudo se utilizan múltiples pruebas simultáneamente. Por ejemplo, un paciente ingresado en un hospital puede ser sometido a una batería de pruebas en el momento del ingreso. Cuando se

utilizan múltiples pruebas simultáneamente para detectar una enfermedad específica, generalmente se considera que el resultado de la prueba en el paciente es «positivo» si ha obtenido un resultado positivo en una o varias de las pruebas. Se considera que el resultado de las pruebas del paciente es «negativo» si los resultados de todas las pruebas son negativos. Los efectos de este abordaje sobre la sensibilidad y la especificidad difieren de los que resultan de las pruebas secuenciales. Con las pruebas secuenciales, cuando volvemos a realizar pruebas a los identificados como positivos con la primera prueba, se produce una pérdida en la sensibilidad neta y una ganancia en la especificidad neta. Cuando se emplean las pruebas simultáneas, como un individuo identificado como positivo en *una* o en *múltiples* pruebas es considerado positivo, se produce una ganancia en la sensibilidad neta. Sin embargo, para ser considerada negativa, una persona debería obtener resultados negativos en *todas* las pruebas realizadas. Como resultado, se produce una pérdida en la especificidad neta.

En resumen, como hemos visto previamente, cuando se utilizan dos pruebas secuenciales, y los individuos que han obtenido resultados positivos en la primera prueba son vueltos a explorar con la segunda prueba, se produce una pérdida neta en la sensibilidad, pero una ganancia neta de la especificidad, en comparación a cuando cada prueba se realiza aisladamente. Sin embargo, cuando se utilizan dos pruebas simultáneamente, se produce una ganancia neta de sensibilidad y una pérdida neta de especificidad, en comparación a cuando las pruebas se realizan aisladamente.

Considerando estos resultados, la decisión de utilizar pruebas secuenciales o simultáneas a menudo se basa en los objetivos de las pruebas (la prueba se realiza con fines diagnósticos o de cribado) y en función de consideraciones prácticas relacionadas con el contexto en el que se realizan las pruebas, como la duración del ingreso hospitalario, los costes y el grado de invasividad de cada prueba, así como el grado de cobertura del seguro a terceros. En la figura 5.10 se muestra a un médico afrontando la sobrecarga de información percibida.



"Whoa—way too much information."

FIG. 5.10 «¡Basta! Me está dando demasiada información.» Comentario de un médico sobre el exceso de información. (Alex Gregory/The New Yorker Collection/The Cartoon Bank.)

Valor predictivo de una prueba

Hasta ahora nos hemos preguntado cómo es de buena la prueba para identificar a las personas que tienen la enfermedad y a las que no la tienen. Este punto es importante, especialmente cuando se realizan pruebas de cribado en poblaciones de la comunidad que no presentan síntomas de la enfermedad que se está estudiando. En efecto, nos preguntamos: «Si realizamos un cribado en una población, ¿qué proporción de las personas que tienen la enfermedad serán identificadas correctamente?». Este aspecto es claramente una consideración de salud pública importante. En el contexto clínico, sin embargo, para el médico puede ser importante otra pregunta: si los resultados de la prueba son positivos en este paciente, ¿cuál es la probabilidad de que dicho paciente tenga la enfermedad? Este es el denominado *valor predictivo positivo* (VPP) de la prueba. En otras palabras, ¿qué proporción de los pacientes con resultados positivos en la prueba tienen realmente la enfermedad en cuestión? Para calcular el VPP, dividimos el número de verdaderos positivos entre el número total de personas con resultados positivos (verdaderos positivos + falsos positivos).

Volvamos al ejemplo que se muestra en la tabla 5.1, en el que se realiza un cribado en una población de 1.000 personas. Como se observa en la tabla 5.7, la tabla de 2 × 2 muestra los resultados de una prueba de cribado dicotómica en dicha población. De las 1.000 personas, el resultado de la prueba es positivo en 180; de estas 180 personas, 80 tienen la enfermedad. Por tanto, el VPP es de 80/180 = 44%.

Tabla 5.7

Valor predictivo de una prueba

Resultados del cribado	Población: Enfermedad	Población: No enfermedad	Totales	
Positivo	80	100	180	Valor predictivo positivo = $\frac{80}{180}$ = 44%
Negativo	20	800	820	Valor predictivo negativo = $\frac{800}{820}$ = 98%
Totales	100	900	1,000	

Sobre los resultados negativos de la prueba puede plantearse una pregunta paralela: «Si el resultado de la prueba es negativo, ¿cuál es la probabilidad de que este paciente no tenga la enfermedad?». Este es el *valor predictivo negativo* (VPN) de la prueba. Se calcula dividiendo el número de verdaderos negativos

entre el total de resultados negativos (verdaderos negativos + falsos negativos). Fijándonos de nuevo en el ejemplo de la tabla 5.7, la prueba arroja un resultado negativo en 820 personas, y de estas, 800 no tienen la enfermedad. Por tanto, el VPN es de 800/820 = 98%.

Cada prueba realizada por un médico (historia clínica, exploración física, pruebas de laboratorio, radiografías, electrocardiogramas y otras intervenciones) se utiliza para facilitar la labor del médico para emitir un diagnóstico correcto. Lo que se quiere saber cuando se realiza una prueba a un paciente es: «Considerando este resultado positivo de la prueba, ¿cuál es la probabilidad de que el paciente tenga la enfermedad?».

A diferencia de la sensibilidad y la especificidad de la prueba, que pueden considerarse características de la prueba que se está utilizando, el VPP se ve afectado por dos factores: la prevalencia de la enfermedad en la población estudiada y, cuando la enfermedad es infrecuente, la especificidad de la prueba que se está empleando. En las siguientes secciones se analizan estas relaciones.

Relación entre el valor predictivo positivo y la prevalencia de la enfermedad

En la siguiente exposición del valor predictivo, el término *valor predictivo* se utiliza para denotar el valor predictivo *positivo* de la prueba.

La relación entre el valor predictivo y la *prevalencia de la enfermedad* puede verse en el ejemplo mostrado en la tabla 5.8. En primer lugar, dirijamos nuestra atención a la parte superior de la tabla. Asumamos que estamos utilizando una prueba con una sensibilidad del 99% y una especificidad del 95% en una población de 1.000 personas en la que la prevalencia de la enfermedad es del 1%. Como la prevalencia es del 1%, 100 de las 1.000 personas presentan la enfermedad y 9.900 no la presentan. Con una sensibilidad del 99%, la prueba identifica correctamente a 99 de las 100 personas que tienen la enfermedad. Con una especificidad del 95%, la prueba identifica correctamente como negativas a 9.405 de las 9.900 personas que no tienen la enfermedad. Por tanto, en esta población con una prevalencia del 1%, la prueba identifica como positivas a 594 personas (99 + 495). Sin embargo, de estas 594 personas, 495 (38%) son falsos positivos y, por tanto, el VPP es de 99/594, o de tan solo el 17%.

Tabla 5.8

Relación entre la prevalencia de la enfermedad y el valor predictivo positivo

EJEMPLO: SENSIBILIDAD = 99%, ESPECIFICIDAD = 95%					
Prevalencia de la enfermedad	Resultados de la prueba	Enfermos	No enfermos	Totales	Valor predictivo positivo
1%	+	99	495	594	$\frac{99}{594} = 17\%$
	–	1	9.405	9.406	
	Totales	100	9.900	10.000	
5%	+	495	475	970	$\frac{495}{970} = 51\%$
	–	5	9.025	9.030	
	Totales	500	9.500	10.000	

Apliquemos ahora la misma prueba (con la misma sensibilidad y especificidad) a una población con una enfermedad de prevalencia más elevada (5%), como se observa en la parte inferior de la tabla 5.8. Realizando cálculos similares a los empleados en la parte superior de la tabla, el VPP es ahora del 51%. Por tanto, la mayor prevalencia en la población cribada ha causado un aumento importante del VPP utilizando la misma prueba. En la figura 5.11 se muestra un ejemplo clásico de la relación entre la prevalencia de la enfermedad y el valor predictivo. Claramente, la mayor parte de la ganancia del valor predictivo se produce cuando aumenta la prevalencia de la enfermedad en los casos en que esta es más baja.

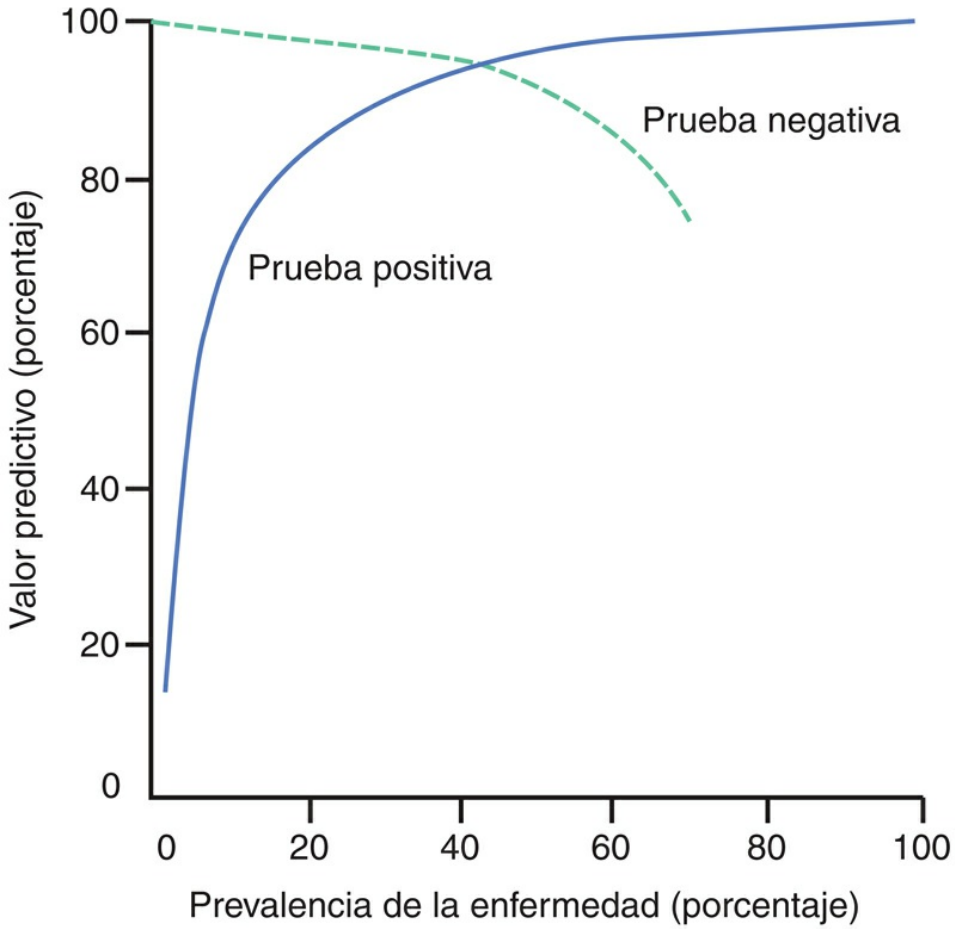


FIG. 5.11 Relación entre la prevalencia de la enfermedad y el valor predictivo en una prueba con una sensibilidad del 95% y una especificidad del 95%. (De Mausner JS, Kramer S. *Mausner and Bahn Epidemiology: An Introductory Text. Filadelfia*: WB Saunders; 1985: 221.)

¿Por qué debe interesarnos la relación entre el valor predictivo y la prevalencia de la enfermedad? Como hemos visto, cuanto más elevada sea la prevalencia, mayor será el valor predictivo. Por tanto, un programa de cribado es más productivo y económico si se dirige a una población objetivo de alto riesgo. El cribado de una población completa para una enfermedad relativamente infrecuente puede suponer un gasto de recursos y puede lograr la detección de pocos casos previamente no detectados en relación con la cantidad de esfuerzo empleado. Sin embargo, si puede identificarse un subgrupo de alto riesgo y el cribado puede centrarse en este subgrupo, es probable que el programa sea mucho más productivo. Además, una población de alto riesgo puede estar más motivada para participar en dicho programa de cribado y es más probable que adopte las acciones recomendadas si los resultados del cribado son positivos.

La relación entre valor predictivo y prevalencia de la enfermedad también muestra que los resultados de cualquier prueba deben interpretarse en el contexto de la prevalencia de la enfermedad en la población a la que pertenece el individuo. Un ejemplo interesante lo constituye la determinación de la concentración de α-fetoproteína sérica materna (AFPSM) para el diagnóstico prenatal de la espina bífida. En la figura 5.12 se muestra la

distribución de las concentraciones de AFPSM en embarazos normales no afectados y en embarazos en los que el feto presentó síndrome de Down, espina bífida, que es un defecto del tubo neural, o anencefalia. A los efectos de este ejemplo, nos centraremos en las curvas para los embarazos no afectados y en aquellos con espina bífida. Aunque la distribución de estas dos curvas es bimodal, existe un tramo en el que la curva se solapa y en dicho tramo no siempre está claro a qué curva pertenecen la madre y el feto. Si la AFPSM se encuentra en el rango más alto para un embarazo no afectado, la prevalencia verdadera de espina bífida será baja para el mismo rango. Por lo tanto, tal superposición de la AFPSM en los embarazos no afectados y en aquellos con fetos con espina bífida ha llevado a que la prueba tenga un VPP muy bajo, de solo del 2 al 6%[1].

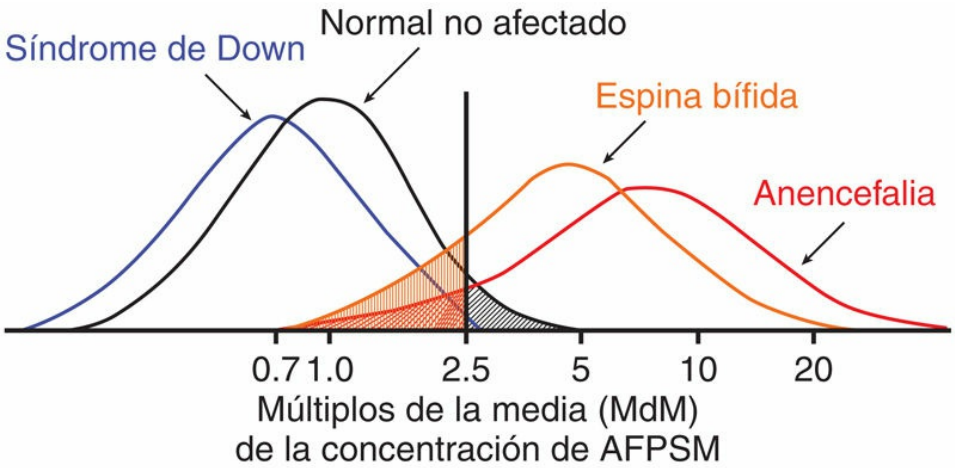


FIG. 5.12 Distribución sérica materna de alfa-fetoproteína *(AFPSM)* para embarazos únicos a las 15 a 20 semanas. Se espera que el valor de corte de pantalla de 2,5 múltiplos de la media dé como resultado una tasa de falsos positivos de hasta el 5% *(área sombreada negra)* y tasas de falsos negativos de hasta el 20% para la espina bífida *(área sombreada naranja)* y 10% para la anencefalia *(área sombreada roja)*. (Modificado de Prenatal diagnosis. En: Cunningham F, Leveno KJ, Bloom SL, et al., eds. Williams Obstetrics. 24th ed. New York: McGraw-Hill; 2013. http://accessmedicine.mhmedical.com.ezp.welch.jhmi.edu/content.aspx?bookid=1057&sectionid=59789152. Consultado del 19 de junio de 2017.)

Es posible que la misma prueba pueda tener un valor predictivo muy diferente cuando se realiza en una población de alto riesgo (prevalencia elevada) o en una población de bajo riesgo (prevalencia baja). Las implicaciones clínicas de esta observación son claras: una mujer puede tomar la decisión de interrumpir un embarazo y un médico puede aconsejar a dicha mujer basándose en los resultados de la prueba. Sin embargo, el mismo resultado de la prueba puede interpretarse de modo diferente, dependiendo de si la mujer pertenece a un grupo de mujeres de alto o bajo riesgo, lo que se reflejará en el VPP de la prueba. Por tanto, el resultado de la prueba de modo aislado puede no ser suficiente para servir de guía sin tener en cuenta las otras consideraciones que acabamos de describir.

Los siguientes ejemplos reales destacan la importancia de este aspecto:

El líder de un sindicato de bomberos consultó a un cardiólogo universitario porque el médico de su unidad había leído un artículo en una revista médica de impacto que describía que cierto hallazgo electrocardiográfico era muy predictivo de la existencia de cardiopatía coronaria grave, generalmente no reconocida. Basándose en este artículo, el médico de la unidad estaba apartando de tareas activas a muchos bomberos jóvenes, en buena condición física. El cardiólogo leyó el artículo y observó que el estudio se había efectuado en pacientes hospitalizados.

¿Cuál fue el problema? Como los pacientes hospitalizados poseen una prevalencia mucho mayor de cardiopatías que el grupo de bomberos jóvenes sanos, el médico del cuerpo de bomberos había tomado erróneamente el elevado valor predictivo obtenido al estudiar una población con una gran prevalencia y lo había aplicado incorrectamente a una población de bomberos jóvenes de baja prevalencia, en los que la misma prueba habría arrojado un valor predictivo mucho más bajo.

Otro ejemplo:

Un médico visitó a su internista general para un examen médico anual rutinario, que incluía una exploración de heces para descartar sangre oculta. Una de las tres muestras de heces examinadas en la prueba fue positiva. El internista dijo a su paciente-médico que el resultado no era significativo porque de manera regular encontraba muchos resultados falsos positivos en su ajetreada consulta. La prueba se repitió en tres nuevas muestras de heces y todas fueron ahora negativas. Sin embargo, percibiendo la preocupación persistente de su paciente, el internista remitió a su paciente-médico a un gastroenterólogo. El gastroenterólogo dijo: «Según mi experiencia, el hallazgo positivo en heces es grave. Dicho hallazgo casi siempre se asocia con trastornos gastrointestinales patológicos. Los resultados negativos posteriores no significan nada, porque podría tener un tumor que únicamente sangra intermitentemente.»

¿Quién tenía razón en este ejemplo? La respuesta es que tanto el internista general como el gastroenterólogo tenían razón. El internista emitió su valoración del valor predictivo basándose en su experiencia en su práctica médica general, una población con una prevalencia baja de enfermedades gastrointestinales graves. Por otra parte, el gastroenterólogo emitió su valoración del valor predictivo de la prueba basándose en su experiencia de

pacientes remitidos, una consulta en la que la mayoría de los pacientes son remitidos debido a la posibilidad de que padezcan una enfermedad gastrointestinal grave (una población con una prevalencia elevada).

Relación entre el valor predictivo positivo y la especificidad de la prueba

En la siguiente exposición, el término *valor predictivo* se utiliza para referirse al VPP de la prueba.

Un segundo factor que afecta al valor predictivo de una prueba es su *especificidad*. Daremos ejemplos de esto primero en forma gráfica y luego en forma de tabla. En la figura 5.13A-D se presentan en un diagrama los resultados del cribado de una población; sin embargo, las tablas 2 × 2 de estas figuras difieren de las presentadas en figuras anteriores. Cada celdilla se dibuja con su tamaño proporcional a la población que representa. En cada figura, las celdillas que representan a personas con resultados positivos en la prueba están coloreadas de azul; estas son las celdillas que serán utilizadas para calcular el VPP.

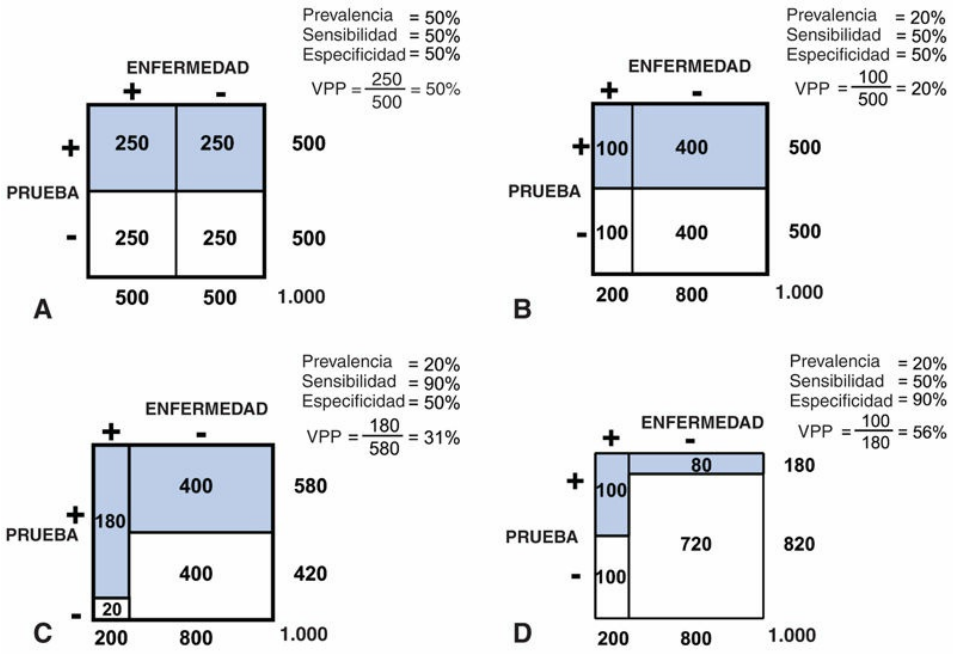


FIG. 5.13 (A-D) Relación entre la especificidad y el valor predictivo positivo *(VPP)*. (V. explicación en el apartado «Relación entre el valor predictivo positivo y la especificidad de la prueba», en la página 109.)

En la figura 5.13A se muestra la población cribada que es utilizada en nuestro análisis: una población de 1.000 personas en la que la prevalencia es del 50%, es decir, 500 personas tienen la enfermedad y 500 no la tienen. Al analizar esta figura, también asumimos que la prueba de cribado que fue utilizada posee una sensibilidad del 50% y una especificidad el 50%. Como el

resultado fue positivo en 500 personas, y 250 de estas tienen la enfermedad, el valor predictivo es de 250/500 o del 50%.

Afortunadamente, la prevalencia de la mayoría de las enfermedades es mucho menor del 50%; por lo general tratamos con enfermedades relativamente infrecuentes. Por tanto, la figura 5.13B asume una prevalencia más baja, del 20% (aunque incluso esta cifra sería una prevalencia inusualmente alta para la mayoría de las enfermedades). Tanto la sensibilidad como la especificidad siguen siendo del 50%. Ahora solo 200 de las 1.000 personas tienen la enfermedad y la línea vertical que separa a los enfermos de los no enfermos se ha desplazado a la izquierda. El valor predictivo ahora se calcula así: 100/500 = 20%.

Dado que estamos realizando un cribado en una población con una tasa de prevalencia más baja, ¿podemos mejorar el valor predictivo? ¿Cuál sería el efecto en el valor predictivo si aumentásemos la sensibilidad de la prueba? En la figura 5.13C se muestran los resultados cuando mantenemos la prevalencia del 20% y la especificidad del 50% pero aumentamos la sensibilidad al 90%. El valor predictivo es ahora 180/850 = 31%, un aumento modesto.

¿Y qué pasaría si en vez de aumentar la sensibilidad de la prueba aumentamos su especificidad? En la figura 5.13D se muestran los resultados cuando se mantiene la prevalencia al 20% y la sensibilidad al 50% pero aumentamos la especificidad al 90%. El valor predictivo ahora es de 100/180 = 56%. Por tanto, el aumento de la especificidad produce un mayor aumento del valor predictivo que el logrado con el mismo aumento de la sensibilidad.

¿Por qué la especificidad tiene mayor influencia sobre el valor predictivo que la sensibilidad? La respuesta es clara si observamos estas figuras. Como estamos tratando con enfermedades infrecuentes, la mayor parte de la población se encuentra a la derecha de la línea vertical. Por tanto, cualquier cambio a la derecha de la línea vertical afecta a un mayor número de personas que un cambio comparable a la izquierda de la línea. Así, un cambio en la especificidad produce un mayor efecto sobre el valor predictivo que un cambio comparable en la sensibilidad. Si estuviéramos tratando con una enfermedad de gran prevalencia, la situación sería diferente.

El efecto de los cambios en la especificidad sobre el valor predictivo también se observa en la tabla 5.9, en una forma similar a la utilizada en la tabla 5.8. Como se observa en este ejemplo, incluso con un 100% de sensibilidad, un cambio en la especificidad del 70% al 95% ejerce un efecto espectacular sobre el VPP.

Tabla 5.9

Relación entre la especificidad y el valor predictivo positivo

EJEMPLO: PREVALENCIA = 10%, SENSIBILIDAD = 100%					
Especificidad	**Resultados de la prueba**	**Enfermos**	**No enfermos**	**Totales**	**Valor predictivo**
70%	+	1.000	2.700	3.700	$\frac{1.000}{3.700} = 27\%$
	–	0	6.300	6.300	
	Totales	1.000	9.000	10.000	
95%	+	1.000	450	1.450	$\frac{1.000}{1.450} = 69\%$
	–	0	8.550	8.550	
	Totales	1.000	9.000	10.000	

Fiabilidad (repetibilidad) de las pruebas

Consideremos otro aspecto de la valoración de las pruebas diagnósticas y de cribado: si una prueba es fiable o repetible. ¿Los resultados obtenidos podrían reproducirse (obteniendo el mismo resultado) si se repitiese la prueba? Claramente, con independencia de la sensibilidad y la especificidad de una prueba, si los resultados de la prueba no son reproducibles, el valor y la utilidad de la prueba son mínimos. El resto del presente capítulo analizará la fiabilidad o repetibilidad de las pruebas diagnósticas y de cribado. Los factores que contribuyen a la variación entre los resultados de la prueba se analizan en primer lugar: variación intraindividual (variaciones en un mismo individuo), variación intraobservador (variación en la lectura de los resultados de la prueba por el mismo observador) y variación interobservador (variación entre varias personas que analizan los resultados de la prueba).

Variación intraindividual

Los valores obtenidos al medir muchas características humanas a menudo varían a lo largo del tiempo, incluso durante un periodo corto de 24 horas o durante periodos más prolongados, como las variaciones estacionales. En la figura 5.14 se muestran los cambios en las mediciones de la presión arterial a lo largo de un periodo de 24 horas en 28 personas normotensas. La variabilidad a lo largo del tiempo es considerable. Este hecho, así como las condiciones en las que se realizan ciertas pruebas (p. ej., poco tiempo tras una comida o tras realizar ejercicio, si se realiza en casa o en la consulta del médico), claramente pueden arrojar diferentes resultados en la misma persona. Por tanto, a la hora de valorar los resultados de cualquier prueba, es importante considerar las condiciones en las que se realizó la prueba, incluida la hora del día.

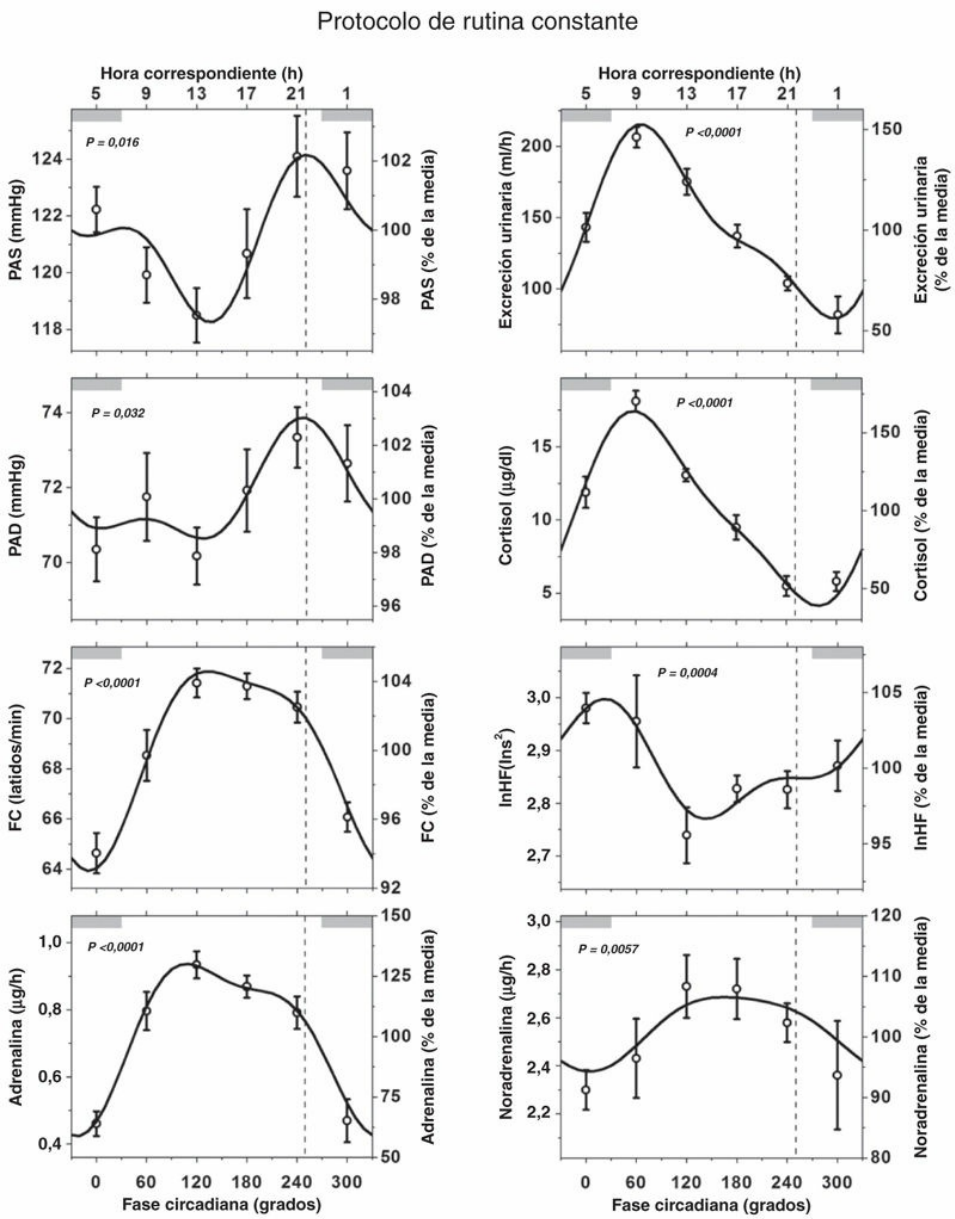


FIG. 5.14 Variación circadiana endógena en la presión arterial durante el protocolo de rutina constante. *PAD*, presión arterial diastólica; *FC*, frecuencia cardiaca; *PAS*, presión arterial sistólica. (De Shea SA, Hilton MF, Hu K, et al. Existence of an endogenous circadian blood pressure rhythm in humans that peaks in the evening. *Circ Res.* 2011;108:980–984.)

Variación intraobservador

En ocasiones se producen variaciones entre dos o más lecturas de los mismos resultados de la prueba valorada por un mismo observador. Por ejemplo, un radiólogo que interprete el mismo grupo de radiografías en dos ocasiones diferentes puede interpretar una o más de las radiografías de modo diferente la segunda vez. Las pruebas y las exploraciones se diferencian según el grado con el que entran en juego factores subjetivos en las conclusiones del observador; cuanto mayor sea el grado de subjetividad en las lecturas, mayor será la probabilidad de que se produzca una variación intraobservador en

ellas (fig. 5.15).



FIG. 5.15 «Esta ya es una segunda opinión. Al principio pensaba que tenía otra cosa.» Una visión de las segundas opiniones. (Leo Cullum/The New Yorker Collection/The Cartoon Bank.)

Variación interobservador

Otra consideración importante es la variación entre observadores. Dos examinadores a menudo no obtienen el mismo resultado. El grado de concordancia o discordancia entre observadores es un aspecto importante, ya sea si consideramos una exploración física, pruebas de laboratorio u otras técnicas de evaluación de características humanas. Necesitamos, por tanto, ser capaces de expresar el grado de concordancia en términos cuantitativos.

Porcentaje de concordancia

En la tabla 5.10 se muestra un esquema para examinar la variación entre observadores. Dos observadores fueron encargados de clasificar cada resultado de una prueba en una de las siguientes cuatro categorías: anormal, sospechoso, dudoso y normal. Este diagrama podría aplicarse, por ejemplo, a las lecturas realizadas por dos radiólogos. En este diagrama, las lecturas del observador 1 se presentan en formato de tabulación cruzada con las del observador 2. El número de lecturas en cada celdilla viene indicado por una

letra del alfabeto. Así, A radiografías fueron consideradas anormales por ambos radiólogos. C radiografías fueron consideradas anormales por el radiólogo 2 y dudosas por el radiólogo 1. M radiografías fueron consideradas anormales por el radiólogo 1 y normales por el radiólogo 2.

Tabla 5.10

Variación por observador o instrumento: porcentaje de concordancia

	Lectura n.º 1			
Lectura n.º 2	Anormal	Sospechosa	Dudosa	Normal
Anormal	A	+ B	C	D
Sospechosa	E	F	+ G	H
Dudosa	I	J	K	+ L
Normal	M	N	O	P

$$\text{Porcentaje de concordancia} = \frac{A + F + K + P}{\text{Lecturas totales}} \times 100$$

Como se observa en la tabla 5.10, para calcular el porcentaje de concordancia global, sumamos los números de todas las celdillas en las que concordaron las interpretaciones de ambos radiólogos (A + F + K + P), dividimos dicha suma entre el número total de radiografías interpretadas y multiplicamos el resultado por 100 para obtener un porcentaje. En la figura 5.16A se muestra el uso de este abordaje para una prueba cuyos resultados posibles son «positivos» o «negativos».

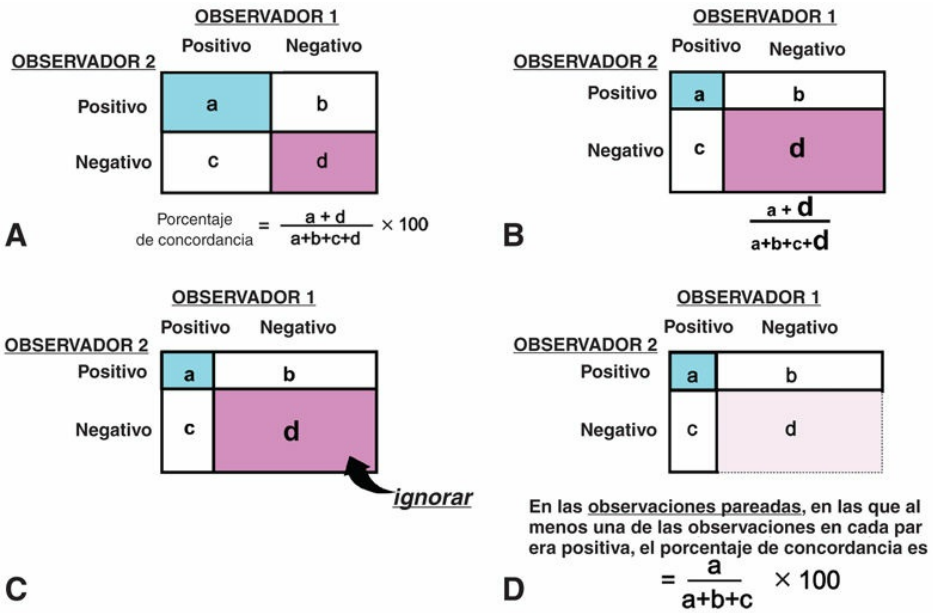


FIG. 5.16 Cálculo del porcentaje de concordancia entre dos observadores. (A) Porcentaje de concordancia cuando se examinan observaciones pareadas entre el observador 1 y el observador 2. (B) Porcentaje de concordancia cuando se examinan observaciones pareadas entre el observador 1 y el observador 2, teniendo en cuenta que la celdilla d (concordancia en los negativos) es muy grande. (C) Porcentaje de concordancia cuando se examinan observaciones pareadas entre el observador 1 y el observador 2, ignorando la celdilla d. (D) Porcentaje de concordancia cuando se examinan observaciones pareadas entre

el observador 1 y el observador 2, utilizando únicamente las celdillas a, b y c para el cálculo.

Por lo general, la mayoría de las personas en las que se realizan pruebas obtienen resultados negativos. Esto se expone en la figura 5.16B, en la que el tamaño de cada celdilla guarda proporción con el número de personas que hay en ella. Probablemente exista una concordancia importante entre los dos observadores acerca de estos individuos, negativos o normales (celdilla d). Así, cuando se calcula el porcentaje de concordancia para todos los sujetos del estudio, su valor puede ser alto debido únicamente al elevado número de hallazgos claramente negativos (celdilla d) en los que concuerdan los observadores. El valor alto puede ocultar, por tanto, una gran falta de concordancia entre los observadores en la identificación de los sujetos que son considerados positivos por al menos un observador.

Un abordaje de este problema, expuesto en la figura 5.16C, es no tener en cuenta a los sujetos etiquetados como negativos por ambos observadores (celdilla d) y calcular el porcentaje de concordancia utilizando como denominador únicamente a los sujetos considerados anormales por al menos uno de los observadores (celdillas a, b y c) (fig. 5.16D).

Así, en las observaciones pareadas en las que al menos uno de los hallazgos de cada par fue positivo, es aplicable la siguiente ecuación:

$$\text{Porcentaje de concordancia} = \frac{a}{a+b+c} \times 100$$

Estadístico kappa

El porcentaje de concordancia entre dos observadores a menudo es valioso para valorar la calidad de sus observaciones. El grado de concordancia entre dos observadores, como, por ejemplo, dos médicos o dos enfermeras, a menudo es un índice importante de la calidad de la asistencia sanitaria que se está proporcionando. Sin embargo, el porcentaje de concordancia entre dos observadores no depende completamente de la calidad de su formación o su experiencia. En el grado de concordancia también influye de manera importante el hecho de que, aunque los dos observadores utilicen criterios completamente diferentes para identificar a sujetos como positivos o negativos, cabría esperar que los observadores coincidieran en las observaciones realizadas, al menos en algunos de los participantes, únicamente debido al azar. Lo que realmente queremos saber es cuánto mejor es su grado de concordancia que el que resultaría debido únicamente al azar. La respuesta a esta pregunta presumiblemente nos dirá, por ejemplo, hasta qué punto la formación y la práctica de los observadores mejoraron la calidad de sus observaciones de modo que el porcentaje de concordancia entre ellos aumentó más de lo que cabría esperar únicamente debido al azar.

Esto puede demostrarse intuitivamente en el siguiente ejemplo: usted es el jefe de un servicio de radiología que un día carece de suficiente personal y todavía tiene pendiente la interpretación de un gran número de radiografías de tórax. Para solucionar el problema, sale a la calle y le pide a algunos residentes del vecindario, sin formación en biología ni en medicina, que interpreten las radiografías que no están informadas y que valoren si son positivas o negativas. La primera persona ojea el montón de radiografías y las interpreta aleatoriamente como positiva, negativa, negativa, positiva, etc. La segunda persona hace lo mismo, siguiendo el mismo patrón, pero de manera completamente independiente respecto a la primera. Dado que ambas personas no poseen conocimientos, criterios o estándares para interpretar radiografías, ¿concordarán sus valoraciones sobre una radiografía específica? La respuesta es claramente afirmativa; en algunos casos coincidirán, únicamente debido al azar.

Sin embargo, si queremos saber cómo de bien han interpretado las radiografías dos observadores, podríamos preguntarnos: «¿Hasta qué punto coinciden sus interpretaciones *más allá de lo que cabría esperar únicamente por el azar*?». En otras palabras, ¿hasta qué punto la concordancia entre los dos observadores supera el grado de concordancia que resultaría únicamente por el azar? Un abordaje para responder a esta pregunta es calcular el estadístico kappa, propuesto por Cohen en 1960[2]. En esta sección analizaremos primero el fundamento del estadístico kappa y las preguntas para cuyas respuestas se diseñó el estadístico kappa. A continuación se expone un cálculo detallado del estadístico kappa para que sirva de ejemplo para los lectores intrépidos. Incluso aunque usted no siga los cálculos detallados que se presentan, es importante asegurarse de que ha comprendido el significado del estadístico kappa, pues se utiliza con frecuencia en la medicina clínica y en el ámbito de la salud pública.

Fundamento del estadístico kappa

Con el fin de comprender kappa, nos planteamos dos preguntas. La primera: «¿Cuánto mejor es la concordancia entre las interpretaciones de los observadores de lo que cabría esperar únicamente por el azar?». Esto puede calcularse como el porcentaje de concordancia observado menos el porcentaje de concordancia que cabría esperar únicamente por el azar. Este es el numerador de kappa:

(Porcentaje de concordancia observado)
−(Porcentaje de concordancia esperado únicamente por el azar)

Nuestra segunda pregunta es: «¿Cuánto es lo máximo que los dos

observadores podrían haber mejorado su concordancia sobre la concordancia que cabría esperar solo por el azar?». Claramente, el máximo de concordancia sería el 100% (concordancia total: los dos observadores coinciden completamente). Por tanto, lo máximo que podemos esperar que sean capaces de mejorar (el denominador de kappa) sería:

$$100 - (\text{Porcentaje de concordancia esperado únicamente por el azar})$$

Kappa expresa el grado en el que la concordancia observada supera a la que cabría esperar únicamente por el azar (es decir, el porcentaje de concordancia observado menos el porcentaje de concordancia esperado únicamente por el azar) [numerador] relativo al máximo que se podría esperar que los observadores mejorasen su concordancia (es decir, 100% menos el porcentaje de concordancia esperado únicamente por el azar) [denominador].

Por tanto, kappa cuantifica el grado en el que la concordancia observada lograda por los observadores supera a la que cabría esperar únicamente por el azar, y lo expresa como la proporción de la mejoría máxima que podría producirse más allá de la concordancia esperada únicamente por el azar. El estadístico kappa puede definirse por la siguiente ecuación:

$$\text{Kappa} = \frac{\left(\begin{array}{c}\text{Porcentaje de}\\ \text{concordancia observado}\end{array}\right) - \left(\begin{array}{c}\text{Porcentaje de}\\ \text{concordancia esperado}\\ \text{únicamente por el azar}\end{array}\right)}{100\% - \left(\begin{array}{c}\text{Porcentaje de concordancia}\\ \text{esperado únicamente por el azar}\end{array}\right)}$$

Cálculo del estadístico kappa: un ejemplo

Para calcular el numerador de kappa, primero debemos calcular el grado de concordancia que podría esperarse únicamente por el azar. Como ejemplo, consideremos los datos sobre densidad mamaria proporcionados por la clasificación radiológica de la densidad mamaria en imágenes sintéticas 2D en comparación con las mamografías digitales 2D publicados por Alshafeiy y cols. en un estudio sobre 309 casos[3].

La primera pregunta es: «¿Cuál es la concordancia observada entre los dos

tipos de mamografías?». La mamografía sintética 2D identificó 179 (o 58%) de los 309 estudios mamarios como no densos y 130 (o 42%) de las imágenes como densas. La mamografía digital 2D identificó 182 (o 59%) de todas las imágenes como no densas (168 en concordancia con la sintética 2D) y 127 (o 41%) de las imágenes como densas (116 en concordancia con la sintética 2D). El porcentaje de concordancia se calcula con la siguiente ecuación:

$$\text{Porcentaje de concordancia observado} = \frac{168+116}{309} = 91,9\%$$

Es decir, los dos dispositivos de mamografía proporcionaban la misma clasificación de imagen mamaria en el 91,9% de las lecturas.

La siguiente pregunta es: «Si los dos tipos de mamografía hubieran usado conjuntos de criterios completamente diferentes para clasificar una imagen mamaria como densa en lugar de no densa, ¿cuánta concordancia cabría esperar *únicamente debido al azar*?» La mamografía sintética 2D consideró que el 58% de las 309 imágenes (179 imágenes) eran no densas y el 42% (130 imágenes) densas. Si estas lecturas hubieran usado criterios independientes de los utilizados por la mamografía digital 2D, esperaríamos que la mamografía sintética 2D considerara como no densas el 58% de las imágenes que la digital había identificado como densas y el 58% de las imágenes que la mamografía digital 2D había identificado como densas. Por lo tanto, esperaríamos que el 58% (73,44) de las 182 imágenes identificadas como no densas por la mamografía digital 2D se identificaran como no densas por la mamografía 2D sintética, y que el 58% (73,44) de las 127 imágenes identificadas como densas por la mamografía digital 2D también serían identificadas como no densas por la mamografía sintética 2D (v. fig. 5.16C). De las 127 imágenes consideradas densas por la mamografía digital 2D, el 42% (53,34) también serían clasificadas como densas por la mamografía sintética 2D.

Así, la concordancia esperada únicamente por el azar sería

$$\frac{105,56}{309} + \frac{53,34}{309} = \frac{158,9}{309} = 51,4\%$$

de todas las imágenes analizadas.

Tras calcular las cifras necesarias para el numerador y el denominador, ya podemos calcular kappa como sigue:

$$\text{Kappa} = \frac{\left(\begin{array}{c}\text{Porcentaje de}\\ \text{concordancia}\\ \text{observado}\end{array}\right) - \left(\begin{array}{c}\text{Porcentaje de}\\ \text{concordancia}\\ \text{esperado únicamente}\\ \text{por el azar}\end{array}\right)}{100\% - \left(\begin{array}{c}\text{Porcentaje de concordancia}\\ \text{esperado únicamente por el azar}\end{array}\right)}$$

$$= \frac{91,9 - 51,4}{100 - 51,4} = \frac{40,5}{48,6} = 83,3\%$$

Landis y Koch[4] sugieren que un kappa mayor de 0,75 representa una concordancia excelente más allá del azar, un kappa menor de 0,40 representa una concordancia baja y un kappa entre 0,40 y 0,75 representa una concordancia de intermedia a buena. Fleiss[5] ha estudiado la significación estadística de kappa. Existe gran controversia acerca del uso apropiado de kappa, un tema estudiado por MacLure y Willet[6].

Validez de las pruebas con resultados multicategóricos

La validez, como concepto, se puede aplicar a cualquier prueba frente a un método de referencia. Como explicamos anteriormente, utilizamos la sensibilidad/especificidad para validar los resultados de las pruebas con resultados dicotómicos frente a un método de referencia. ¿Qué ocurre con las pruebas con resultados multicategóricos? En este caso, podemos calcular el estadístico kappa, que hemos demostrado anteriormente que se trata de una herramienta que sirve para evaluar la fiabilidad.

Validez de los autoinformes

A menudo obtenemos información sobre la salud y el estado de la enfermedad preguntando directamente a los pacientes o a los participantes de un estudio sobre su historial médico, sus hábitos y otros factores de interés. La mayoría de las personas hoy en día conocen su fecha de nacimiento, por lo que la evaluación de la edad generalmente no se acompaña de errores significativos. Sin embargo, muchas personas infraestiman su peso, sus prácticas de consumo de alcohol y tabaco, y otros tipos de riesgos. Los autoinformes sobre prácticas sexuales se consideran sujetos a errores considerables. Para superar estos sesgos de informe, los biomarcadores se han utilizado con frecuencia en estudios de campo. Por ejemplo, Zenilman y cols.[7] utilizaron un ensayo de reacción en cadena de la polimerasa (PCR) para detectar fragmentos del cromosoma Y en muestras vaginales obtenidas

mediante hisopos por las propias pacientes. Este biomarcador puede detectar el coito en mujeres durante un periodo de 2 semanas, y puede validar los autoinformes acerca del uso de preservativos[8].

Relación entre validez y fiabilidad

Para finalizar este capítulo, comparemos la validez y la fiabilidad utilizando una representación gráfica.

La línea horizontal de la figura 5.17 es una escala de los valores para una variable determinada, como la concentración de glucosa en sangre, en la que se indica el valor real. Los resultados obtenidos con la prueba se muestran mediante la curva. La curva es estrecha, lo que indica que los resultados son bastante fiables (repetibles); desafortunadamente, sin embargo, se agrupan lejos del valor real, por lo que no son válidos. En la figura 5.18 se muestra una curva que es ancha y, por tanto, poco fiable. Sin embargo, los valores obtenidos se agrupan alrededor del valor real, por lo que son válidos. Claramente, lo que querríamos lograr son resultados válidos y fiables (fig. 5.19).

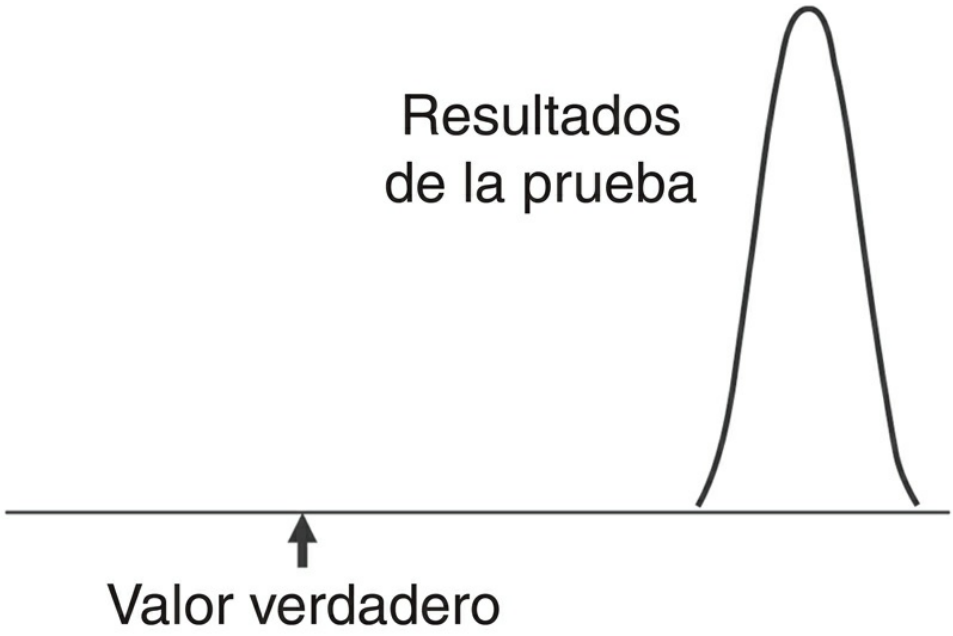


FIG. 5.17 Gráfico de los resultados de una prueba hipotética que son fiables, pero no válidos.

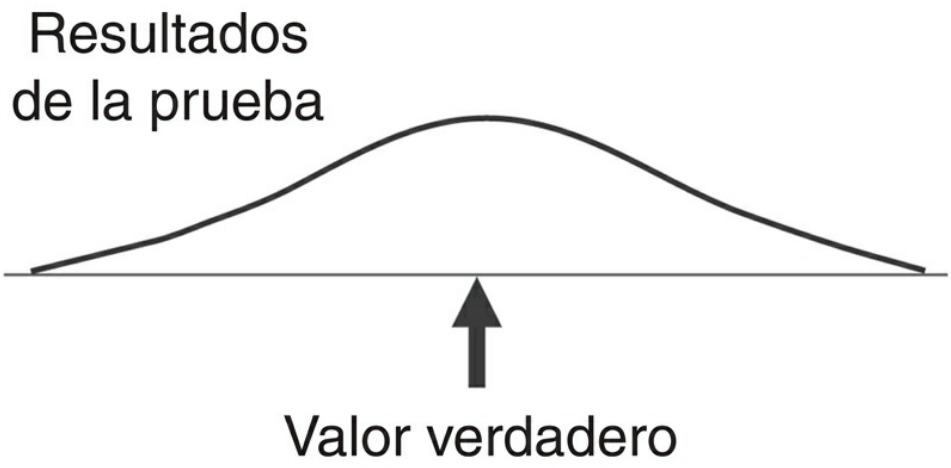


FIG. 5.18 Gráfico de los resultados de una prueba hipotética que son válidos, pero no fiables.

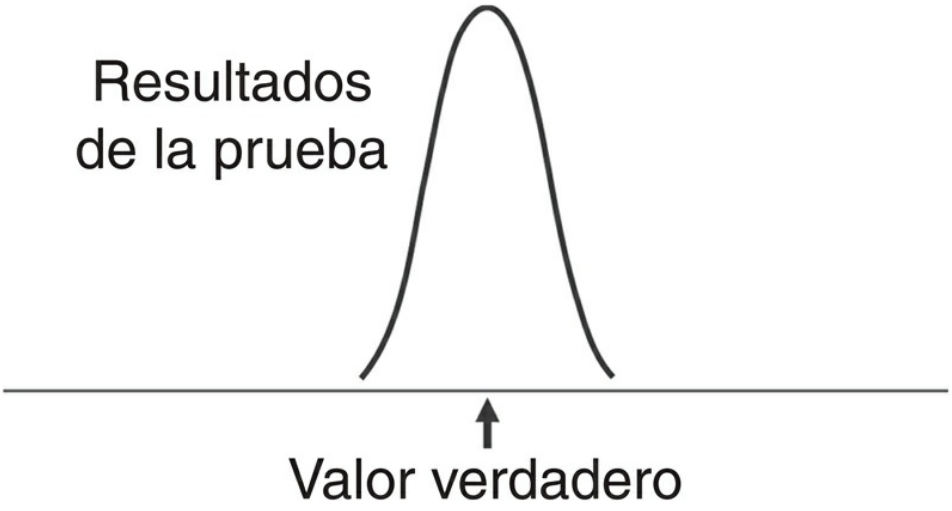


FIG. 5.19 Gráfico de los resultados de una prueba hipotética que son válidos y fiables.

Es importante destacar que en la figura 5.18, en la que la distribución de los resultados es una curva ancha centrada sobre el valor real, describimos los resultados como válidos. Sin embargo, los resultados son válidos solo para un grupo (es decir, tienden a agruparse alrededor del valor real). No hay que olvidar que lo que puede ser válido para un grupo o una población puede no serlo para un individuo en un contexto clínico. Cuando la fiabilidad o repetibilidad de una prueba es baja, la validez de la prueba para un individuo concreto también puede ser mala. Por tanto, es importante tener en cuenta la distinción entre validez grupal y validez individual a la hora de valorar la calidad de las pruebas diagnósticas y de cribado.

Conclusión

Este capítulo ha estudiado la validez de las pruebas diagnósticas y de cribado analizando la sensibilidad y la especificidad, el valor predictivo y la fiabilidad o repetibilidad. Claramente, con independencia de la sensibilidad y la especificidad de una prueba, si sus resultados no pueden repetirse, la prueba es poco útil. Por tanto, todas estas características deben tenerse en cuenta cuando se valora una prueba, junto con la finalidad para la que se quiere utilizar dicha prueba.

Bibliografía

1. Prenatal diagnosis. In: Cunningham F, Leveno KJ, Bloom SL, et al, eds. *Williams Obstetrics*. 24th ed. New York: McGraw-Hill; 2013 [fecha de última consulta: 19 de junio de 2017]. http://accessmedicine.mhmedical.com.ezp.welch.jhmi.edu/con bookid=1057&sectionid=59789152.
2. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20:37.
3. Alshafeiy TI, Wadih A, Nicholson BT, et al. Comparison between digital and synthetic 2D mammograms in breast density interpretation. *AJR Am J Roentgenol*. 2017;209:W36–W41.
4. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–174.
5. Fleiss JL. Statistical Methods for Rates and Proportions. 2nd ed. New York: John Wiley & Sons; 1981.
6. MacLure M, Willett WC. Misinterpretation and misuse of the kappa statistic. *Am J Epidemiol*. 1987;126:161–169.
7. Zenilman JM, Yeunger J, Galai N, et al. Polymerase chain reaction detection of Y chromosome sequences in vaginal fluid: preliminary studies of a potential biomarker. *Sex Transm Dis*. 2005;32:90–94.
8. Ghanem KG, Melendez JH, McNeil-Solis C, et al. Condom use and vaginal Y-chromosome detection: the specificity of a potential biomarker. *Sex Transm Dis*. 2007;34:620.

Véanse las preguntas de repaso en las páginas Véanse las preguntas de repaso en las páginas 121-122.

Apéndices del capítulo 5

El texto del capítulo 5 se centra en la lógica que respalda el cálculo de la sensibilidad, la especificidad y el valor predictivo. En el Apéndice 1 se resumen medidas de validez para las pruebas de cribado para detectar la ausencia o la presencia de una enfermedad determinada; primero se dedican una páginas en el texto a las medidas y a la interpretación de cada medida. Los que prefieran ver las fórmulas de cada medida pueden consultar la columna derecha de esta tabla; no obstante, no son esenciales para comprender la lógica que respalda el cálculo de cada medida.

Apéndice 1 del capítulo 5. Medidas de la validez de una prueba y su interpretación

	Medidas de validez de una prueba	Números de página	Interpretación	Fórmula
Pruebas de cribado INDIVIDUALES	Sensibilidad	95	La proporción de los que *tienen* la enfermedad en los que la prueba es *positiva*	$\frac{VP}{VP+FN}$
	Especificidad	95	La proporción de los que *no tienen* la enfermedad en los que la prueba es *negativa*	$\frac{VN}{VN+FP}$
	Valor predictivo positivo	106-107	La proporción de los que tienen resultados *positivos* que tienen la enfermedad	$\frac{VP}{VP+FP}$
	Valor predictivo negativo	107	La proporción de los que tienen resultados *negativos* que NO tienen la enfermedad	$\frac{VN}{VN+FN}$
Pruebas de cribado SECUENCIALES	Sensibilidad neta	99-101	La proporción de los que *tienen* la enfermedad y obtienen resultados *positivos* en AMBAS pruebas (1 y 2)	(Sensibilidad de la prueba1) ×(Sensibilidad de la prueba 2)
	Especificidad neta	99-101	La proporción de los que *no tienen* la enfermedad y obtienen resultados *negativos* en la prueba 1 *o* en la prueba 2	$\left(\begin{array}{c}\text{Especificidad}\\ \text{de la prueba 1}\\ +\\ \text{Especificidad}\\ \text{de la prueba 2}\end{array}\right) - \left(\begin{array}{c}\text{Especificidad}\\ \text{de la prueba 1}\\ \times\\ \text{Especificidad}\\ \text{de la prueba 2}\end{array}\right)$
Pruebas de cribado SIMULTÁNEAS	Sensibilidad neta	101-104	La proporción de los que *tienen* la enfermedad y obtienen resultados *positivos* en la prueba 1 *o* en la prueba 2	$\left(\begin{array}{c}\text{Sensibilidad}\\ \text{de la prueba 1}\\ +\\ \text{Sensibilidad}\\ \text{de la prueba 2}\end{array}\right) - \left(\begin{array}{c}\text{Sensibilidad}\\ \text{de la prueba 1}\\ \times\\ \text{Sensibilidad}\\ \text{de la prueba 2}\end{array}\right)$
	Especificidad neta	105-106	La proporción de los que *no tienen* la enfermedad y obtienen resultados *negativos* en la prueba 1 *y* en la prueba 2	(Especificidad de la prueba 1) × (Especificidad de la prueba 2)

FN, falsos negativos; *FP,* falsos positivos; *VN,* verdaderos negativos; *VP,* verdaderos positivos.

En el Apéndice 2 se resumen los tres pasos necesarios para calcular el estadístico kappa.

Apéndice 2 del capítulo 5. Los tres pasos necesarios para calcular el estadístico kappa (κ)

Componentes	Pasos
NUMERADOR ¿En qué cuantía la concordancia observada es mejor que la que cabría esperar únicamente por el azar?	PASO 1: (Porcentaje de concordancia observado) − (Porcentaje de concordancia esperado únicamente por el azar)
DENOMINADOR ¿Cuál es la máxima mejoría que podrían haber logrado los observadores sobre la concordancia esperada únicamente por el azar?	PASO 2: 100% − (Porcentaje de concordancia esperado únicamente por el azar)
$\frac{\text{NUMERADOR}}{\text{DENOMINADOR}} = \text{ESTADÍSTICO KAPPA}\,(\kappa)$ Del máximo aumento en la concordancia esperado más allá del debido únicamente al azar que podría haberse producido, ¿qué proporción se ha producido realmente?	PASO 3: $\kappa = \frac{(\text{Porcentaje de concordancia observado}) - (\text{Porcentaje de concordancia esperado únicamente por el azar})}{100\% - (\text{Porcentaje de concordancia esperado únicamente por el azar})}$

En la página 113 se expone una explicación detallada de kappa y un ejemplo de su cálculo.

Preguntas de repaso del capítulo 5

Las preguntas 1, 2 y 3 se basan en la siguiente información:

Se realizó una exploración física como cribado del cáncer de mama en 2.500 mujeres con adenocarcinoma de mama demostrado mediante biopsia y en 5.000 mujeres controles de edad y raza similares. Los resultados de la exploración fueron positivos (es decir, se palpó una masa) en 1.800 casos y en 800 de las mujeres controles, todas las cuales carecían de signos de cáncer en la biopsia.

1. *La sensibilidad de la exploración física fue*: _____

2. *La especificidad de la exploración física fue*: _____

3. *El valor predictivo positivo de la exploración física fue*: _____

La pregunta 4 se basa en la siguiente información:

Una prueba de cribado se utiliza del mismo modo en dos poblaciones similares, pero la proporción de resultados falsos positivos entre los que obtienen resultados positivos en la población A es menor que

entre los que obtienen resultados positivos en la población B.

4. *¿Cuál es la explicación probable de este hallazgo?*

a. Es imposible determinar la causa de esta diferencia.

b. La especificidad de la prueba es menor en la población A.

c. La prevalencia de la enfermedad es menor en la población A.

d. La prevalencia de la enfermedad es mayor en la población A.

e. La especificidad de la prueba es mayor en la población A.

La pregunta 5 se basa en la siguiente información:

Se realizó una exploración física y una audiometría a 500 personas en las que se sospechaban problemas auditivos; fueron encontrados en 300 de ellas. Los resultados de la exploración fueron los siguientes:

	PROBLEMAS AUDITIVOS	
Resultado	**Presentes**	**Ausentes**
Exploración física		
Positivo	240	40
Negativo	60	160
Audiometría		
Positivo	270	60
Negativo	30	140

5. *En comparación con la exploración física, la audiometría es:*

a. Igual de sensible y específica.

b. Menos sensible y menos específica.

c. Menos sensible y más específica.

d. Más sensible y menos específica.

e. Más sensible y más específica.

La pregunta 6 se basa en la siguiente información:

Dos pediatras quieren estudiar una nueva prueba de laboratorio que identifica las infecciones estreptocócicas. El Dr. Kidd utiliza la prueba de cultivo estándar, que posee una sensibilidad del 90% y una especificidad del 96%. El Dr. Childs utiliza la prueba nueva, que posee un 96% de sensibilidad y un 96% de especificidad.

6. *Si realizamos el cultivo en 200 pacientes con ambas pruebas, ¿cuál de las siguientes afirmaciones es correcta?*

a. El Dr. Kidd identificará correctamente a más personas con infección estreptocócica que el Dr. Childs.

b. El Dr. Kidd identificará correctamente a menos personas con infección estreptocócica que el Dr. Childs.

c. El Dr. Kidd identificará correctamente a más personas sin infección estreptocócica que el Dr. Childs.

d. Se necesita conocer la prevalencia de la infección estreptocócica para determinar qué pediatra identificará correctamente a un mayor número de personas con la enfermedad.

Las preguntas 7 y 8 se basan en la siguiente información:

En Nottingham, Inglaterra, se está llevando a cabo un estudio de cribado de cáncer de colon. Se estudiarán individuos de 50-75 años con la prueba Hemoccult. En esta prueba se estudia la presencia de sangre en una muestra de heces.

7. *La prueba Hemoccult posee una sensibilidad del 70% y una especificidad del 75%. Si la prevalencia del cáncer de colon en Nottingham es de 12/1.000, ¿cuál es el valor predictivo positivo de la prueba?*

8. *Si el resultado de la prueba Hemoccult es negativo, no se realizan nuevas pruebas. Si el resultado de la prueba Hemoccult es positivo, se volverá a analizar una segunda muestra de heces del individuo con la prueba Hemoccult II. Si el resultado en esta segunda muestra también es positivo, el individuo será remitido para realizar un estudio más extenso. ¿Cuál es el efecto sobre la sensibilidad neta y la especificidad neta de este método de cribado?*

a. Tanto la sensibilidad neta como la especificidad neta aumentan.

b. La sensibilidad neta se reduce y la especificidad neta aumenta.

c. La sensibilidad neta no cambia y la especificidad neta aumenta.

d. La sensibilidad neta aumenta y la especificidad neta disminuye.

e. El efecto sobre la sensibilidad neta y la especificidad neta no puede determinarse a partir de estos datos.

Las preguntas 9-12 se basan en la siguiente información:

Se pidió a dos médicos que clasificasen 100 radiografías de tórax como anormales o normales independientemente. La comparación de su clasificación se expone en la siguiente tabla:

Comparación entre la clasificación de las radiografías de tórax por el médico 1 y el médico 2

	Médico 2		
MÉDICO 1	**Anormal**	**Normal**	**Total**
Anormal	40	20	60
Normal	10	30	40
Total	50	50	100

9. *El porcentaje de concordancia simple entre los dos médicos respecto al total es*:

10. *El porcentaje de concordancia entre los dos médicos, excluyendo las radiografías clasificadas como normales por ambos médicos es*:

11. *El valor de kappa es*:

12. *Este valor de kappa, ¿qué grado de concordancia representa?*

a. Excelente.

b. Intermedio-bueno.

c. Bajo.

CAPÍTULO 6

La historia natural de la enfermedad: formas de expresar el pronóstico

Palabras clave

tasa de letalidad
personas-años
tabla de vida
análisis de supervivencia
método de Kaplan-Meier

Objetivos de aprendizaje

- Comparar cinco formas diferentes de describir la historia natural de la enfermedad: tasa de letalidad, supervivencia a 5 años, supervivencia observada, mediana de supervivencia y supervivencia relativa.
- Describir dos abordajes para calcular la supervivencia observada a lo largo del tiempo: el abordaje de la tabla de vida y el método Kaplan-Meier.
- Ilustrar el uso de tablas de vida para estudiar cambios de la supervivencia.
- Describir cómo las mejoras en los métodos diagnósticos disponibles pueden afectar a la estimación del pronóstico (migración de estadios).

Hasta ahora hemos aprendido cómo las pruebas diagnósticas y de cribado permiten la diferenciación entre individuos sanos y enfermos. Una vez que se identifica que una persona tiene una enfermedad, la pregunta que surge es: *¿Cómo podemos describir la historia natural de la enfermedad en términos cuantitativos?* Dicha cuantificación es importante por varios motivos. En

primer lugar, es necesario describir la gravedad de una enfermedad para establecer prioridades en los servicios clínicos y en los programas de salud pública. En segundo lugar, los pacientes a menudo plantean preguntas acerca del pronóstico (fig. 6.1). En tercer lugar, dicha cuantificación es importante para establecer una línea basal de la historia natural, de modo que, a medida que se disponga de nuevos tratamientos, los efectos de estos tratamientos puedan compararse con el resultado esperado sin estos. Esto también es importante para identificar diferentes tratamientos o estrategias terapéuticas para diferentes etapas de la enfermedad. Además, si se dispone de diferentes tipos de tratamientos para una cierta enfermedad, como tratamientos médicos o quirúrgicos, o dos tipos diferentes de intervenciones quirúrgicas, queremos ser capaces de comparar la eficacia de las diferentes modalidades terapéuticas. Por tanto, para poder realizar dicha comparación, necesitamos medios cuantitativos para expresar el pronóstico en grupos que reciben diferentes tratamientos.



"How much time do I have, Doc?"

FIG. 6.1 «¿Cuánto tiempo me queda, doctor?». Preocupación acerca del pronóstico. (Charles Barsotti/The New Yorker Collection/The Cartoon Bank.)

Este capítulo expone algunas de las formas de describir el pronóstico de un grupo de pacientes en términos cuantitativos. Por tanto, este capítulo estudia la historia natural de la enfermedad (y, en consecuencia, su pronóstico). En

capítulos posteriores se analiza cómo se puede intervenir en la historia natural de la enfermedad para mejorar el pronóstico: en los capítulos 10 y 11 se estudia cómo se utilizan los ensayos clínicos aleatorizados para seleccionar la intervención más apropiada (fármaco, intervención quirúrgica o estilo de vida) y en el capítulo 18 se estudia cómo, a través del proceso de cribado, puede detectarse una enfermedad en un momento más temprano de lo habitual en su historia natural para maximizar la eficacia del tratamiento. Para estudiar el pronóstico, comencemos con una representación esquemática de la historia natural de la enfermedad en un paciente, según se muestra en la figura 6.2.

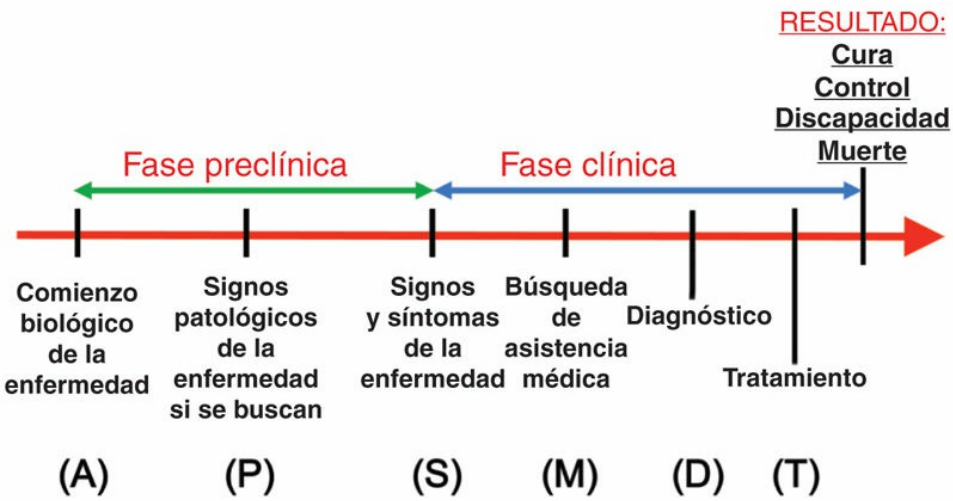


FIG. 6.2 La historia natural de la enfermedad en un paciente.

El punto A marca el comienzo biológico de la enfermedad. A menudo, este punto no puede identificarse porque se produce de manera subclínica, quizá como un cambio subcelular, como una alteración del ADN. En algún punto en la progresión del proceso de la enfermedad (punto P), podrían obtenerse pruebas patológicas de la enfermedad si estas se buscaran mediante cribado de la población o por un médico, probablemente durante un examen de rutina; estas pruebas también pueden ser un hallazgo casual descubierto al tratar otra enfermedad o patología en el mismo paciente. Posteriormente, el paciente presenta los signos y los síntomas de la enfermedad (punto S) y, algún tiempo después, el paciente puede buscar asistencia médica (punto M). A continuación, el paciente puede ser diagnosticado (punto D), tras lo que puede pautarse un tratamiento (punto T). La evolución posterior de la enfermedad podría terminar con la curación o la remisión, su control (con o sin discapacidad) o incluso la muerte.

¿En qué momento comenzamos a cuantificar el tiempo de supervivencia? De modo ideal, preferiríamos hacerlo desde el comienzo de la enfermedad. Sin embargo, por lo general, esto no es posible porque el momento del comienzo biológico en un individuo es desconocido. Si quisiéramos contar desde el momento en el que comienzan los síntomas, introduciríamos una

gran variabilidad subjetiva al medir la duración de la supervivencia porque ignoramos inadvertidamente el tiempo entre el inicio biológico de la enfermedad y los primeros síntomas y signos, que puede variar desde horas o días (para una infección aguda) hasta meses o años (p. ej., en el cáncer de próstata). Por lo general, para estandarizar los cálculos, la duración de la supervivencia se mide desde el momento del diagnóstico. Sin embargo, incluso con el uso de este punto de comienzo, sigue produciéndose variabilidad porque los pacientes difieren en el momento en el que buscan asistencia médica. Además, algunas enfermedades, como ciertos tipos de artritis, son indoloras y se desarrollan lentamente, de modo que puede que los pacientes no sean capaces de detallar con precisión el comienzo de los síntomas o recordar el punto en el tiempo en el que solicitaron asistencia médica. Además, cuando la supervivencia se cuenta desde el momento del diagnóstico, todo paciente que haya fallecido antes de ser diagnosticado es excluido del recuento. ¿Cómo afectaría esto a nuestras estimaciones sobre el pronóstico?

Una pregunta relacionada importante es: «¿Cómo se realiza el diagnóstico?». ¿Existe una prueba patognomónica clara para la enfermedad en cuestión? Con frecuencia no disponemos de dicha prueba. En ocasiones, una enfermedad puede ser diagnosticada tras el aislamiento de un microorganismo infeccioso, pero, como las personas pueden ser portadoras de microorganismos sin estar realmente infectadas, no siempre sabemos si el microorganismo aislado es la causa de la enfermedad. En algunas enfermedades preferiríamos alcanzar el diagnóstico mediante confirmación tisular por biopsia, pero con frecuencia existe variabilidad en la interpretación de las muestras de tejido por diferentes anatomopatólogos. Un problema adicional es que, en ciertos problemas de salud, como las cefaleas, las lumbalgias y la dismenorrea, puede no ser posible un diagnóstico tisular específico. Por tanto, cuando decimos que la supervivencia se mide desde el momento del diagnóstico, la franja temporal no siempre está clara. Estos aspectos deben tenerse en cuenta cuando avancemos en el análisis de los diferentes abordajes para estimar el pronóstico.

El pronóstico puede expresarse en función de las muertes debidas a la enfermedad o en función de los que sobreviven a la enfermedad. Aunque en la siguiente exposición empleamos ambos abordajes, el punto final empleado para los propósitos de nuestro análisis en este ejemplo es la muerte. Como la muerte es inevitable, no nos referimos a morir frente a no morir, sino a prolongar el intervalo hasta que se produce la muerte tras el diagnóstico. Se pueden utilizar otros puntos finales, como el intervalo desde el diagnóstico hasta la recurrencia de la enfermedad o desde el diagnóstico hasta el momento en el que aparece afectación funcional, complicaciones específicas de la enfermedad, discapacidad o cambios en la calidad de vida del paciente, todos los cuales pueden verse afectados por la invasividad de los tratamientos disponibles, el momento en el que se inició el tratamiento o por el grado de mejoría alcanzable en algunos de los síntomas, incluso aunque no

pueda aumentarse la esperanza de vida del paciente. Todas estas son medidas importantes, pero no se tratan en este capítulo.

Tasa de letalidad

La primera forma de expresar el pronóstico es la *tasa de letalidad* (se expuso en el cap. 4). La tasa de letalidad se define como el número de personas que mueren por una enfermedad dividido entre el número de personas que tienen la enfermedad. Cuando una persona tiene una enfermedad, ¿cuál es la probabilidad de que muera de dicha enfermedad? Obsérvese que el denominador de la tasa de letalidad es el número de personas que tienen la enfermedad, lo que representa una proporción, aunque en ocasiones se refiere a la misma incorrectamente como una tasa. En esto se diferencia de la *tasa de mortalidad,* en la que el denominador incluye a cualquier persona con riesgo de morir de la enfermedad: tanto personas que tienen la enfermedad como personas que (todavía) no tienen la enfermedad, pero que podrían presentarla.

La tasa de letalidad no incluye ninguna mención explícita del tiempo. Sin embargo, el tiempo es expresado implícitamente, porque la tasa de letalidad suele usarse en enfermedades agudas en las que la muerte, si se produce, ocurre relativamente pronto tras el diagnóstico. Por tanto, si se conoce la historia natural habitual de la enfermedad, el término *tasa de letalidad* se refiere al periodo tras el diagnóstico durante el que cabría esperar que el paciente falleciera.

La tasa de letalidad es apropiada para enfermedades agudas de corta duración. Para las enfermedades crónicas en las que la muerte puede producirse muchos años tras el diagnóstico y la posibilidad de morir de otras causas se vuelve más probable, la tasa de letalidad es una medida menos útil. Por ejemplo, en el estudio del cáncer de próstata, la mayoría de los hombres con este diagnóstico mueren por alguna otra causa, debido a la progresión muy lenta de este cáncer. Por tanto, usamos diferentes abordajes para expresar el pronóstico en dichas enfermedades.

Personas-años

Una forma útil de expresar la mortalidad es mediante el número de muertes dividido entre las personas-años a lo largo de los que se observa un grupo. Como los individuos a menudo son observados durante diferentes periodos de tiempo, la unidad usada para contar el tiempo de observación es personas-años. (Las personas-años se abordaron en el cap. 3, págs. 47-50.) El número de personas-años para dos personas, cada una de las cuales es observada durante 5 años, es igual al de 10 personas, cada una de las cuales es observada durante 1 año, es decir, 10 personas-años. Los números de personas-años pueden sumarse y el número de acontecimientos, como las muertes, pueden calcularse para el número de personas-años observado.

Un problema de utilizar las personas-años es que se asume que cada persona-año es equivalente al resto de personas-años (es decir, que el riesgo es el mismo en cualquier persona-año observado). Sin embargo, puede que esto no sea así. Consideremos la situación de la figura 6.3, que muestra dos ejemplos de 10 personas-años: dos personas observadas durante 5 años y cinco personas observadas durante 2 años. ¿Son equivalentes?

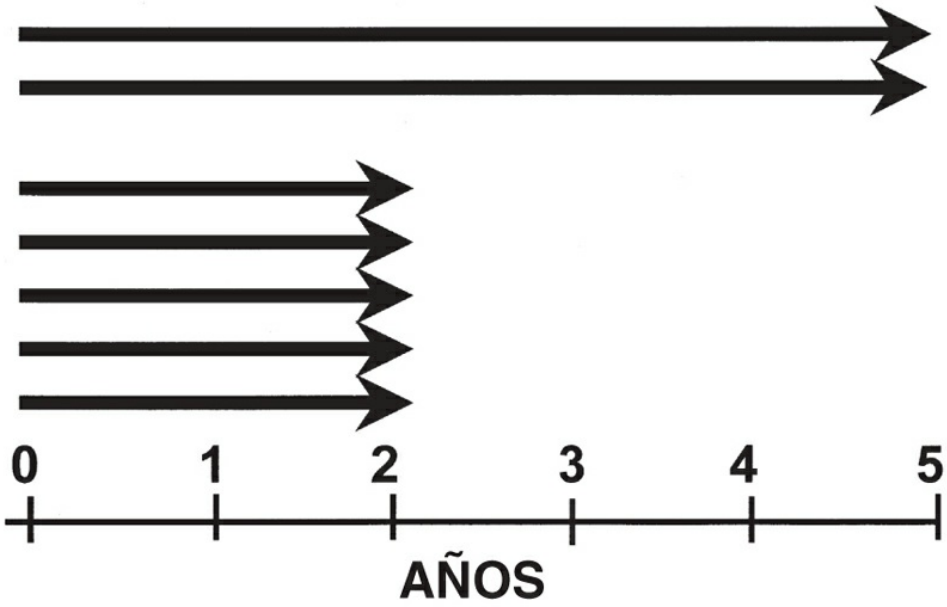


FIG. 6.3 Dos ejemplos de 10 personas-años: dos personas, cada una de ellas observada durante 5 años, y cinco personas, cada una de ellas observada durante 2 años.

Supongamos la situación que se muestra en la figura 6.4; observamos que el periodo de mayor riesgo de morir es desde poco tiempo después del diagnóstico hasta aproximadamente 20 meses después del diagnóstico. Claramente, la mayor parte de las personas-años del primer ejemplo (es decir, dos personas observadas durante 5 años) se encontrarán fuera del periodo de

mayor riesgo (fig. 6.5), de los 20 a los 60 meses. Por el contrario, la mayor parte de los intervalos de 2 años de las 5 personas mostradas en el segundo ejemplo tendrán lugar durante el periodo de mayor riesgo (fig. 6.6). Por tanto, cuando comparamos los dos ejemplos (fig. 6.7), cabría esperar más muertes en el ejemplo de las cinco personas observadas durante 2 años que en el ejemplo de las dos personas observadas durante 5 años. A pesar de este aspecto, las personas-años resultan útiles como denominadores de tasas de acontecimientos en muchas situaciones, como en ensayos clínicos aleatorizados (v. caps. 10 y 11) y en estudios de cohortes (v. cap. 8). Tenga en cuenta que, como se expone en otros libros de texto[1], una tasa por persona-año es equivalente a una tasa promedio anual. Por lo tanto, una tasa por persona-año puede compararse con una tasa anual de estadísticas vitales basada en la estimación de la población del punto medio del periodo. Esto resulta útil cuando interesa comparar las tasas por persona-año en un estudio con las tasas de población.

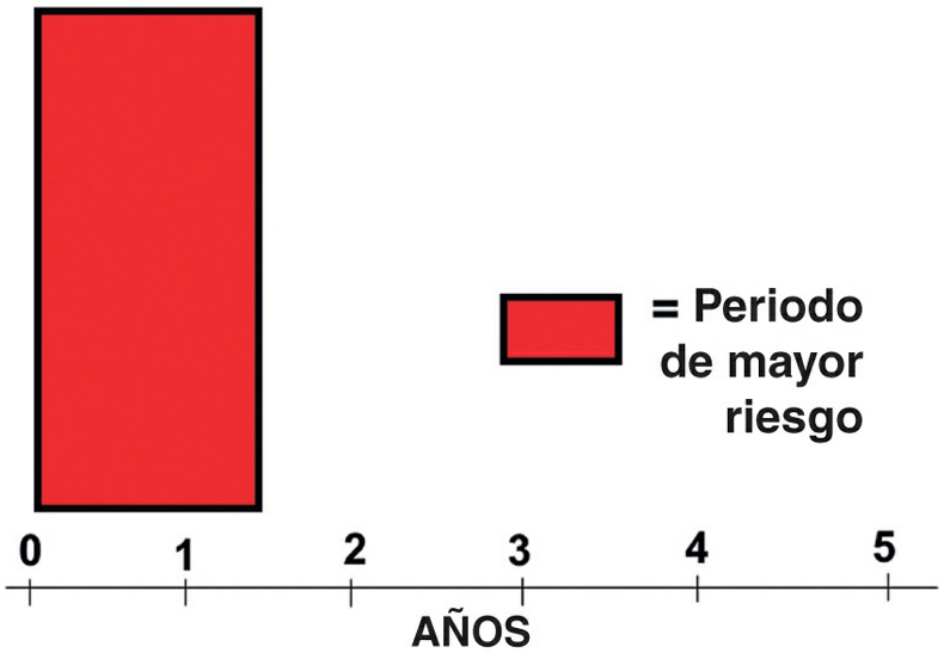


FIG. 6.4 El momento de mayor riesgo es desde poco después del diagnóstico hasta aproximadamente 20 meses después de este.

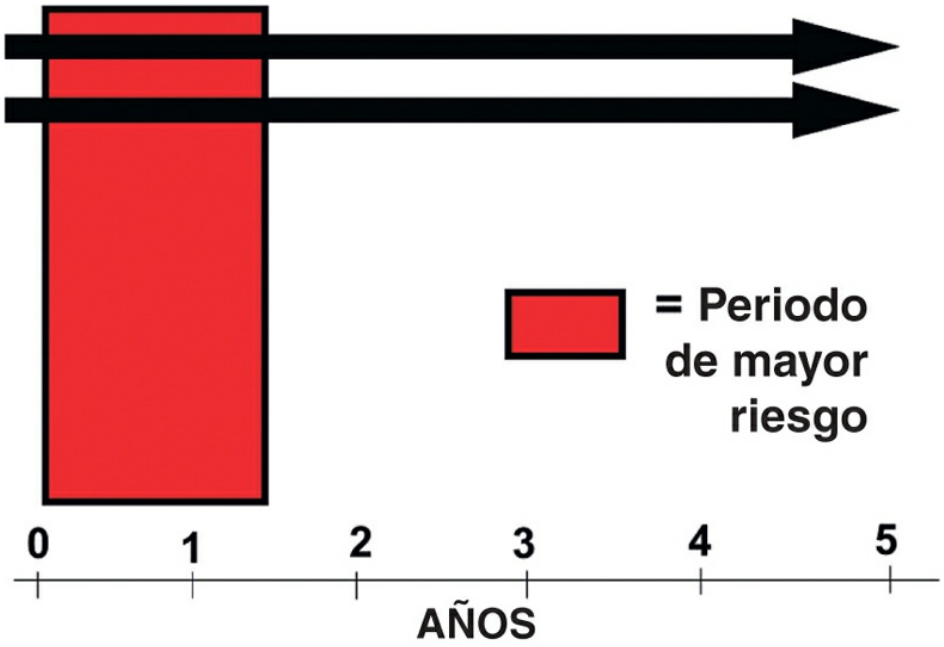


FIG. 6.5 Dos personas, cada una de ellas observada durante 5 años, y la relación con el periodo de mayor riesgo.

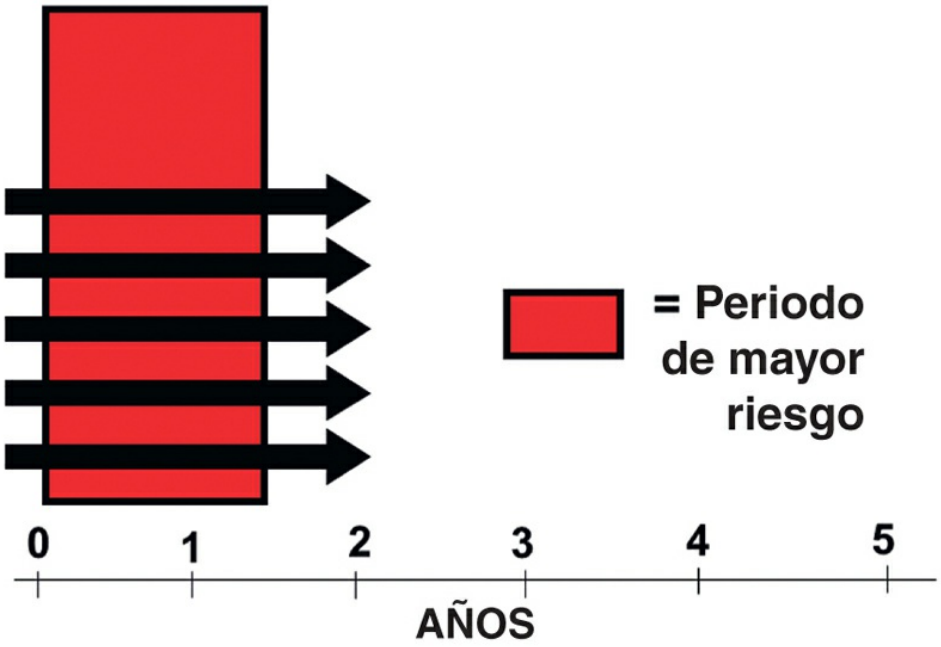


FIG. 6.6 Cinco personas, cada una de ellas observada durante 2 años, y la relación con el periodo de mayor riesgo.

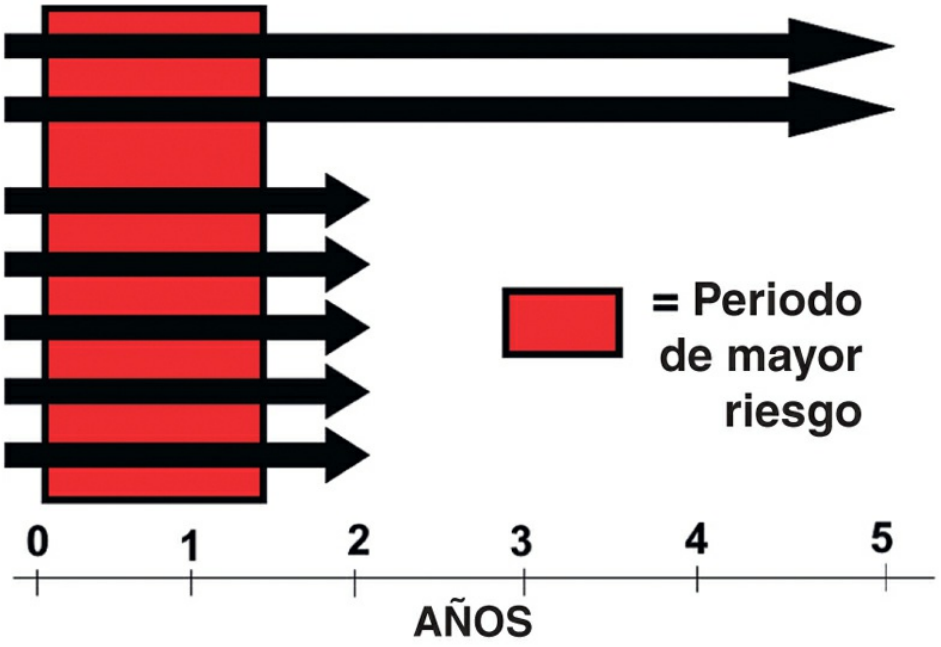


FIG. 6.7 Dos ejemplos de 10 personas-años en los que el periodo de mayor riesgo es desde poco después del diagnóstico hasta aproximadamente 20 meses tras este.

Supervivencia a cinco años

La *supervivencia a cinco años* es otra medida empleada para expresar el pronóstico. Este término se utiliza con frecuencia en la medicina clínica, especialmente para evaluar tratamientos del cáncer.

La supervivencia a 5 años es el porcentaje de pacientes que están vivos 5 años después del comienzo del tratamiento o 5 años después del diagnóstico. (Aunque a menudo se habla de la supervivencia a 5 años como una tasa, realmente es una proporción.) A pesar del uso extendido del intervalo de 5 años, se debe precisar que no tiene nada de mágico. Ciertamente, en la historia natural de una enfermedad no se produce ningún cambio biológico significativo de forma abrupta a los 5 años que justifique su uso como punto final. Sin embargo, la mayoría de las muertes por cáncer se producían por lo general durante este periodo tras el diagnóstico cuando este comenzó a utilizarse en la década de 1950, por lo que desde entonces la supervivencia a 5 años se ha utilizado como índice de éxito del tratamiento del cáncer.

Un problema con el uso de la supervivencia a 5 años se ha vuelto más importante en los últimos años con el empleo de mejores programas de cribado. Estudiemos un ejemplo hipotético: en la figura 6.8 se muestra la cronología de una mujer con cáncer de mama de comienzo biológico en el año 2005. Como la enfermedad era subclínica en esa fecha, se encontraba asintomática. En 2013 notó un bulto en la mama que la llevó a consultar a su médico, que realizó el diagnóstico. La paciente fue sometida posteriormente a una mastectomía. En 2015 falleció por un cáncer metastásico. Si utilizamos como medida la supervivencia a 5 años, que se emplea con frecuencia en oncología como medida del éxito del tratamiento, esta paciente no ha sido un «éxito» porque solo sobrevivió 2 años tras el diagnóstico.



FIG. 6.8 El problema de la supervivencia a 5 años en una población cribada: I. Situación sin cribado.

Imaginemos ahora que esta mujer vivía en una comunidad en la que existía una campaña agresiva de cribado del cáncer de mama mediante mamografías (cronología inferior en la fig. 6.9). Al igual que antes, el comienzo biológico de la enfermedad tuvo lugar en el año 2005, pero en 2010 se identificó una masa muy pequeña en su mama por medio del programa de cribado. Fue

intervenida quirúrgicamente en 2010, pero falleció en 2015. Como sobrevivió 5 años tras el diagnóstico y el tratamiento, sería identificada como un «éxito» terapéutico en términos de la supervivencia a 5 años. Sin embargo, esta supervivencia aparentemente más prolongada es un artefacto. La muerte siguió ocurriendo en 2015; la vida de la paciente no fue más prolongada tras la detección y el tratamiento más tempranos. Lo que ha ocurrido es que el intervalo entre el diagnóstico (y el tratamiento) y su muerte aumentó por el diagnóstico más precoz, pero no se retrasó la fecha de su muerte. (El intervalo entre el diagnóstico más temprano en 2010, hecho posible por el estudio de cribado, y el momento habitual de diagnóstico más tardío en 2013 se denomina *adelanto en el momento del diagnóstico*. Este concepto se aborda en detalle en el capítulo 18 en el contexto de la evaluación de los programas de cribado.) Es engañoso concluir que, teniendo en cuenta la supervivencia a 5 años de la paciente, el resultado del segundo escenario es mejor que el del primero, porque no se ha producido un cambio en la historia natural de la enfermedad, como refleja el año en el que se produjo la muerte. De hecho, el único cambio que ha tenido lugar es que, cuando se realizó el diagnóstico 3 años antes (2010 frente a 2013), la paciente recibió cuidados médicos para su cáncer de mama, con todas las dificultades acompañantes, durante 3 años adicionales. Así pues, cuando se realizan pruebas de cribado, puede observarse una supervivencia a 5 años más alta, no porque los pacientes vivan más tiempo sino únicamente porque el diagnóstico se ha realizado más precozmente. Este tipo de sesgo potencial (conocido como *sesgo por adelanto en el momento del diagnóstico*) debe tenerse en cuenta cuando se valora cualquier programa de cribado antes de poder concluir que el cribado es beneficioso para aumentar la supervivencia.

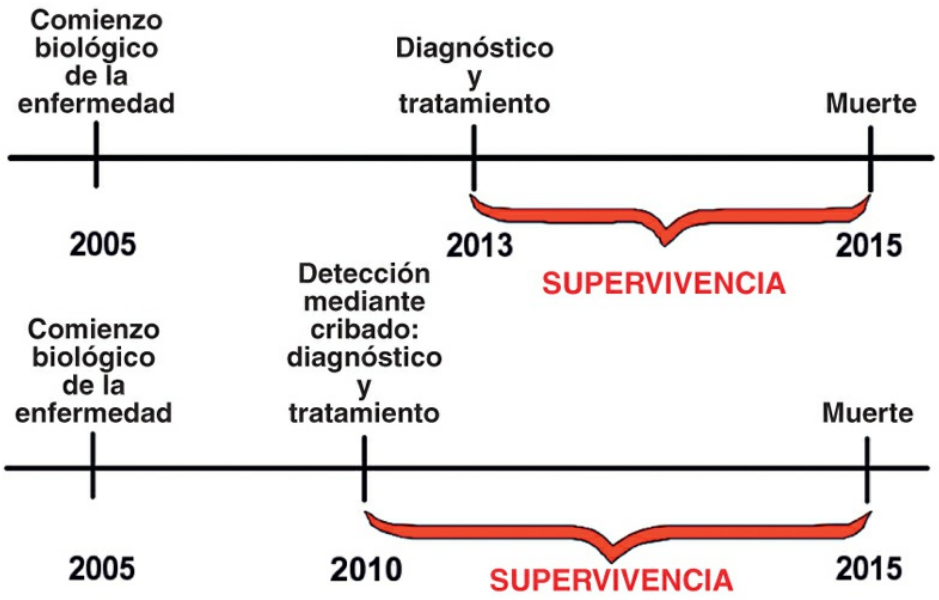


FIG. 6.9 El problema de la supervivencia a 5 años en una población cribada: II. Detección más temprana de la enfermedad gracias al cribado.

Otro problema con la supervivencia a 5 años es que, si queremos fijarnos en la experiencia de supervivencia de un grupo de pacientes que fueron diagnosticados hace menos de 5 años, claramente no podemos utilizar este criterio, porque en estos pacientes se necesitan 5 años de observación para calcular la supervivencia a 5 años. Por tanto, si queremos valorar un tratamiento que fue iniciado hace menos de 5 años, la supervivencia a 5 años no es una medida apropiada.

Un último aspecto relacionado con la supervivencia a 5 años se muestra en la figura 6.10. En esta figura observamos curvas de supervivencia de dos poblaciones, A y B. La supervivencia a 5 años es de aproximadamente el 10%. Sin embargo, las curvas que dan lugar a la misma supervivencia a 5 años son bastante diferentes. Aunque la supervivencia a 5 años sea la misma en ambos grupos, la mayoría de las muertes en el grupo A no se produjeron hasta el quinto año, mientras que la mayoría de las muertes en el grupo B se produjeron en el primer año, ya que en ellos el periodo de tiempo hasta el desenlace (muerte) fue más corto en comparación con el grupo A. Así, a pesar de supervivencias a 5 años idénticas, la supervivencia durante los 5 años es claramente mejor para los pacientes del grupo A.

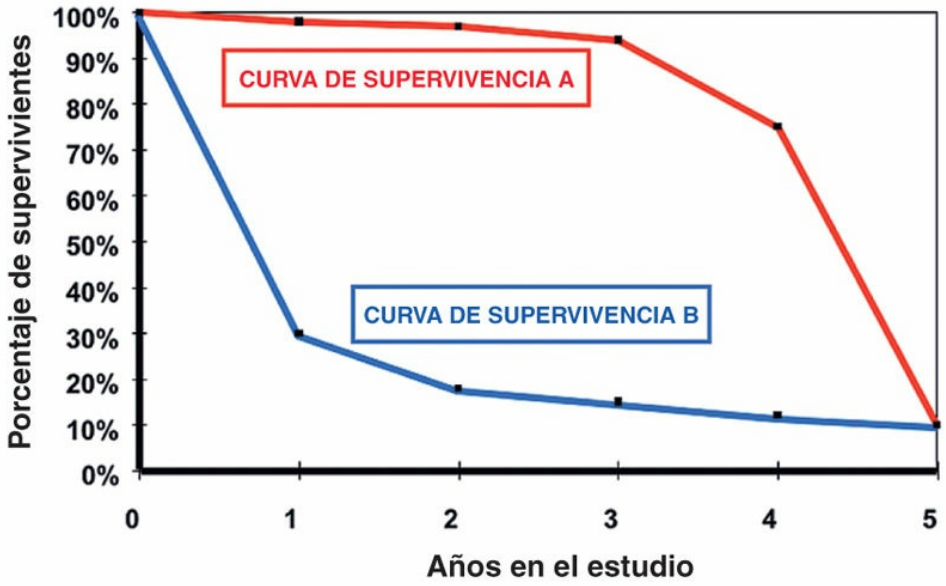


FIG. 6.10 Curvas de supervivencia a 5 años en dos poblaciones hipotéticas.

Supervivencia observada

Fundamento de la tabla de vida

Otro enfoque para cuantificar el pronóstico es utilizar la supervivencia real observada en los pacientes seguidos a lo largo del tiempo, basándose en saber el intervalo dentro del cual ocurrió el acontecimiento. Para ello, utilizamos una *tabla de vida.* Las tablas de vida han sido utilizadas por los actuarios para estimar el riesgo en poblaciones durante siglos cuando no había datos sobre individuos. Los métodos y modelos actuariales se han aplicado en un gran número de situaciones, entre las que se incluyen propiedad/accidentes, seguros de vida, pensiones y seguros de salud, entre otros. Los actuarios están acreditados, con una base de estadísticas y probabilidad, procesos estocásticos y métodos y modelos actuariales.

Examinemos el marco conceptual que subyace en el cálculo de las tasas de supervivencia usando una tabla de vida, especialmente cuando no se conoce el tiempo exacto del acontecimiento, sino que usamos el intervalo dentro del cual tuvo lugar el acontecimiento.

En la tabla 6.1 se muestra un estudio hipotético de los resultados del tratamiento en pacientes tratados de 2010 a 2014 y seguidos hasta 2015. (Simplemente mirando esta tabla, usted se dará cuenta de que el ejemplo es hipotético, porque el título indica que ¡no se perdió a ningún paciente durante el seguimiento!)

Tabla 6.1

Estudio hipotético de los resultados del tratamiento de pacientes tratados de 2010 a 2014 y seguidos hasta 2015 (ninguna pérdida de seguimiento)

		NÚMERO DE VIVOS EN EL ANIVERSARIO DEL TRATAMIENTO				
Año de tratamiento	N.° de pacientes tratados	2001	2002	2003	2004	2005
2010	84	44	21	13	10	8
2011	62		31	14	10	6
2012	93			50	20	13
2013	60				29	16
2014	76					43

Para cada año de calendario de tratamiento, la tabla muestra el número de pacientes que reciben el tratamiento y el número de pacientes vivos en cada año de calendario tras el inicio de dicho tratamiento. Por ejemplo, de los 84 pacientes que iniciaron el tratamiento en el año 2010, 44 estaban vivos en

2011, un año después de comenzar el tratamiento; 21 estaban vivos en 2012, y así sucesivamente.

Los resultados de la tabla 6.1 incluyen todos los datos disponibles para valorar el tratamiento. Si queremos describir el pronóstico en estos pacientes tratados utilizando todos los datos de la tabla, evidentemente no podemos emplear la supervivencia a 5 años, porque todo el grupo de 375 pacientes no ha sido observado durante 5 años. Podríamos calcular la supervivencia a 5 años a partir únicamente de los primeros 84 pacientes que iniciaron el tratamiento en 2010 y fueron observados hasta 2015, porque fueron los únicos observados durante 5 años. Sin embargo, esto nos obligaría a descartar el resto de los datos, lo que sería inapropiado, dado el esfuerzo y los gastos involucrados en la obtención de los datos, y también debido a la luz adicional que la experiencia de supervivencia de esos pacientes arrojaría sobre la eficacia del tratamiento. La pregunta es: ¿cómo podemos utilizar *toda* la información de la tabla 6.1 para describir la experiencia de supervivencia de los pacientes de este estudio?

Para utilizar todos los datos, reestructuramos los datos de la tabla 6.1 como se muestra en la tabla 6.2. En esta tabla, los datos muestran el número de pacientes que comenzaron el tratamiento en cada año del calendario y el número de aquellos que seguían vivos en cada aniversario del inicio del tratamiento. Los pacientes que comenzaron el tratamiento en 2014 fueron observados únicamente durante un solo año, porque el estudio finalizó en 2015.

Tabla 6.2

Reestructuración de los datos de la Tabla 6.1 mostrando la supervivencia tabulada por años desde el inicio del tratamiento (ninguna pérdida de seguimiento)

		NÚMERO DE VIVOS AL FINAL DEL AÑO				
Año de tratamiento	N.° de pacientes tratados	1.er año	2.° año	3.er año	4.° año	5.° año
2010	84	44	21	13	10	8
2011	62	31	14	10	6	
2012	93	50	20	13		
2013	60	29	16			
2014	76	43				

Con los datos en este formato, ¿cómo utilizamos la tabla? En primer lugar, preguntémonos: «¿Cuál es la probabilidad de sobrevivir 1 año tras el inicio del tratamiento?». Para responder a esta pregunta, dividimos el número total de pacientes que estaban vivos el primer año después del inicio del tratamiento (197) entre el número total de pacientes que comenzaron el tratamiento (375; tabla 6.3).

Tabla 6.3

Análisis de la supervivencia de los pacientes tratados de 2010 a 2014 y seguidos hasta 2015 (ninguna pérdida de seguimiento): I

		NÚMERO DE VIVOS AL FINAL DEL AÑO				
Año de tratamiento	**N.° de pacientes tratados**	**1.er año**	**2.° año**	**3.er año**	**4.° año**	**5.° año**
2010	84	44	21	13	10	8
2011	62	31	14	10	6	
2012	93	50	20	13		
2013	60	29	16			
2014	76	43				
Totales	375	197				

$$P_1 = \text{Probabilidad de sobrevivir el } 1.^{er} \text{ año} = \frac{197}{375} = 0{,}525$$

La probabilidad de sobrevivir el primer año (P_1) es:

$$P_1 = \frac{197}{375} = 0{,}525$$

A continuación, nos preguntamos: «¿Cuál es la probabilidad de que, tras sobrevivir el primer año tras iniciar el tratamiento, los pacientes sobrevivan el segundo año?». En la tabla 6.4 observamos que 197 personas sobrevivieron el primer año, pero de 43 de ellos (los que iniciaron el tratamiento en 2014) no tenemos más información porque fueron observados durante solo 1 año. Como 71 sobrevivieron el segundo año, calculamos la probabilidad de sobrevivir el segundo año si el paciente sobrevivió el primer año (P_2) del siguiente modo:

$$P_2 = \frac{71}{197 - 43} = 0,461$$

Tabla 6.4

Análisis de la supervivencia de los pacientes tratados de 2010 a 2014 y seguidos hasta 2015 (ninguna pérdida de seguimiento): II

Año de tratamiento	N.º de pacientes tratados	NÚMERO DE VIVOS AL FINAL DEL AÑO				
		1.er año	2.º año	3.er año	4.º año	4.º año
2010	84	44	21	13	10	8
2011	62	31	14	10	6	
2012	93	50	20	13		
2013	60	29	16			
2014	76	43				
Totales		197	71			

P_2 = Probabilidad de sobervivir el 2.º año $= \frac{71}{197 - 43} = 0{,}461$

En el denominador restamos los 43 pacientes de los que no tenemos datos durante el segundo año.

Siguiendo este patrón, nos preguntamos: «Dado que una persona ha sobrevivido hasta el final del segundo año, ¿cuál es la probabilidad media de que sobreviva hasta el final del tercer año?».

En la tabla 6.5 observamos que 36 sobrevivieron el tercer año. Aunque 71 habían sobrevivido el segundo año, no disponemos de más información sobre la supervivencia de 16 de ellos porque fueron incorporados tarde al estudio. Por tanto, restamos 16 a 71 y calculamos la probabilidad de sobrevivir el tercer año, teniendo en cuenta la supervivencia al final del segundo año (P_3), del siguiente modo:

$$P_3 = \frac{36}{71-16} = 0{,}655$$

Tabla 6.5

Análisis de la supervivencia de los pacientes tratados de 2010 a 2014 y seguidos hasta 2015 (ninguna pérdida de seguimiento): III

Año de tratamiento	N.º de pacientes tratados	NÚMERO DE VIVOS AL FINAL DEL AÑO				
		1.er año	2.º año	3.er año	4.º año	5.º año
2010	84	44	21	13	10	8
2011	62	31	14	10	6	
2012	93	50	20	13		
2013	60	29	16			
2014	76	43				
Totales			71	36		

P_3 = Probabilidad de sobrevivir el 3.er año $= \frac{36}{71 - 16} = 0{,}655$

Seguidamente nos preguntamos: «Si una persona sobrevive hasta el final del tercer año, ¿cuál es la probabilidad de que sobreviva hasta el final del cuarto año?».

Como se observa en la tabla 6.6, un total de 36 personas sobrevivieron el tercer año, pero carecemos de información para 13 de ellos. Como 16 sobrevivieron el cuarto año, la probabilidad de sobrevivir el cuarto año, si la persona había sobrevivido el tercer año (P_4), es:

$$P_4 = \frac{16}{36-13} = 0,696$$

Tabla 6.6

Análisis de la supervivencia de los pacientes tratados de 2010 a 2014 y seguidos hasta 2015 (ninguna pérdida de seguimiento): IV

		NÚMERO DE VIVOS AL FINAL DEL AÑO				
Año de tratamiento	N.º de pacientes tratados	1.er año	2.º año	3.er año	4.º año	5.º año
2010	84	44	21	13	10	8
2011	62	31	14	10	6	
2012	93	50	20	13		
2013	60	29	16			
2014	76	43				
Totales				36	16	

P_4 = Probabilidad de sobrevivir el 4.º año $= \frac{16}{36-13} = 0,696$

Por último, realizamos la misma operación para el quinto año (tabla 6.7). Observamos que 16 personas sobrevivieron el cuarto año, pero carecemos de más información para 6 de ellos.

Tabla 6.7

Análisis de la supervivencia de los pacientes tratados de 2010 a 2014 y seguidos hasta 2015 (ninguna pérdida de seguimiento): V

		NÚMERO DE VIVOS AL FINAL DEL AÑO				
Año de tratamiento	N.º de pacientes tratados	1.er año	2.º año	3.er año	4.º año	5.º año
2010	84	44	21	13	10	8
2011	62	31	14	10	6	
2012	93	50	20	13		
2013	60	29	16			
2014	76	43				
Totales					16	8

P_5 = Probabilidad de sobrevivir el 5.º año $= \frac{8}{16-6} = 0,800$

Como 8 personas estaban vivas al final del quinto año, la probabilidad de sobrevivir el quinto año, cuando se ha sobrevivido el cuarto año (P_5), es:

$$P_5 = \frac{8}{16-6} = 0,800$$

Utilizando todos los datos que hemos calculado, nos preguntamos: «¿Cuál es la probabilidad de sobrevivir los 5 años?». En el cuadro 6.1 se muestran todas las probabilidades que hemos calculado de sobrevivir cada año individual.

Cuadro 6.1 Probabilidad de supervivencia en cada año del estudio

$$P_1 = \text{Probabilidad de sobrevivir el 1.}^{er}\text{ año} = \frac{197}{375} = 0{,}525 = 52{,}5\%$$

$$P_2 = \text{Probabilidad de sobrevivir el 2.º año dada la supervivencia al final del 1.}^{er}\text{ año} = \frac{71}{197 - 43} = 0{,}461 = 46{,}1\%$$

$$P_3 = \text{Probabilidad de sobrevivir el 3.}^{er}\text{ año dada la supervivencia al final del 2.º año} = \frac{36}{71 - 16} = 0{,}655 = 65{,}5\%$$

$$P_4 = \text{Probabilidad de sobrevivir el 4.º año dada la supervivencia al final del 3.}^{er}\text{ año} = \frac{16}{36 - 13} = 0{,}696 = 69{,}6\%$$

$$P_5 = \text{Probabilidad de sobrevivir el 5.º año dada la supervivencia al final del 4.º año} = \frac{8}{16 - 6} = 0{,}800 = 80{,}0\%$$

Ahora podemos responder a esta pregunta: «Si una persona es incorporada al estudio, ¿cuál es la probabilidad de que sobreviva 5 años tras iniciar el tratamiento?». La probabilidad de sobrevivir 5 años es el producto de las probabilidades de sobrevivir cada año, mostradas en el cuadro 6.1. Por tanto, la probabilidad de sobrevivir 5 años es:

$$\begin{aligned} &= P_1 \times P_2 \times P_3 \times P_4 \times P_5 \\ &= 0{,}525 \times 0{,}461 \times 0{,}655 \times 0{,}696 \times 0{,}800 \\ &= 0{,}088 \, o \, 8{,}8\% \end{aligned}$$

Las probabilidades de sobrevivir diferentes periodos de tiempo se muestran en el cuadro 6.2. Estos cálculos pueden presentarse gráficamente en una curva de supervivencia, como se observa en la figura 6.11. Obsérvese que estos cálculos utilizan todos los datos que hemos obtenido, incluidos los datos de los pacientes que no fueron observados durante los 5 años del estudio. Como resultado, el uso de los datos es económico y eficiente.

Cuadro 6.2 Probabilidades acumuladas de sobrevivir diferentes periodos de tiempo

Probabilidad de sobrevivir 1 año = $P_1 = 0{,}525 = 52{,}5\%$.
Probabilidad de sobrevivir 2 años = $P_1 \times P_2 = 0{,}525 \times 0{,}461 = 0{,}242 = 24{,}2\%$.
Probabilidad de sobrevivir 3 años = $P_1 \times P_2 \times P_3 = 0{,}525 \times 0{,}461 \times 0{,}655 = 0{,}159 = 15{,}9\%$.
Probabilidad de sobrevivir 4 años = $P_1 \times P_2 \times P_3 \times P_4 = 0{,}525 \times 0{,}461 \times 0{,}655 \times 0{,}696 = 0{,}110 = 11{,}0\%$.
Probabilidad de sobrevivir 5 años = $P_1 \times P_2 \times P_3 \times P_3 \times P_5 = 0{,}525 \times 0{,}461 \times 0{,}655 \times 0{,}696 \times 0{,}800 = 0{,}088 = 8{,}8\%$.

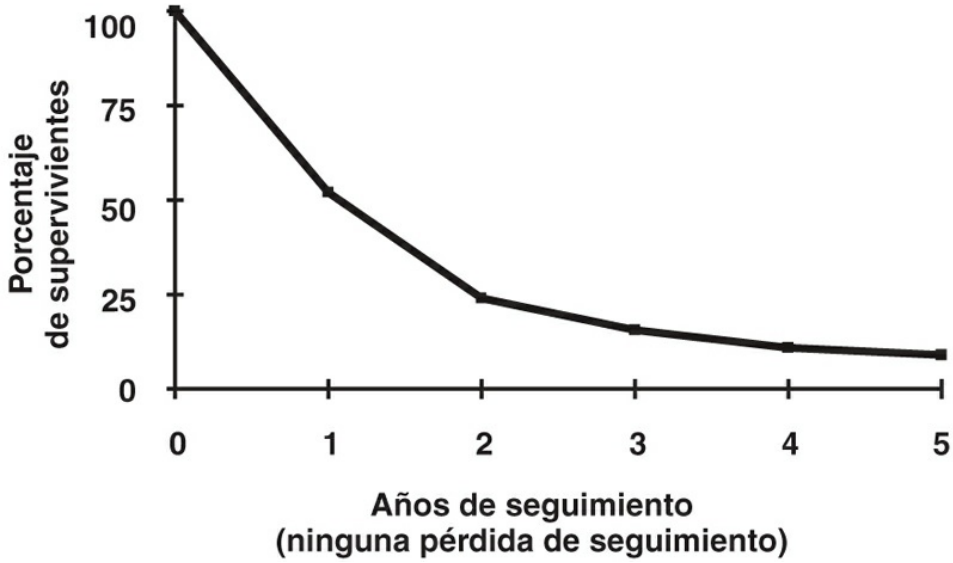


FIG. 6.11 Curva de supervivencia para un ejemplo hipotético de pacientes tratados de 2010 a 2014 y seguidos hasta 2015.

Cálculo de una tabla de vida

Fijémonos ahora en los datos de este ejemplo en la forma de tabla estándar en la que suelen presentarse para calcular una tabla de vida. En el ejemplo que acabamos de analizar, las personas de las que no se disponían datos para los 5 años del estudio fueron las que se incorporaron tiempo después de que el estudio hubiese comenzado, por lo que no fueron seguidas durante el periodo total de 5 años. En prácticamente todos los estudios de supervivencia, sin embargo, también se pierden individuos durante el periodo de seguimiento. Puede ocurrir que se pierdan o que declinen seguir participando en el estudio. Para calcular la tabla de vida, las personas de las que carecemos de datos durante el periodo completo de seguimiento (bien

porque el seguimiento no fue posible o porque se incorporaron al estudio una vez que este ya había comenzado) se denominan *pérdidas* (o *perdidos durante el seguimiento*).

En la tabla 6.8 se muestran los datos de este ejemplo con información sobre el número de muertes y pérdidas en cada intervalo. Las columnas se numeran únicamente para tener una referencia (es decir, no hay ningún significado inherente a la numeración). En la fila directamente inferior a los números de las columnas se muestran los términos empleados con frecuencia en los cálculos de las tablas de vida. Las cinco filas siguientes de la tabla proporcionan los datos de los 5 años del estudio.

Tabla 6.8

Reestructuración de datos en formato estándar para calcular una tabla de vida

(1) Intervalo desde el comienzo del tratamiento	(2) Vivos al comienzo del intervalo	(3) Muertos durante el intervalo	(4) Perdidos durante el intervalo
x	l_x	d_x	w_x
1.er año	375	178	0
2.° año	197	83	43
3.er año	71	19	16
4.° año	36	7	13
5.° año	16	2	6

Las columnas son las siguientes:

Columna (1): el intervalo desde el comienzo del tratamiento.
Columna (2): el número de individuos del estudio que estaban vivos al comienzo de cada intervalo.
Columna (3): el número de individuos del estudio que murieron durante dicho intervalo.
Columna (4): el número que se «perdió» durante el intervalo, es decir, el número de individuos del estudio que no fueron seguidos durante todo el periodo del estudio, porque se perdieron durante el seguimiento o porque se incorporaron al estudio una vez que el mismo ya había comenzado.

La tabla 6.9 incorpora columnas adicionales a la tabla 6.8. Estas columnas muestran los cálculos y son las siguientes:

Columna (5): el número de personas que tienen efectivamente riesgo de morir durante el intervalo. Se supone que las pérdidas de seguimiento (perdidos) durante cada intervalo de tiempo han ocurrido

uniformemente durante todo el intervalo. (Esta suposición es más probable que se cumpla cuando el intervalo es corto.) Por tanto, asumimos que tenían riesgo durante la mitad del intervalo. Así, para calcular el número de personas con riesgo durante cada intervalo restamos la mitad de los perdidos durante dicho intervalo, como se indica en el encabezado de la columna 5.

Columna (6): la proporción que murió durante el intervalo, que se calcula dividiendo:

$$\frac{\text{El número que falleció durante el intervalo (columna 3)}}{\text{El número que efectivamente tenía riesgo de morir durante el intervalo (columna 5)}}$$

Columna (7): la proporción que no murió durante el intervalo, es decir, la proporción de los que estaban vivos al inicio del intervalo y que sobrevivieron dicho intervalo = 1,0 − proporción que murió durante el intervalo (columna 6).

Columna (8): la proporción que sobrevivió desde el punto en el que se incorporaron al estudio hasta el final de este intervalo (supervivencia acumulada). Se obtiene multiplicando la proporción de los que estaban vivos al inicio de este intervalo y los que sobrevivieron a este intervalo por la proporción que había sobrevivido desde la incorporación hasta el final del intervalo previo. Así, cada una de las cifras de la columna 8 informa de la proporción de personas que iniciaron el estudio que sobrevivió hasta el final de cada intervalo. Esto se demostrará calculando las dos primeras filas de la tabla 6.9.

Tabla 6.9

Cálculo de una tabla de vida

(1) Intervalo desde el comienzo del tratamiento	(2) Vivos al comienzo del intervalo	(3) Muertos durante el intervalo	(4) Perdidos durante el intervalo	(5) Número efectivo de expuestos al riesgo de morir durante el intervalo: Col (2) − ½ [Col (4)]	(6) Proporción que falleció durante el intervalo: Col (3)/ Col (5)	(7) Proporción que no falleció durante el intervalo: 1 − Col (6)	(8) Proporción acumulada que sobrevivió desde la incorporación al final del intervalo: supervivencia acumulada
x	l_x	d_x	w_x	l'_x	q_x	p_x	P_x
1.er año	375	178	0	375,0	0,475	0,525	0,525
2.° año	197	83	43	175,5	0,473	0,527	0,277
3.er año	71	19	16	63,0	0,302	0,698	0,193
4.° año	36	7	13	29,5	0,237	0,763	0,147
5.° año	16	2	6	13,0	0,154	0,846	0,124

Fijémonos en los datos del primer año. (En estos cálculos, redondearemos los resultados en cada paso y utilizaremos las cifras redondeadas para el próximo cálculo. En realidad, sin embargo, cuando se calculan las tablas de vida, se utilizan las cifras no redondeadas para calcular cada intervalo posterior y, al final de todos los cálculos, todas las cifras se redondean con el fin de presentar los resultados.) Había 375 individuos incorporados al estudio que estaban vivos al comienzo del primer año tras su incorporación (columna 2). De estos, 178 murieron durante el primer año (columna 3). Todos los individuos fueron seguidos durante el primer año, por lo que no hubo pérdidas (columna 4). Por tanto, 375 personas tenían efectivamente riesgo de morir durante este intervalo (columna 5). La proporción que murió durante este intervalo fue 0,475: 178 (el número que murió [columna 3]) dividido entre 375 (el número que tenía riesgo de morir [columna 5]). La proporción que no falleció durante el intervalo es 1,0 − [la proporción que falleció (1,0 − 0,475)] = 0,525 (columna 7). Para el primer año tras la incorporación, esta también es la proporción que sobrevivió desde la incorporación hasta el final del intervalo (columna 8).

A continuación, fijémonos en los datos del segundo año. Es importante que comprendamos estos cálculos, ya que sirven de modelo para calcular cada año sucesivo en la tabla de vida.

Para calcular el número de individuos vivos al comienzo del segundo año, comenzamos con el número de vivos al comienzo del *primer* año y restamos a ese número la cifra de muertos y perdidos durante dicho año. Por tanto, al comienzo del segundo año, 197 individuos estaban vivos al comienzo del intervalo (columna 2 [375 − 178 − 0]). De estos, 83 murieron durante el segundo año (columna 3). Se produjeron 43 pérdidas de individuos que habían sido observados durante solo 1 año (columna 4). Como se ha expuesto anteriormente, restamos la mitad de las pérdidas, 21,5 (43/2), a los 197 que estaban vivos al inicio del intervalo; el resultado son 175,5 personas que tenían efectivamente riesgo de morir durante este intervalo (columna 5). La proporción que murió durante este intervalo (columna 6) fue 0,473, es decir,

83 (el número que murió [columna 3]) dividido entre 175,5 (el número con riesgo de morir [columna 5]). La proporción que no murió durante el intervalo es 1,0 − la proporción que murió (1,0 − 0,473) = 0,527 (columna 7). La proporción de individuos que sobrevivieron desde el comienzo del tratamiento hasta el final del segundo año es el producto de 0,525 (la proporción de los que habían sobrevivido desde el comienzo del tratamiento hasta el final del primer año, es decir, el comienzo del segundo año) por 0,527 (la proporción de personas que estaban vivas al comienzo del segundo año y sobrevivieron hasta el final del segundo año) = 0,277 (columna 8). Por tanto, un 27,7% de los individuos sobrevivieron desde el comienzo del tratamiento hasta el final del segundo año. Fijándonos en la última entrada de la columna 8, observamos que el 12,4% de todos los sujetos que iniciaron el estudio sobrevivieron hasta el final del quinto año.

Analice los años restantes de la tabla 6.9 para asegurarse de que entiende los conceptos y los cálculos.

El método Kaplan-Meier

A diferencia del abordaje con tablas de vida que acabamos de exponer, en el método de Kaplan-Meier[2] no se utilizan intervalos predeterminados (como 1 mes o 1 año). Con este método identificamos el punto exacto en el tiempo en el que se produjo cada muerte, o el acontecimiento de interés, de modo que cada muerte o acontecimiento termina el intervalo previo y comienza un nuevo intervalo (y una nueva fila en la tabla de Kaplan-Meier). El número de personas que murieron en dicho punto se utiliza como numerador y el número de vivos hasta ese punto (incluidos los que murieron en ese punto en el tiempo) se emplea como denominador, después de restar los perdidos producidos antes de ese punto.

Fijémonos en el pequeño estudio hipotético que se muestra en la figura 6.12. Seis pacientes fueron estudiados, de los que cuatro murieron y dos fueron perdidos durante el seguimiento («perdidos»). Las muertes se produjeron 4, 10, 14 y 24 meses después de la incorporación en el estudio. Los datos se organizan como se muestra en la tabla 6.10:

Columna (1): los tiempos hasta las muertes desde el momento de la incorporación (tiempo en el que se inició el tratamiento).
Columna (2): el número de pacientes que estaban vivos y eran seguidos en el momento de esa muerte, incluidos los que murieron es ese tiempo.
Columna (3): el número de muertos en ese tiempo.
Columna (4): la proporción entre los que estaban vivos y eran seguidos (columna 2) y los que murieron en ese tiempo (columna 3) (columna 3 / columna 2).
Columna (5): la proporción de los que estaban vivos y sobrevivieron (1,0 − columna 4).
Columna (6): supervivencia acumulada (la proporción de los que participaron desde el inicio y sobrevivieron hasta ese punto).

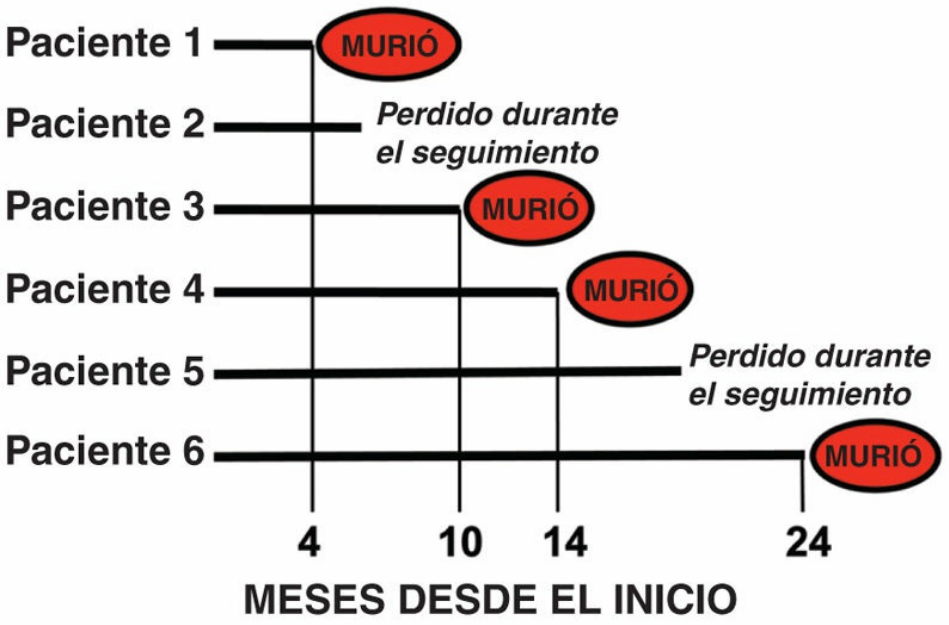


FIG. 6.12 Ejemplo hipotético de un estudio de seis pacientes analizados con el método Kaplan-Meier.

Tabla 6.10

Cálculo de la supervivencia empleando el método Kaplan-Meier[a]

(1) Tiempos hasta las muertes desde el inicio del tratamiento (meses)	(2) Números de vivos en cada momento	(3) Número de fallecidos en cada momento	(4) Proporción de fallecidos en ese momento: Col (3)/Col (2)	(5) Proporción que sobrevivió en ese momento: 1 − Col (4)	(6) Proporción acumulada que sobrevivió hasta ese momento: supervivencia acumulada
4	6	1	0,167	0,833	0,833
10	4	1	0,250	0,750	0,625
14	3	1	0,333	0,667	0,417
24	1	1	1,000	0,000	0,000

[a] Véase el texto y la figura 6.12 en relación con las pérdidas.

Fijémonos en la primera fila de la tabla. La primera muerte se produjo a los 4 meses, cuando 6 pacientes estaban vivos y eran seguidos (v. fig. 6.12). En ese punto se produjo una muerte (columna 3), para una proporción de 1/6 = 0,167 (columna 4). La proporción de los que sobrevivieron en ese momento es de 1,0 − columna 4, o 1,0 − 0,167 = 0,833 (columna 5), que también es la supervivencia acumulada en ese punto (columna 6).

La siguiente muerte tuvo lugar 10 meses después de la incorporación inicial de los 6 pacientes en el estudio, y los datos para este tiempo se observan en la siguiente fila de la tabla. Aunque antes de esta muerte solo se había producido otro fallecimiento, el número de vivos y seguidos es de solo 4 porque también se había producido una pérdida antes de este punto (no se

muestra en la tabla pero puede verse en la fig. 6.12). Por tanto, se produjo una muerte (columna 3) y, como se observa en la tabla 6.10, la proporción que murió es ¼ o 0,250 (columna 4). La proporción que sobrevivió es 1,0 − columna 4, o 1,0 − 0,250 = 0,750 (columna 5). Por último, la proporción acumulada de supervivientes (columna 6) es el producto de la proporción que sobrevivió hasta el final del intervalo previo (hasta justo antes de la muerte previa), mostrada en la columna 6 de la primera fila (0,833), por la proporción que sobrevivió desde ese momento hasta justo antes de la segunda muerte (segunda fila en la columna 5: 0,750). El producto es 0,625, es decir, un 62,5% de los que iniciaron el estudio sobrevivieron hasta este punto. Revise las siguientes dos filas de la tabla para asegurarse de que ha entendido los conceptos y los cálculos.

Los valores calculados en la columna 6 se representan como se observa en la figura 6.13. Obsérvese que los datos se representan escalonadamente en vez de en una pendiente suave, ya que, tras la disminución de la supervivencia resultante de cada muerte, la supervivencia permanece sin cambios hasta que tiene lugar el siguiente fallecimiento.

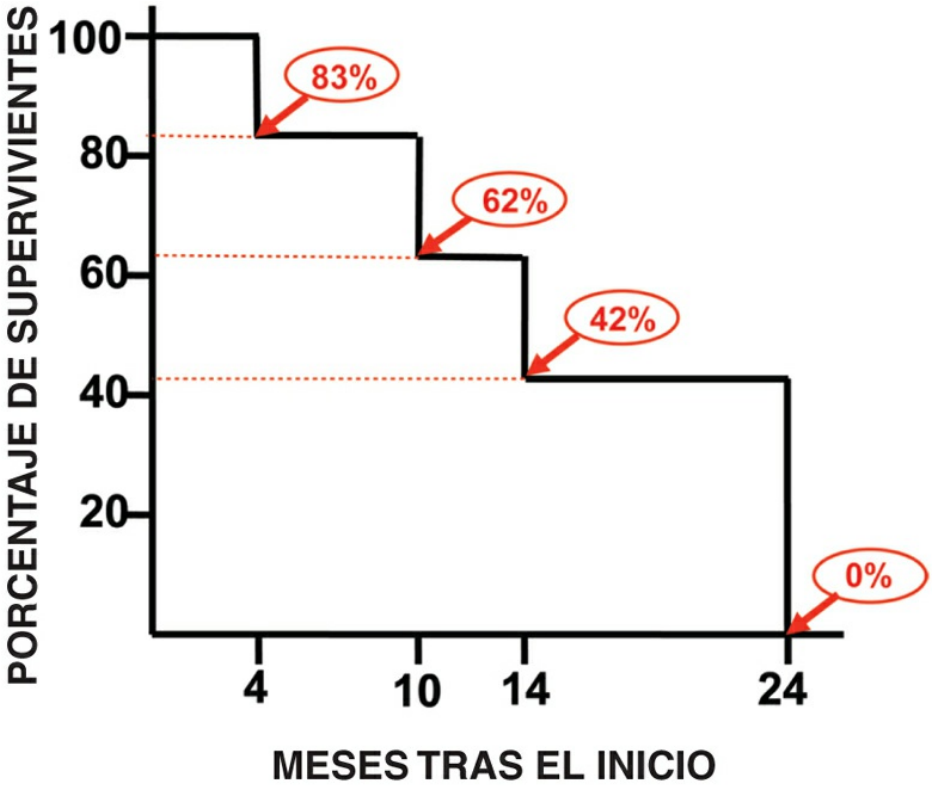


FIG. 6.13 Gráfico de Kaplan-Meier del estudio de supervivencia hipotético de seis pacientes mostrados en la figura 6.12. Los porcentajes en rojo indican las proporciones acumuladas de supervivientes tras las muertes mostradas en la figura 6.12 y se han tomado de la columna 6 de la tabla 6.10. (V. explicación del método de Kaplan-Meier en la pág. 134.)

Cuando se dispone de información acerca del momento exacto de la muerte, el método de Kaplan-Meier claramente hace pleno uso de ella,

porque los datos se usan para definir los intervalos, en lugar de intervalos arbitrarios predeterminados utilizados en el método de tablas de vida. El uso de tecnología moderna para comunicarse con los pacientes, llevado a cabo simultáneamente en diferentes sitios de estudio, y vincular electrónicamente los datos de mortalidad a las bases de datos de investigación permite a los investigadores identificar el examen del tiempo del acontecimiento. Además, existen programas computarizados fácilmente disponibles que hacen que el método de Kaplan-Meier sea aplicable también a grupos de datos extensos. La mayoría de los estudios longitudinales de los trabajos publicados comunican en la actualidad datos de supervivencia empleando el método de Kaplan-Meier. Por ejemplo, en el año 2000 Rosenhek y cols. publicaron un estudio de pacientes con estenosis aórtica grave, pero asintomática[3]. Un aspecto no resuelto era si los pacientes con enfermedad asintomática debían ser sometidos a un recambio valvular aórtico. Los investigadores examinaron la historia natural de esta enfermedad para valorar la supervivencia global de estos pacientes e identificar factores predictivos del resultado. Gibson y cols.[4] estudiaron la asociación entre el calcio arterial coronario (CAC) y los eventos cerebrovasculares (ECV) en 6.779 participantes en el Multi-Ethnic Study of Atherosclerosis (MESA) y luego los siguieron durante un promedio de 9,5 años. La figura 6.14A muestra el análisis de Kaplan-Meier de supervivencia libre de ECV por la presencia o ausencia de CAC al inicio del estudio. Los participantes con CAC presente durante el examen de referencia tuvieron una tasa de supervivencia sin ECV más baja en comparación con los participantes sin CAC en la visita de referencia. En la figura 6.14B, los autores dividieron a los participantes en cuatro grupos según su CAC en la visita de referencia (CAC: 0, 0 a 100, > 100 a 400 y > 400 unidades Agatston), y podemos ver claramente una curva distinta para cada grupo que muestra una supervivencia significativa sin ECV graduada.

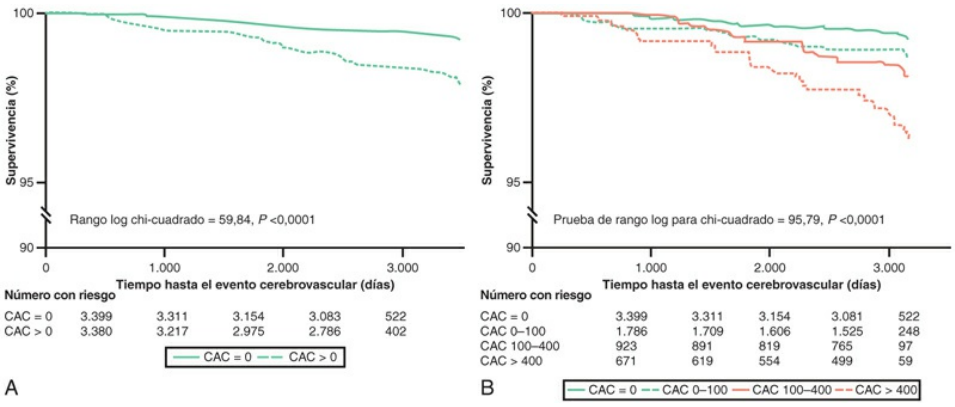


FIG. 6.14 (A) Análisis de Kaplan-Meier que muestra la supervivencia sin eventos de los participantes con y sin calcio arterial coronario *(CAC)* y eventos cerebrovasculares *(ECV)* incidentes en la cohorte MESA (Multi-Ethnic Study of Atherosclerosis). (B) Análisis de Kaplan-Meier que muestra la supervivencia sin eventos ECV de los participantes con 0, de 0 a 100, >100 a 400 y >400 CAC (unidades Agatston) y ECV incidentes en la cohorte MESA. (De Gibson AO, Blaha MJ, Arnan MK, et al. Coronary artery calcium and incident cerebrovascular events

in an asymptomatic cohort. The MESA Study. *JACC Cardiovasc Imaging.* 2014;7:1108–1115.)

Suposiciones empleadas al usar tablas de vida y el método de Kaplan-Meier

Cuando se utilizan tablas de vida y el método de Kaplan-Meier estamos suponiendo dos aspectos importantes. En primer lugar, suponemos que no se han producido cambios seculares (temporales) en la eficacia del tratamiento o en la supervivencia a lo largo del tiempo de calendario. Es decir, asumimos que durante el periodo del estudio no se han producido mejorías en el tratamiento y que la supervivencia en un año de calendario del estudio es la misma que en otro año de calendario del estudio. Claramente, si el estudio se realiza a lo largo de muchos años, esta suposición puede no ser válida, porque afortunadamente los tratamientos mejoran con el paso del tiempo. Si creemos que la eficacia del tratamiento puede haber cambiado durante el periodo del estudio, podríamos examinar los datos iniciales separadamente de los datos más tardíos. Si encontramos diferencias, podríamos analizar separadamente los periodos iniciales y tardíos y comparar los efectos.

La segunda suposición se relaciona con el seguimiento de las personas incorporadas al estudio. En prácticamente todos los estudios reales se pierde el seguimiento de algún paciente. Esto puede ocurrir por diversos motivos. Algunos pueden morir y no pueden ser seguidos. Algunos pueden cambiar de residencia o buscar asistencia médica en otro centro. Algunos pueden perderse porque la enfermedad desaparece y se encuentran bien. En la mayoría de los estudios desconocemos los motivos reales de las pérdidas de seguimiento. ¿Cómo podemos abordar el problema de las personas que perdemos durante el seguimiento y de las cuales no tenemos, por tanto, más información sobre su supervivencia? Como disponemos de datos basales de estas personas, podríamos comparar las características de las personas perdidas durante el seguimiento con las de las personas que continuaron en el estudio. Si se pierde el seguimiento de una gran proporción de la población del estudio, los hallazgos del estudio serán menos válidos. El reto es minimizar las pérdidas de seguimiento. En cualquier caso, la segunda suposición asumida en el análisis mediante tablas de vida es que la experiencia de supervivencia de las personas de las que se perdió el seguimiento es la misma que la experiencia de los que continuaron el seguimiento. Aunque esta suposición se asume con el fin de realizar los cálculos, la realidad es que su validez a menudo puede ser cuestionable. Sin embargo, para la mortalidad, la asunción se puede verificar mediante la consulta del United States National Death Index, que permite comparar la mortalidad de los perdidos durante el seguimiento con los que continúan siendo estudiados.

Aunque el término *tabla de vida* puede sugerir que estos métodos son útiles únicamente para calcular la supervivencia, en realidad no es así. La muerte no tiene por qué ser el punto final de estos cálculos. Por ejemplo, la

supervivencia puede calcularse como el tiempo que transcurre hasta la aparición de hipertensión, de una recurrencia de un cáncer o el tiempo de supervivencia sin efectos adversos del tratamiento. Además, aunque podemos fijarnos en una sola curva de supervivencia, a menudo el mayor interés reside en la comparación entre dos o más curvas de supervivencia, como las de los tratados y no tratados en un ensayo clínico aleatorizado. Al realizar dichas comparaciones, existen métodos estadísticos disponibles para determinar si una curva es significativamente diferente de otra.

Una tercera suposición es específica de las tablas de vida tradicionales, pero no del método de Kaplan-Meier, y trata del uso de intervalos predeterminados al calcular las tablas de vida. El motivo principal para emplear el método de la tabla de vida en vez del método de Kaplan-Meier es que si no podemos identificar el momento exacto en el que tuvo lugar el acontecimiento, debemos usar un intervalo arbitrario dentro del cual ocurrió dicho evento. Posteriormente, no podemos identificar el momento exacto en el que se produjeron las pérdidas de seguimiento del estudio. Por lo tanto, es importante asumir que existe una distribución uniforme del riesgo y de las pérdidas durante cada intervalo de tiempo, y que no hay un cambio rápido en el riesgo o en las pérdidas dentro de un intervalo de tiempo. Una forma razonable de lograr esta asunción es hacer que el intervalo sea lo más corto posible.

Ejemplo de utilización de una tabla de vida

Las tablas de vida se emplean en prácticamente todas las áreas clínicas. Sin embargo, en la actualidad se utilizan con menor frecuencia y han sido reemplazadas con el método de Kaplan-Meier, en el que los investigadores pueden identificar el momento exacto del evento para cada participante en el estudio. Las tablas de vida fueron la forma estándar de expresar y comparar la supervivencia durante mucho tiempo, antes de la aparición del método de Kaplan-Meier. Examinemos algunos ejemplos. Uno de los grandes triunfos de la pediatría en las últimas décadas ha sido el tratamiento de la leucemia infantil. Sin embargo, la mejoría ha sido mucho mayor en los blancos que en los negros, y los motivos de estas diferencias no están claros. En un periodo en el que las tasas de supervivencia de la leucemia aguda infantil estaban aumentando rápidamente se llevó a cabo un estudio para explorar las diferencias raciales en la supervivencia. Los datos de este estudio se muestran en las figuras 6.15 a 6.17[5]. Las curvas se basan en tablas de vida que fueron realizadas empleando el abordaje expuesto anteriormente.

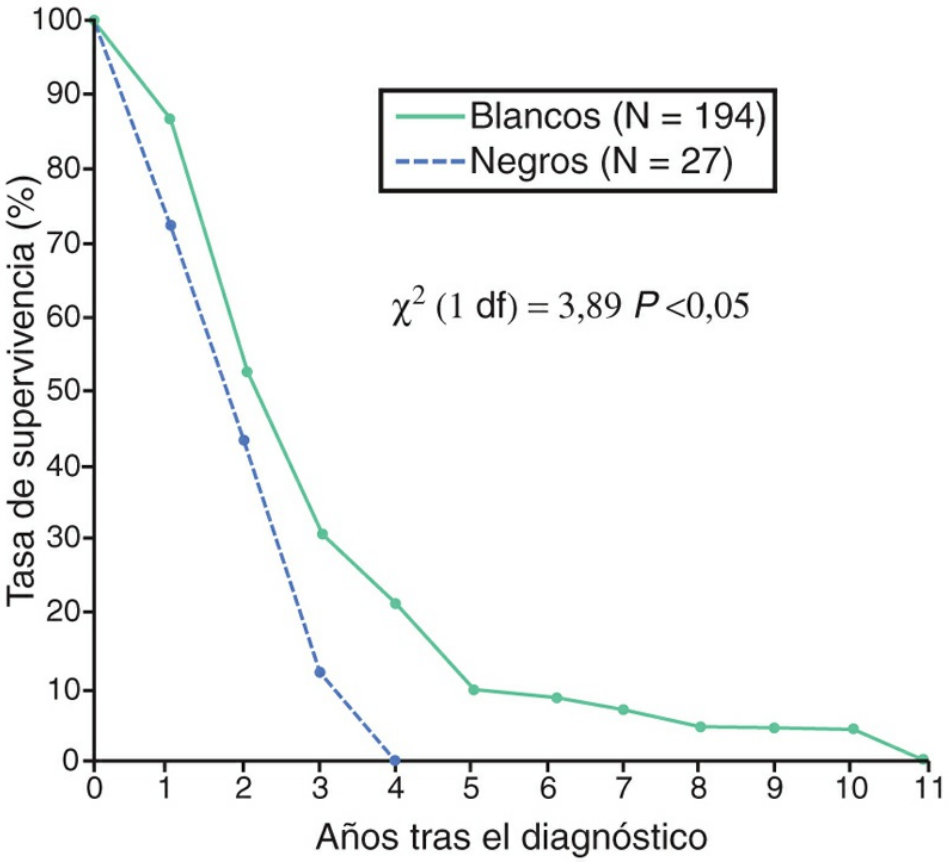


FIG. 6.15 Supervivencia de niños de 0 a 19 años con leucemia linfocítica aguda por raza, área metropolitana de Baltimore, 1960-1975. (De Szklo M, Gordis L, Tonascia J, Kaplan E. The changing survivorship of white and black children with leukemia. *Cancer.* 1978;42:59–66. Copyright © 1978 American Cancer Society. Reproducido con autorización de Wiley-Liss, Inc., una filial de John Wiley & Sons, Inc.)

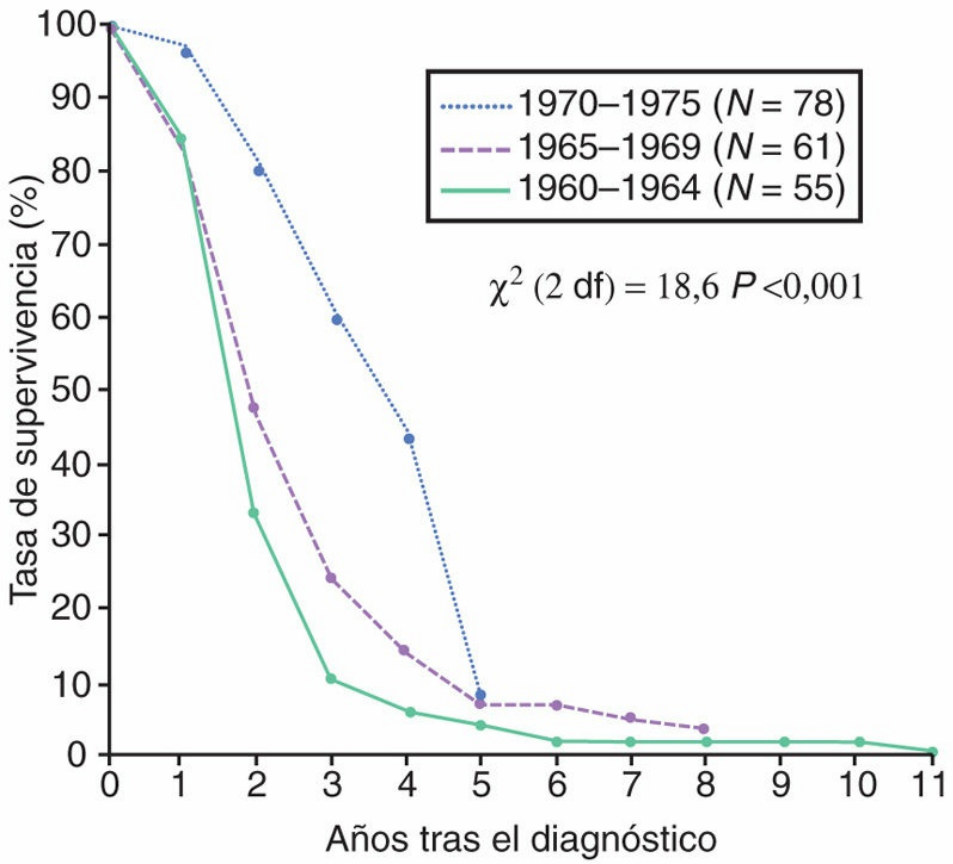


FIG. 6.16 Cambios temporales en la supervivencia de niños blancos de 0 a 19 años con leucemia linfocítica aguda, área metropolitana de Baltimore, 1960-1975. (De Szklo M, Gordis L, Tonascia J, Kaplan E. The changing survivorship of white and black children with leukemia. *Cancer.* 1978;42:59–66. Copyright © 1978 American Cancer Society. Reproducido con autorización de Wiley-Liss, Inc., una filial de John Wiley & Sons, Inc).

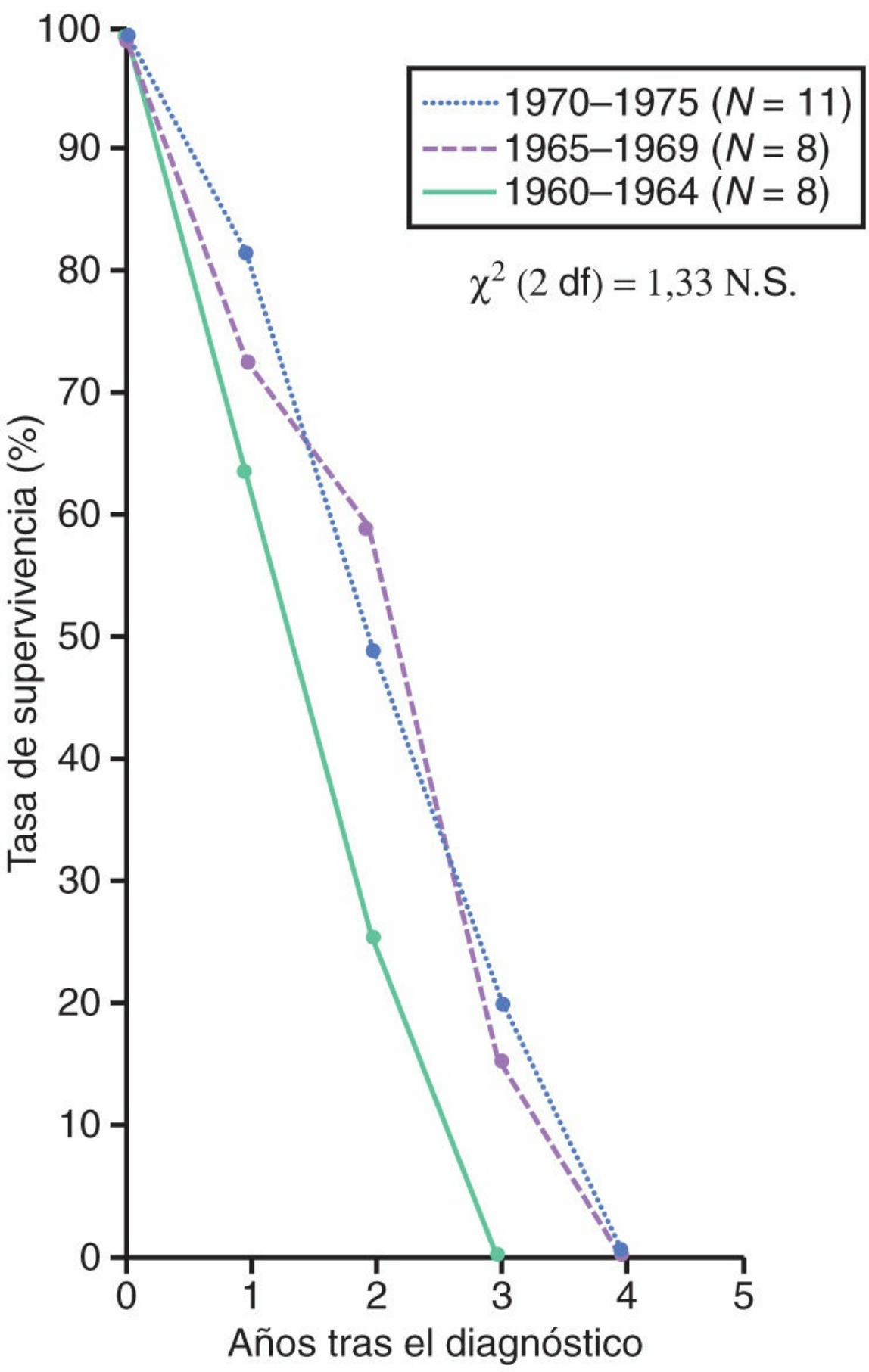


FIG. 6.17 Cambios temporales en la supervivencia de niños negros de 0 a 19 años con leucemia linfocítica aguda, área metropolitana de Baltimore, 1960-1975. (De Szklo M, Gordis L, Tonascia J, Kaplan E. The changing survivorship of white

and black children with leukemia. *Cancer.* 1978;42:59–66. Copyright © 1978 American Cancer Society. Reproducido con autorización de Wiley-Liss, Inc., una filial de John Wiley & Sons, Inc.)

En la figura 6.15 se muestra la supervivencia de los niños blancos y negros con leucemia en Baltimore a lo largo de un periodo de 16 años. Ningún niño negro sobrevivió más de 4 años, pero algunos niños blancos sobrevivieron hasta 11 años en este periodo de observación de 16 años.

¿Qué cambios tuvieron lugar en la supervivencia durante los 16 años del estudio? En las figuras 6.16 y 6.17 se muestran los cambios en la mortalidad por leucemia a lo largo del tiempo en los niños blancos y negros, respectivamente. El periodo de 16 años fue dividido en tres periodos: de 1960 a 1964 *(línea continua),* de 1965 a 1969 *(línea discontinua)* y de 1970 a 1975 *(línea de puntos).*

En los blancos (v. fig. 6.16), la supervivencia aumentó en cada periodo sucesivo. Por ejemplo, si examinamos la supervivencia a 3 años fijándonos en el punto de 3 años en cada curva sucesiva, observamos que la supervivencia mejoró del 8% al 25% y al 58%. Por el contrario, en los negros (v. fig. 6.17) se produjo una mejoría más leve de la supervivencia a lo largo del tiempo; las curvas de los dos periodos tardíos de los 5 años casi se superponen.

¿Qué explica esta diferencia racial? En primer lugar, debemos tener en cuenta los pequeños números involucrados y la posibilidad de que las diferencias pudieran haberse debido al azar. Asumamos, sin embargo, que las diferencias son reales. Durante las últimas décadas se han producido varios avances en el tratamiento de la leucemia a través de terapias combinadas, como la radiación del sistema nervioso central añadida a la quimioterapia. ¿Por qué existen entonces diferencias raciales en la supervivencia? ¿Por qué las mejoras terapéuticas que han sido tan efectivas en los niños blancos no han tenido un beneficio comparable en los niños negros? Análisis posteriores del intervalo desde el momento en el que la madre notó los síntomas hasta el momento del diagnóstico y el tratamiento indicaban que las diferencias en la supervivencia no parecían ser debidas a un retraso de los padres negros en buscar u obtener asistencia médica. Como la leucemia aguda es más grave en los negros y se encuentra más avanzada en el momento del diagnóstico, la diferencia racial podría reflejar las diferencias biológicas de la enfermedad, como una forma más agresiva y rápidamente progresiva de la enfermedad. La explicación definitiva es desconocida.

Efectos aparentes sobre el pronóstico de las mejoras diagnósticas

Hemos analizado la suposición asumida al usar una tabla de vida de que *no se ha producido mejora en la eficacia del tratamiento* a lo largo del tiempo de calendario durante el periodo del estudio. Otro aspecto del cálculo y la interpretación de las tasas de supervivencia es el posible efecto de las *mejoras en los métodos diagnósticos* a lo largo del tiempo de calendario.

Un ejemplo interesante fue comunicado por Feinstein y cols.[6] Compararon la supervivencia en una cohorte de pacientes con cáncer de pulmón tratados por primera vez en 1977 con la supervivencia en una cohorte de pacientes con cáncer de pulmón tratados de 1953 a 1964. La supervivencia a seis meses fue superior en el segundo grupo tanto para la totalidad del grupo como para los subgrupos creados según el estadio de la enfermedad. Los autores encontraron que la aparente mejora en la supervivencia se debió en parte a la *migración de estadios,* un fenómeno que se muestra en la figura 6.18A-C.

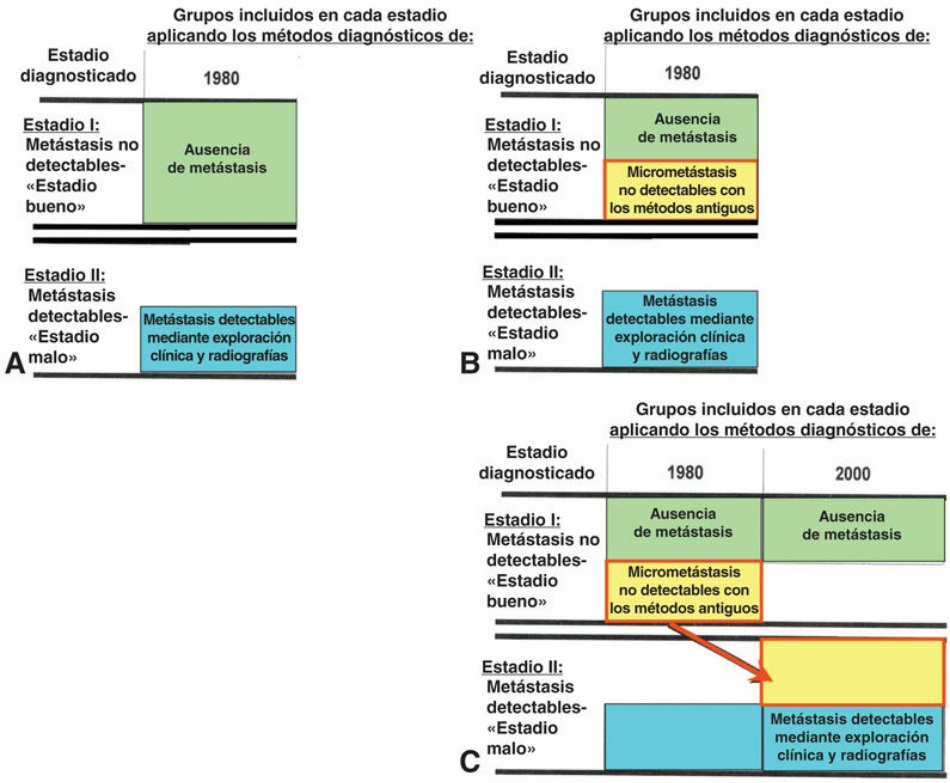


FIG. 6.18 Migración de estadios. (A) Clasificación de los casos según la presencia o ausencia de metástasis detectables en 1980. (B) Presencia de micrometástasis indetectables en 1980. (C) Impacto de las mejoras diagnósticas de las micrometástasis en el año 2000 sobre la clasificación de los casos en función de la presencia o ausencia de metástasis detectables.

En la figura 6.18A, los pacientes con cáncer son divididos en estadios

«buenos» y «malos» en función de si tenían metástasis detectables en 1980. Algunos pacientes que habrían sido asignados al estadio «bueno» en 1980 puede que tuvieran micrometástasis en ese momento que habrían pasado desapercibidas (v. fig. 6.18B). Sin embargo, en el año 2000, a medida que mejoraron las técnicas diagnósticas, muchos de estos pacientes habrían sido asignados al estadio «malo», porque sus micrometástasis ahora se habrían identificado utilizando las nuevas técnicas diagnósticas ya disponibles (fig. 6.18C). Si esto se hubiera producido, parecería que la supervivencia por estadio habría mejorado incluso aunque no hubiese aumentado la eficacia del tratamiento durante este tiempo.

Consideremos un ejemplo hipotético que ilustra este efecto de la migración de estadios. En la figura 6.19A-C se muestra un estudio hipotético de la tasa de letalidad en 300 pacientes con cáncer en dos periodos de tiempo, 1980 y 2000, *suponiendo que no se han producido mejoras en la eficacia del tratamiento disponible entre los dos periodos*. Asumiremos, como se muestra en la figura 6.19A, que en ambos periodos de tiempo la tasa de letalidad es del 10% para los pacientes sin metástasis, del 30% para los pacientes con micrometástasis y del 80% para los pacientes con metástasis. Fijándonos en la figura 6.19B, observamos que, en 1980, 200 pacientes fueron clasificados en el estadio I. Cien de estos pacientes no tenían metástasis y 100 presentaban micrometástasis ocultas. La tasa de letalidad en estos casos era del 10% y el 30%, respectivamente. En 1980, 100 pacientes presentaban claramente metástasis evidentes y fueron clasificados en el estadio II; su tasa de letalidad era del 80%.

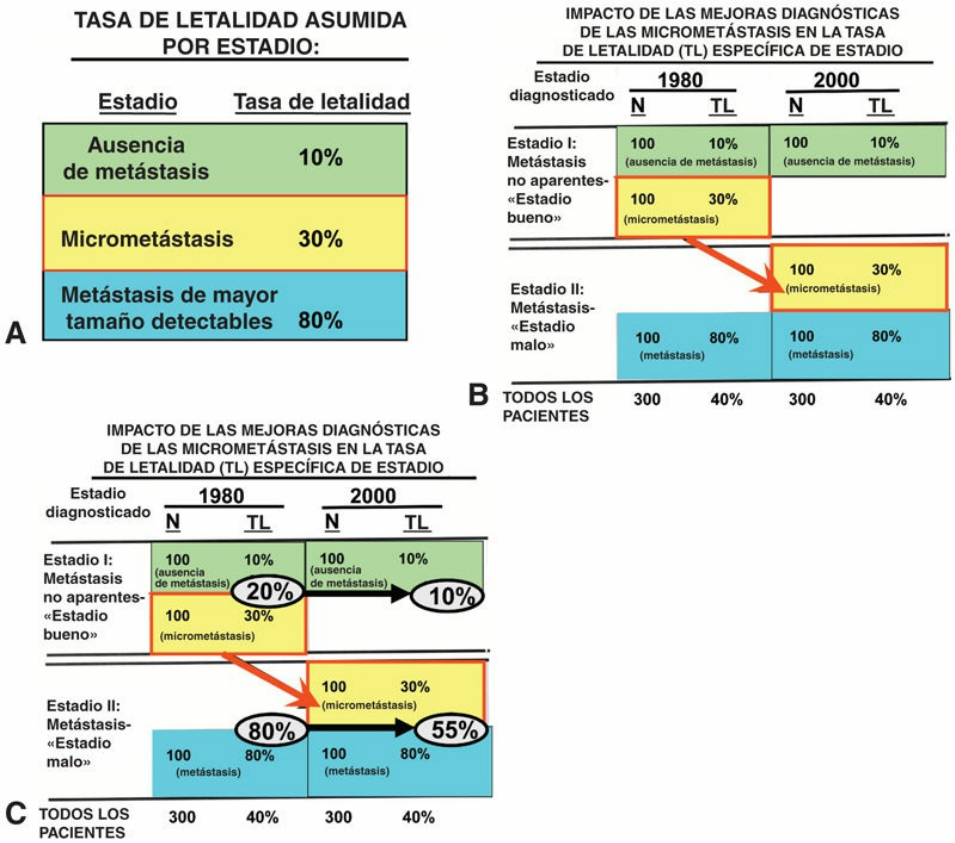


FIG. 6.19 Ejemplo hipotético de migración de estadios. (A) Tasa de letalidad asumida por estadio. (B) Impacto de las mejoras diagnósticas de las micrometástasis en la tasa de letalidad *(TL)* específica de estadio. (C) Mejoras aparentes en la supervivencia específica de estadio como resultado de la migración de estadios incluso sin mejoras en la eficacia del tratamiento.

Como resultado de las mejoras en las técnicas diagnósticas en el año 2000, se detectaron micrometástasis en los 100 pacientes afectados, y estos pacientes fueron clasificados en el estadio II (v. fig. 6.19C). Como el pronóstico de los pacientes con micrometástasis es peor que el de los otros pacientes del estadio I, y como, en el periodo tardío del estudio, los pacientes con micrometástasis ya no son incluidos en el grupo de estadio I (porque han migrado al estadio II), la tasa de letalidad de los pacientes del estadio I parece haber disminuido desde el 20% en el periodo inicial al 10% en el periodo tardío. Sin embargo, aunque el pronóstico de los pacientes que migraron del estadio I al estadio II fue peor que el de los otros pacientes en estadio I, el pronóstico de estos pacientes seguía siendo mejor que el de los otros pacientes en el estadio II, que tenían metástasis de mayor tamaño, de diagnóstico más fácil y una tasa de letalidad del 80%. Por tanto, la tasa de letalidad de los pacientes en estadio II también parece haber mejorado, habiendo disminuido desde el 80% en el periodo inicial hasta el 55% en el periodo tardío, *incluso en ausencia de mejora en la eficacia del tratamiento*.

Las mejoras aparentes en la supervivencia tanto en los pacientes en estadio I como en los pacientes en estadio II se deben solo al cambio de clasificación de los pacientes con micrometástasis en el periodo tardío. Si nos fijamos en la última línea de la figura, observamos que la tasa del letalidad del 40% para el

total de los 300 pacientes no ha cambiado desde el periodo inicial hasta el periodo tardío. Únicamente han cambiado las tasas de letalidad específicas de estadio aparente. Por tanto, es importante excluir la posibilidad de que se haya producido migración de estadios antes de atribuir la mejora aparente del pronóstico a la mayor eficacia de la asistencia médica.

A la migración de estadios los autores la denominan fenómeno de Will Rogers, en referencia a Will Rogers, un humorista americano durante la época de la depresión económica de la década de 1930. En esa época, debido a las dificultades económicas, muchos residentes de Oklahoma abandonaron su estado y emigraron a California. Rogers comentó: «Cuando los habitantes de Oklahoma abandonaron su estado y emigraron a California, aumentó el nivel medio de inteligencia en ambos estados.»

Mediana de supervivencia

Otra forma de expresar el pronóstico es mediante la *mediana de supervivencia*, que se define como el periodo de tiempo en el que sobrevive la mitad (50%) de la población del estudio. ¿Por qué deberíamos emplear la mediana de supervivencia en vez del tiempo medio de supervivencia, que es la media de los tiempos de supervivencia? La mediana de supervivencia ofrece dos ventajas sobre la supervivencia media. En primer lugar, se ve menos afectada por los extremos, mientras que la media puede verse muy afectada incluso por un solo valor extremo. Una o dos personas con un tiempo de supervivencia muy prolongado podrían afectar significativamente a la media, incluso aunque todos los otros tiempos de supervivencia fuesen mucho más cortos. En segundo lugar, si utilizáramos la supervivencia media, deberíamos observar todas las muertes del estudio antes de poder calcular la media. Sin embargo, para calcular la mediana de supervivencia, solo debemos observar las muertes de la mitad del grupo estudiado.

Supervivencia relativa

Consideremos la supervivencia a 5 años para un grupo de varones de 30 años con cáncer colorrectal. ¿Qué supervivencia a 5 años esperaríamos que tuvieran si no padeciesen un cáncer colorrectal? Claramente, sería casi del 100%. Por tanto, estamos comparando la supervivencia observada en varones jóvenes con cáncer colorrectal con una supervivencia de casi el 100% que es la esperada en los que no padecen cáncer colorrectal. ¿Qué pasaría si consideramos un grupo de varones de 80 años con cáncer colorrectal? En una población de esta edad no esperaríamos nada próximo a una supervivencia a 5 años del 100%, incluso aunque no padeciesen un cáncer colorrectal. Querríamos comparar la supervivencia observada en varones de 80 años con cáncer colorrectal con la supervivencia esperada en varones de 80 años sin cáncer colorrectal. Así, en todo grupo de personas con una enfermedad, queremos comparar su supervivencia con la supervivencia que cabría esperar en ese grupo de edad aunque *no* tuviese la enfermedad. Esta es la denominada *supervivencia relativa*.

La supervivencia relativa se define, por tanto, como el cociente entre la supervivencia observada y la supervivencia esperada:

$$\text{Supervivencia relativa} = \frac{\text{Supervivencia observada en personas con la enfermedad}}{\text{Supervivencia esperada si la enfermedad no estuviese presente}}$$

¿Tiene alguna importancia la supervivencia relativa? En la tabla 6.11 se muestran datos de supervivencia relativa y supervivencia observada en pacientes con cáncer de colon y recto, desde 1990 hasta 1998. Cuando nos fijamos en los grupos de edad más avanzada, que presentan altas tasas de mortalidad por otras causas, existe una gran diferencia entre la supervivencia observada y la supervivencia relativa. Sin embargo, en las personas jóvenes, que generalmente no se mueren de otras causas, la supervivencia observada y la supervivencia relativa en el cáncer de colon y recto no difieren de modo significativo.

Tabla 6.11

Supervivencia observada y relativa (%) a cinco años por edad para el cáncer de colon y

recto, 1990-1998 programa SEER, 1970-2011

Edad (años)	Supervivencia observada (%)	Supervivencia relativa (%)
<50	64	65
50-64	61,9	65,4
65-74	54,3	62,9
> 75	35,5	55,8

SEER, Surveillance, Epidemiology and End Results (Study).

Por cortesía del Dr. Louise Brinton y Mr. Jake Thistle del National Cancer Institute, empleando el programa SEER, 1970–2011.

Otra forma de ver la supervivencia relativa es examinando las curvas hipotéticas de supervivencia a 10 años en varones de 80 años, que se muestran en la figura 6.20A-D. Como referencia, en la figura 6.20A se muestra una curva de supervivencia perfecta del 100% (la curva horizontal de la parte superior) a lo largo de los 10 años del periodo del estudio. En la figura 6.20B se añade una curva de supervivencia observada, es decir, la supervivencia real observada en este grupo de pacientes con la enfermedad a lo largo de un periodo de 10 años. Como se observa en la figura 6.20C, la supervivencia esperada en este grupo de varones de 80 años es claramente menor del 100% porque en este grupo de edad las muertes por otras causas son importantes. La supervivencia relativa es el cociente entre la supervivencia observada y la supervivencia esperada. Como la supervivencia esperada se aleja de la supervivencia perfecta (100%), y la supervivencia esperada es el denominador para estos cálculos, la supervivencia relativa será mayor que la supervivencia observada (v. fig. 6.20D).

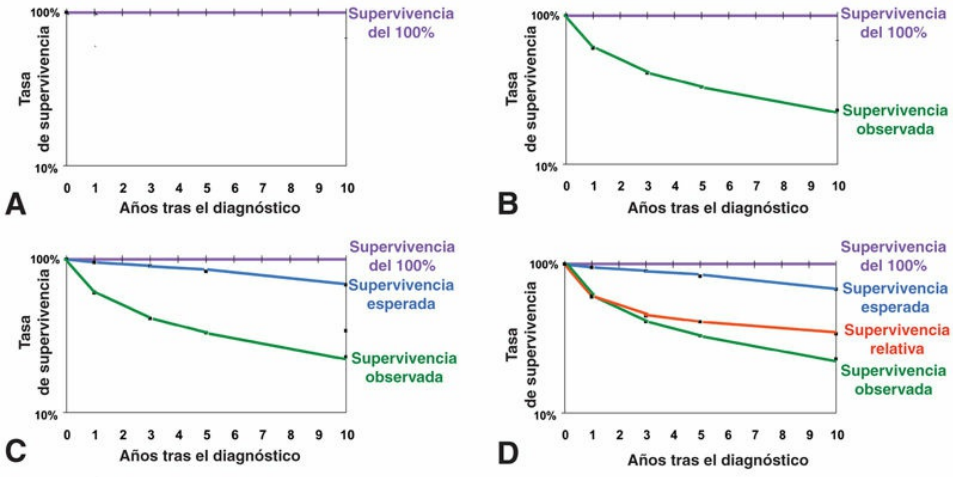


FIG. 6.20 Supervivencia relativa. (A) Supervivencia del 100% a lo largo de 10 años. (B) Supervivencia observada. (C) Supervivencia observada y esperada. (D) Supervivencia observada, esperada y relativa.

Generalización de los datos de supervivencia

Un último aspecto relacionado con la historia natural y el pronóstico de la enfermedad es la cuestión de qué pacientes son seleccionados para el estudio. Fijémonos en un ejemplo.

Las convulsiones febriles son frecuentes en los lactantes. Los niños por lo demás sanos a menudo sufren convulsiones asociadas con la fiebre elevada. La duda se plantea acerca de si estos niños deberían tratarse con un régimen de fenobarbital u otra medicación anticonvulsivante a largo plazo. Es decir, ¿las convulsiones febriles son un signo premonitorio de una epilepsia futura o se trata simplemente de un fenómeno asociado con la fiebre en los lactantes, en cuyo caso es poco probable que los niños sufran posteriormente convulsiones no febriles?

Para tomar una decisión lógica acerca del tratamiento, la pregunta que nos debemos plantear es: «¿Cuál es el riesgo de que un niño que ha presentado una convulsión febril sufra posteriormente convulsiones no febriles?». En la figura 6.21 se muestran los resultados de un análisis de Ellenberg y Nelson de los estudios publicados[7].

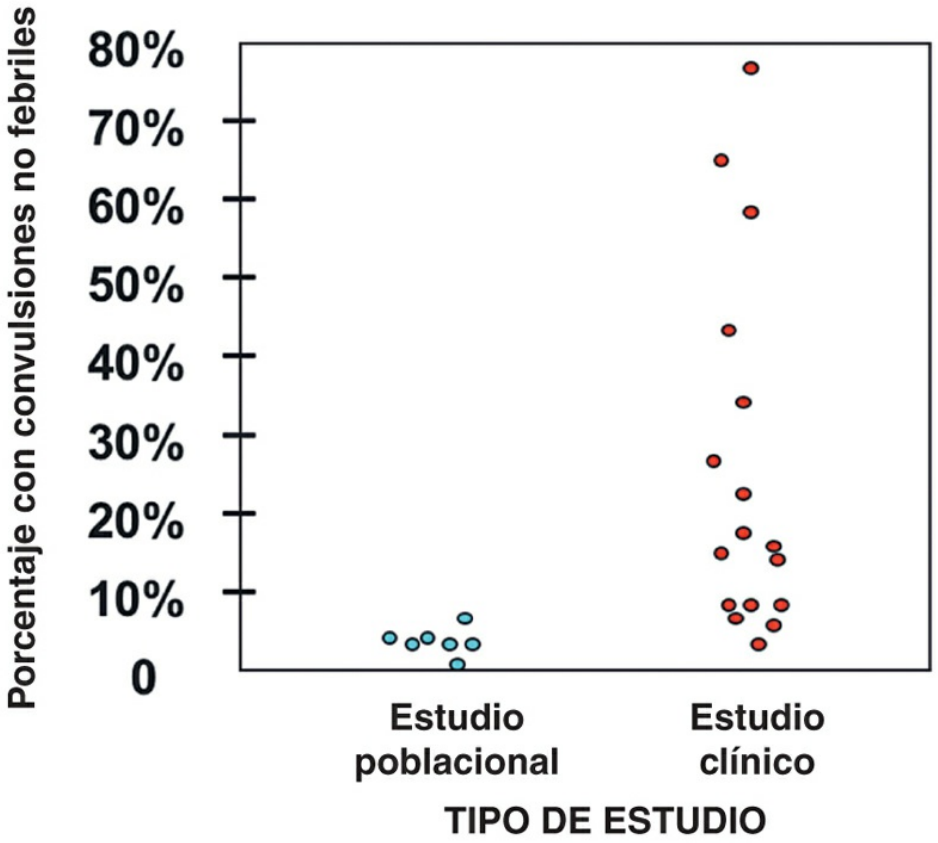


FIG. 6.21 Porcentaje de niños que sufrieron convulsiones no febriles tras uno o

más episodios de convulsiones febriles, por diseño de estudio. (Modificado de Ellenberg JH, Nelson KB. Sample selection and the natural history of disease: studies on febrile seizures. *JAMA.* 1980;243:1337–1340.)

Cada punto indica el porcentaje de niños con convulsiones febriles que posteriormente desarrollaron convulsiones no febriles en un estudio diferente. Los autores dividieron los estudios en dos grupos: estudios poblacionales y estudios clínicos basados en clínicas pediátricas o de epilepsia. Los resultados de diferentes estudios clínicos muestran un riesgo considerable de sufrir posteriormente convulsiones no febriles. Sin embargo, los resultados de los estudios poblacionales muestran poca variación en el riesgo, y los resultados de todos estos estudios suelen agruparse alrededor de un nivel de riesgo bajo.

¿Por qué deberían diferenciarse los dos tipos de estudios? ¿Qué resultados creería usted? Es probable que cada una de las clínicas tuviera diferentes criterios de selección y diferentes patrones de remisión. Por tanto, los diferentes riesgos observados en los diferentes estudios basados en clínicas son probablemente resultado de la selección de poblaciones diferentes en cada una de las clínicas. Por el contrario, en los estudios poblacionales (que, de hecho, pueden realizar una selección aleatoria), este tipo de variación debida a la selección se ve reducida o eliminada, lo que explica el agrupamiento cercano de los datos y el hallazgo resultante de que el riesgo de convulsiones no febriles es muy bajo. El punto importante es que puede resultar muy tentador analizar historiales de pacientes hospitalarios y generalizar los hallazgos para todos los pacientes en la población general. Sin embargo, este no es un abordaje válido porque los pacientes que acuden a una cierta clínica u hospital a menudo no son representativos de todos los pacientes de la comunidad. Esto no significa que los estudios realizados en un solo hospital o en una sola clínica carezcan de valor. De hecho, hay mucho que aprender de los estudios realizados en un solo hospital. Sin embargo, estos estudios son especialmente tendentes a sesgos de selección, y esta posibilidad siempre debe tenerse en cuenta cuando se interpretan los hallazgos de dichos estudios y su potencial para generalizar sus resultados.

Conclusión

Este capítulo ha expuesto cinco formas de expresar el pronóstico (cuadro 6.3). El mejor abordaje depende del tipo de datos disponibles, de los métodos de recogida de los datos y de la finalidad del análisis de los datos.

Cuadro 6.3 Cinco formas de expresar el pronóstico

1. Tasa de letalidad.
2. Supervivencia a 5 años.
3. Supervivencia observada.
4. Mediana de supervivencia.
5. Supervivencia relativa.

Bibliografía

1. Szklo M, Nieto FJ. Epidemiology: Beyond the Basics. 3rd ed. Burlington, MA: Jones & Bartlett; 2014.
2. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc*. 1958;53:457–481.
3. Rosenhek R, Binder T, Porenta G, et al. Predictors of outcome in severe, asymptomatic aortic stenosis. *N Engl J Med*. 2000;343:611–617.
4. Gibson AO, Blaha MJ, Arnan MK, et al. Coronary artery calcium and incident cerebrovascular events in an asymptomatic cohort. The MESA study. *JACC Cardiovasc Imaging*. 2014;7:1108–1115.
5. Szklo M, Gordis L, Tonascia J, et al. The changing survivorship of white and black children with leukemia. *Cancer*. 1978;42:59–66.
6. Feinstein AR, Sosin DM, Wells CK. The Will Rogers phenomenon: stage migration and new diagnostic techniques as a source of misleading statistics for survival in cancer. *N Engl J Med*. 1985;312:1604–1608.
7. Ellenberg JH, Nelson KB. Sample selection and the natural history of disease: studies on febrile seizures. *JAMA*. 1980;243:1337–1340.

Preguntas de repaso del capítulo 6

La pregunta 1 se basa en la información proporcionada en la siguiente tabla:

		N.° DE PACIENTES VIVOS EN CADA ANIVERSARIO DEL INICIO DEL TRATAMIENTO		
Año de tratamiento	N.° de pacientes tratados	1.°	2.°	3.°
2012	75	60	56	48
2014	63	55	31	
2015	42	37		
Total	180	152	87	48

Un total de 180 pacientes fueron tratados de la enfermedad X desde 2012 a 2014, y su evolución fue seguida hasta 2015. Los resultados

del tratamiento se presentan en la tabla. Ningún paciente se perdió durante el seguimiento.

1. *¿Cuál es la probabilidad de sobrevivir 3 años?* _____

2. *Una suposición importante en este tipo de análisis es que:*

a. El tratamiento ha mejorado durante el periodo del estudio.

b. La calidad del mantenimiento de los datos ha mejorado durante el periodo del estudio.

c. No se han producido cambios en la eficacia del tratamiento durante el periodo del estudio.

d. Cada año se incorporaron al estudio un número igual de varones y mujeres.

e. Ninguna de las anteriores.

3. *¿Cuál de los siguientes es un buen índice de la gravedad de una enfermedad aguda de corta evolución?*

a. Tasa de mortalidad específica de causa.

b. Supervivencia a 5 años.

c. Tasa de letalidad.

d. Razón de mortalidad estandarizada.

e. Ninguno de los anteriores.

4. *Se dispone de una prueba diagnóstica que detectará cierta enfermedad 1 año antes de lo que se detecta habitualmente. ¿Qué es lo más probable que le suceda a la enfermedad 10 años después de la aparición de la prueba? (Suponga que la detección precoz no ejerce ningún efecto sobre la historia natural de la enfermedad. Suponga también que no se han producido cambios en los certificados de defunción durante los 10 años.)*

a. La tasa de prevalencia de periodo disminuirá.

b. La supervivencia aparente a 5 años aumentará.

c. La tasa de mortalidad ajustada por edad disminuirá.

d. La tasa de mortalidad ajustada por edad aumentará.

e. La tasa de incidencia disminuirá.

5. *¿Cuál de las siguientes afirmaciones sobre la supervivencia relativa es verdadera?*

a. Se refiere a la supervivencia de los parientes de primer grado.

b. Suele ser más parecida a la supervivencia observada en las

poblaciones de edad avanzada.

c. Suele ser más parecida a la supervivencia observada en las poblaciones jóvenes.

d. Generalmente se diferencia de la supervivencia observada en una cantidad constante, independientemente de la edad.

e. Ninguna de las anteriores.

Las preguntas 6 a 8 se basan en los datos de la tabla que se muestra abajo. Los datos se obtuvieron de un estudio de 248 pacientes con síndrome de inmunodeficiencia adquirida (SIDA) que recibieron un nuevo tratamiento y fueron seguidos para determinar la supervivencia. La población del estudio fue seguida durante 36 meses.

Nota: realice los cálculos en la tabla con cuatro decimales (es decir; 0,1234), pero para la respuesta final use tres decimales (p. ej., 0,123 o 12,3%).

6. *En las personas que sobrevivieron el segundo año, ¿cuál es la probabilidad de morir en el tercer año?*

7. *¿Cuál es la probabilidad de que una persona incorporada al estudio sobreviva hasta el final del tercer año?*

Supervivencia de pacientes con síndrome de inmunodeficiencia adquirida tras el diagnóstico

(1) Intervalo desde el comienzo del tratamiento (meses)	(2) Vivos al comienzo del intervalo	(3) Muertos durante el intervalo	(4) Perdidos durante el intervalo	(5) Número efectivo expuestos al riesgo de morir durante el intervalo: Col (2) – ½[Col (4)]	(6) Proporción que murió durante el intervalo: Col (3)/ Col (5)	(7) Proporción que no murió durante el intervalo: 1 – Col (6)	(8) Proporción acumulada que sobrevivió desde la incorporación al final del intervalo: supervivencia acumulada
x	*lx*	*dx*	*wx*	*l'x*	*qx*	*px*	*Px*
1-12	248	96	27				
13-24	125	55	13				
25-36	57	55	2				

8. *Antes de comunicar los resultados de este análisis de supervivencia, los investigadores compararon las características basales de las 42 personas de*

las que se perdió el seguimiento antes de que acabara el estudio con las de los participantes que finalizaron el seguimiento. ¿Cuál fue el motivo de esta comparación?

a. Comprobar si la aleatorización fue exitosa.

b. Estudiar si se produjeron cambios en el pronóstico a lo largo del tiempo.

c. Comprobar si los que continuaron en el estudio representan a la población total del estudio.

d. Determinar si los resultados de los que continuaron en el estudio son los mismos que los de la población general.

e. Comprobar si existen factores de confusión en los grupos expuestos y no expuestos.

9. *Esta pregunta se basa en un estudio de Faraday y cols. que examinó la asociación entre el antecedente de infección cutánea y la infección del sitio quirúrgico (ISQ) después de una cirugía electiva. Siguieron a 613 pacientes durante aproximadamente 6 meses. La siguiente figura muestra las estimaciones de Kaplan-Meier de la incidencia acumulada de ISQ o muerte de etiología infecciosa en función de los antecedentes de infección de la piel. Según la figura, la mediana de supervivencia es:*

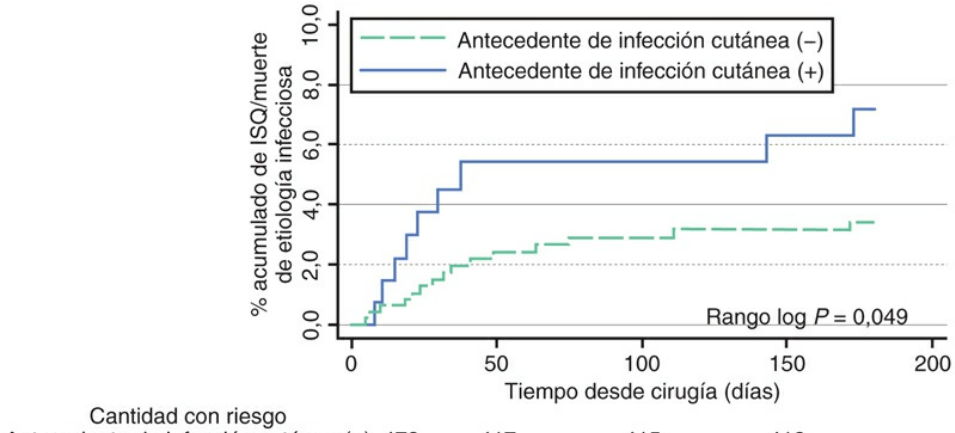
a. 30–50 días.

b. 60–80 días.

c. 90–110 días.

d. 120–130 días.

e. No se puede obtener información de esta figura.



(Faraday N, Rock P, Lin EE, et al. Past history of skin infection and risk of SSI

after elective surgery. *Ann Surg.* 2013;257:150–154.)

10. *En el estudio de Faraday (v. pregunta 9), ¿cuál de los siguientes es/son supuesto(s) necesario(s) cuando se usa el método de Kaplan-Meier para estimar la incidencia acumulada?*

a. La incidencia de eventos de ISQ es inferior al 10% en la población estudiada.

b. Aquellos que se pierden al seguimiento antes de los 6 meses tienen la misma experiencia de supervivencia que los que permanecen en el estudio.

c. Los eventos y la pérdida de seguimiento ocurren a una tasa constante durante cada intervalo de tiempo.

d. Aquellos censurados antes de los 6 meses tienen más probabilidades de desarrollar ISQ que aquellos que no están censurados antes de los 6 meses.

SECCIÓN 2

Empleo de la Epidemiología Para Identificar las Causas de las Enfermedades

Introducción
Capítulo 7: Estudios observacionales
Capítulo 8: Estudios de cohortes
Capítulo 9: Comparación de los estudios de casos y controles y de cohortes
Capítulo 10: Valoración de las medidas preventivas y terapéuticas: ensayos aleatorizados
Capítulo 11: Ensayos aleatorizados: algunos aspectos adicionales
Capítulo 12: Estimación del riesgo: ¿existe una asociación?
Capítulo 13: Conceptos adicionales sobre el riesgo: estimación de las posibilidades de prevención
Capítulo 14: De la asociación a la causalidad: derivación de inferencias a partir de estudios epidemiológicos
Capítulo 15: Más sobre inferencias causales: sesgos, confusión e interacción
Capítulo 16: Identificación de los factores genéticos y ambientales en la causalidad de la enfermedad

Introducción

En la sección 1 se abordaron las cuestiones de la definición y el diagnóstico de las enfermedades y la descripción de su transmisión, adquisición e historia natural en las poblaciones. La sección 2 se ocupará de un tema diferente: ¿cómo se diseñan y se llevan a cabo estudios para dilucidar la etiología y los factores de riesgo de las enfermedades humanas? Si organizamos una intervención preventiva ¿cómo saber si resultará efectiva? Este tipo de estudios son fundamentales tanto en la medicina clínica como en la práctica de la salud pública.

Esta sección comienza con una exposición de los diseños de estudios básicos que se utilizan en los estudios epidemiológicos (caps. 7 a 11). Comenzamos con las observaciones que puede realizar un médico en ejercicio al identificar un agrupamiento inusual de enfermedades. Describimos los componentes de los estudios observacionales (cap. 7), abordando primero las asociaciones a nivel comunitario y luego los estudios transversales. A continuación, pasamos a los estudios de cohortes (cap. 8) y las formas en que medimos las asociaciones. Sigue una breve comparación de los estudios de cohortes y de casos y controles (cap. 9). Los dos siguientes capítulos (10 y 11) exponen los ensayos aleatorizados, que son experimentos verdaderos. Luego describimos cómo se pueden usar los hallazgos de tales estudios para estimar los riesgos de enfermedad asociados con exposiciones específicas (caps. 12 y 13). Finalmente, abordamos los aspectos de la inferencia causal (caps. 15 y 16).

¿Por qué el médico debe prestar atención a la etiología de las enfermedades? ¿No ha sido el papel tradicional del médico tratar la enfermedad una vez que se ha hecho evidente? Para responder a esta pregunta, se deben hacer varias puntualizaciones. En primer lugar, la *prevención* es una de las principales responsabilidades del médico y de la comunidad de salud pública en general; tanto la prevención como el tratamiento deben ser vistos por el médico como elementos esenciales de su rol profesional. De hecho, muchos pacientes toman la iniciativa y plantean a sus médicos preguntas sobre qué medidas adoptar para mantener la salud y prevenir ciertas enfermedades. «¿Debo tomar aspirina infantil para prevenir enfermedades cardiovasculares?» «¿Realmente necesito realizarme mamografías con regularidad para la detección temprana del cáncer de mama?» «¿Cuál es la lectura de presión arterial más alta que aceptará antes de recetarme medicamentos para reducir mi presión sanguínea?» La mayoría de las oportunidades para prevenir los procesos patológicos requieren una

comprensión de la *etiología* o *causa* de dichas enfermedades, de modo que la exposición a un factor de riesgo causante pueda reducirse o la cadena patogénica que va desde el factor causal hasta el desarrollo de la enfermedad clínica se pueda interrumpir.

En segundo lugar, los pacientes y sus familias a menudo plantean preguntas al médico sobre el *riesgo* de padecer la enfermedad. ¿Cuál es el riesgo de que la enfermedad reaparezca? ¿Cuál es el riesgo de que otros miembros de la familia puedan desarrollar la enfermedad? Por ejemplo:

Un varón que sufre un infarto de miocardio a una edad temprana puede preguntar: «¿Por qué me sucedió a mí? ¿Puedo evitar tener un segundo infarto? ¿Mis hijos también tienen un riesgo elevado de sufrir un infarto a una edad temprana? En tal caso, ¿se puede hacer algo para reducir el riesgo?»

Una mujer que tiene un hijo con una malformación congénita puede preguntar: «¿Por qué ha sucedido? ¿Es por algo que hice durante el embarazo? Si me quedo embarazada nuevamente, ¿es probable que mi nuevo hijo también tenga una malformación?»

En tercer lugar, durante la práctica clínica y la realización de observaciones a la cabecera del paciente, un médico a menudo «tiene una corazonada» respecto a una posible relación todavía no conocida entre un factor y el riesgo de sufrir una enfermedad. Por ejemplo, Alton Ochsner, el famoso cirujano, señaló que prácticamente todos los pacientes a los que operó de cáncer de pulmón eran fumadores; esta observación le llevó a sugerir que el tabaquismo tenía una relación causal con el desarrollo de cáncer de pulmón y le indicó la necesidad de aclarar la naturaleza de esta relación mediante estudios llevados a cabo de forma rigurosa en poblaciones humanas definidas.

Mientras que la práctica clínica se centra en las personas, la práctica de la salud pública se centra en poblaciones que viven en comunidades. A la vista del enorme impacto potencial de las acciones de salud pública, que a menudo afectan a comunidades enteras, los profesionales de la salud pública deben entender cómo se llega a conclusiones con respecto a los riesgos sanitarios de una comunidad y cómo se desarrollan los fundamentos de las medidas y las acciones preventivas a partir de datos centrados en la población que se interpretan correctamente en su contexto biológico. Solo de esta manera se pueden adoptar políticas racionales para la prevención de las enfermedades y para mejorar la salud de las poblaciones con el menor coste posible.

Los médicos atentos y perspicaces, así como otros profesionales de la salud pública de los ámbitos académico, clínico y de los departamentos de salud, tienen muchas oportunidades para realizar estudios sobre la etiología de las enfermedades o sobre el riesgo de sufrir enfermedades con el fin de confirmar o refutar las impresiones preliminares clínicas o de otro tipo con respecto a los orígenes de las enfermedades. Los hallazgos podrían ser cruciales a la hora de proporcionar los fundamentos de la prevención de dichas enfermedades, para mejorar nuestra comprensión de su patogenia y para sugerir direcciones de las futuras investigaciones de laboratorio y

epidemiológicas. En consecuencia, la comprensión de los tipos de diseño de los estudios que se utilizan para la investigación de la etiología y la identificación de factores de riesgo, junto con una apreciación de los problemas metodológicos implicados en tales estudios, son fundamentales tanto para la medicina clínica como para la práctica de la salud pública.

Por último, esta sección finaliza con una exposición de cómo puede utilizarse la epidemiología para valorar las contribuciones relativas de los factores genéticos y ambientales en la etiología de la enfermedad humana, valoración que tiene unas implicaciones fundamentales para la práctica clínica y las políticas de salud pública (cap. 16).

CAPÍTULO 7

Estudios observacionales

Palabras clave

casos clínicos y series de casos
estudios ecológicos
estudios transversales
estudios de casos y controles
sesgo de información
estudios de casos cruzados
emparejamiento

Objetivos de aprendizaje

- Describir las motivaciones y el diseño de los estudios observacionales.
- Exponer los orígenes de los métodos de investigación, incluidos los casos clínicos, las series de casos y los estudios ecológicos.
- Describir el diseño del estudio transversal y su importancia.
- Analizar los estudios de casos y controles, incluyendo la selección de casos y controles.
- Analizar posibles sesgos de selección en los estudios de casos y controles.
- Analizar los sesgos de información en los estudios de casos y controles, incluidas las limitaciones en el recuerdo y el sesgo de recuerdo.
- Describir otros problemas en los estudios de casos y

controles, incluida la comparación y el uso de múltiples controles.

- Introducir el diseño de estudio de casos cruzados.

Casos clínicos y series de casos

Quizás uno de los orígenes más comunes y tempranos de las investigaciones médicas es la observación cuidadosa por parte de los médicos y otros proveedores de atención médica de lo que ven durante su práctica clínica. Dichas observaciones a nivel individual pueden documentarse en un caso clínico, que describe un fenómeno clínico particular en un solo paciente, o en una serie de casos, que describe más de un paciente con problemas similares. Tanto los casos clínicos como las series de casos se consideran los diseños de estudio más simples (aunque hay quien afirma que son simplemente «diseños pre-estudio»). El objetivo principal de los casos clínicos y las series de casos es proporcionar una descripción completa y detallada de los casos bajo observación. Esto permite a otros médicos identificar y potencialmente comunicar casos similares de su práctica, especialmente cuando comparten características clínicas específicas o geográficas. Por ejemplo, 2015 fue testigo de un brote del virus Zika en América Latina. El virus Zika es un flavivirus transmitido por mosquitos *Aedes*, más comúnmente *Aedes aegypti* y posiblemente *Aedes albopictus*, y originalmente aislado de un mono rhesus en el bosque Zika en Uganda en 1947[1]. A principios de 2016, después de un número creciente de bebés nacidos con microcefalia en áreas afectadas por el virus Zika, los Centros para el Control y la Prevención de Enfermedades (CDC, *Centers for Disease Control and Prevention*) publicaron una serie de casos descriptivos de Brasil sobre la posible asociación entre la infección por el virus del Zika y la microcefalia, una patología caracterizada porque la cabeza del lactante es significativamente más pequeña de lo esperado, posiblemente debido a un desarrollo cerebral incompleto[2]. Se publicó otro caso clínico sobre la descendencia de una mujer eslovena que vivía y trabajaba en Brasil y se quedó embarazada en febrero de 2015[3]. La paciente presentó un cuadro con fiebre alta, seguida de dolor musculoesquelético y retroocular intenso, y una erupción maculopapular pruriginosa generalizada. No se realizaron pruebas virológicas para detectar el virus del Zika. Regresó a Europa en la semana 28 de gestación, cuando las imágenes ecográficas mostraron anomalías fetales. El embarazo finalizó en la semana 32 de gestación a petición de la madre, tras la aprobación de los comités éticos nacionales e institucionales, y se aisló el virus Zika en el tejido cerebral fetal.

A pesar del hecho de que los casos clínicos y las series de casos son de naturaleza meramente descriptiva, sin un grupo de referencia para hacer una comparación estricta, las series de casos brasileñas fueron fundamentales para el desarrollo de las directrices de los CDC[4] (fig. 7.1) para la evaluación y la petición de pruebas, por parte de los profesionales sanitarios, de lactantes cuyas madres viajaron o residieron en un área con transmisión continua del virus Zika durante sus embarazos (fig. 7.2).

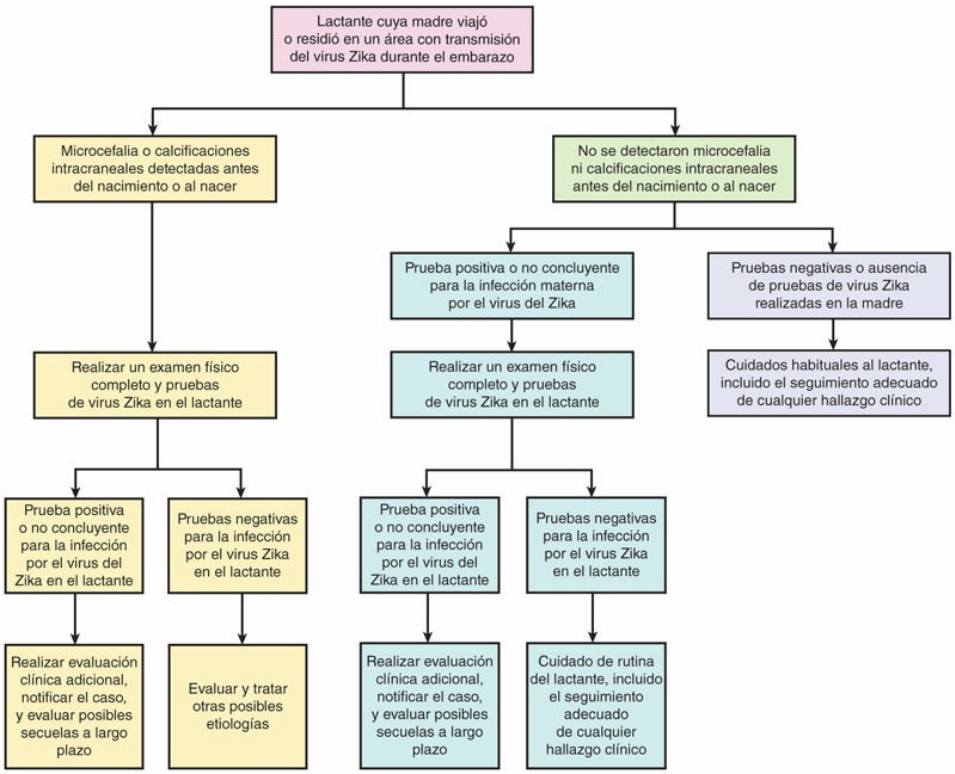


FIG. 7.1 Directrices provisionales para la evaluación y solicitud de pruebas de los lactantes cuyas madres viajaron o residieron en un área con transmisión continua del virus Zika durante sus embarazos. (Modificado de Staples JE, Dziuban EJ, Fischer M, et al. Interim guidelines for the evaluation and testing of infants with possible congenital Zika virus infection—United States, 2016. *MMWR Morb Mortal Wkly Rep*. 2016;65:63–67.)

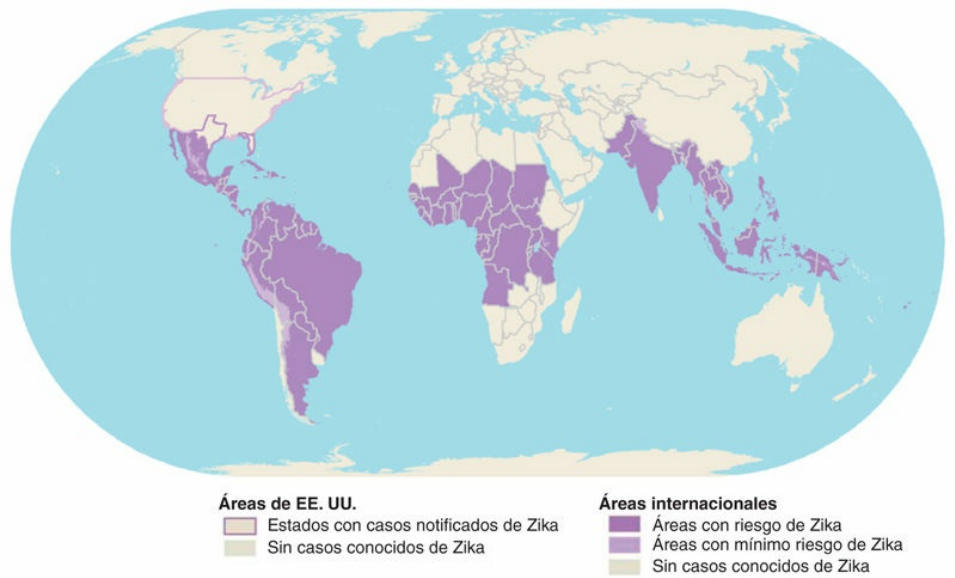


FIG. 7.2 Mapa mundial de zonas con riesgo de virus Zika. (De Centers for Disease Control and Prevention. https://wwwnc.cdc.gov/travel/files/zika-areas-of-risk.pdf. Consultado el 24 de julio de 2017.)

Los casos clínicos y las series de casos son herramientas clave para generar

hipótesis, especialmente cuando son simples, de bajo costo y fáciles de llevar a cabo en entornos clínicos de gran actividad. Sin embargo, la falta de un grupo de comparación es una desventaja importante. Además, la validez externa (generalización) es limitada, dada la selección sesgada de casos (todos identificados en la práctica clínica). Por último, cualquier asociación observada en un caso clínico o en una serie de casos puede verse influida por factores de confusión potencialmente no medidos sin el conocimiento de los investigadores.

Estudios ecológicos

El primer abordaje para determinar si existe una asociación puede ser el estudio de las características del grupo, los llamados *estudios ecológicos.* La figura 7.3 muestra la correlación entre el nivel de consumo de chocolate de cada país y su número de premios Nobel per cápita[5]. En esta figura, cada punto representa un país diferente. Como se desprende de esta figura, cuanto mayor es el consumo promedio de chocolate en un país, mayor es el número de premios Nobel per cápita. Se cree que el chocolate, rico en flavonoides dietéticos, mejora la función cognitiva y reduce el riesgo de demencia. Por lo tanto, podríamos sentirnos tentados a concluir que el consumo de chocolate puede ser un factor causal para obtener un premio Nobel. ¿Cuál es el problema de sacar tal conclusión de este tipo de estudio? Considere a Suiza, por ejemplo, que tiene el mayor número de premios Nobel per cápita y el mayor consumo promedio de chocolate. El problema es que no sabemos si las *personas* que ganaron el premio Nobel en ese país realmente consumían una cantidad elevada de chocolate. Todo lo que conocemos son valores *promedio* del consumo de chocolate y el número de premios Nobel per cápita en cada país. De hecho, se podría argumentar que, dado el mismo contexto general, es posible que aquellos que ganaron el Premio Nobel comieran muy poco chocolate. La figura 7.3 sola no revela si esto pudiera ser cierto; en efecto, los individuos en cada país se caracterizan por las cifras promedio (nivel de consumo y premios Nobel per cápita) para ese país. No se tiene en cuenta la variabilidad entre individuos en ese país con respecto al consumo de chocolate. Este problema se llama *falacia ecológica*: podemos atribuir a los miembros de un grupo algunas características que, de hecho, no poseen como individuos. Este problema surge en un estudio ecológico porque los datos solo están disponibles para grupos; no tenemos datos de exposición y resultados para cada individuo en la población.

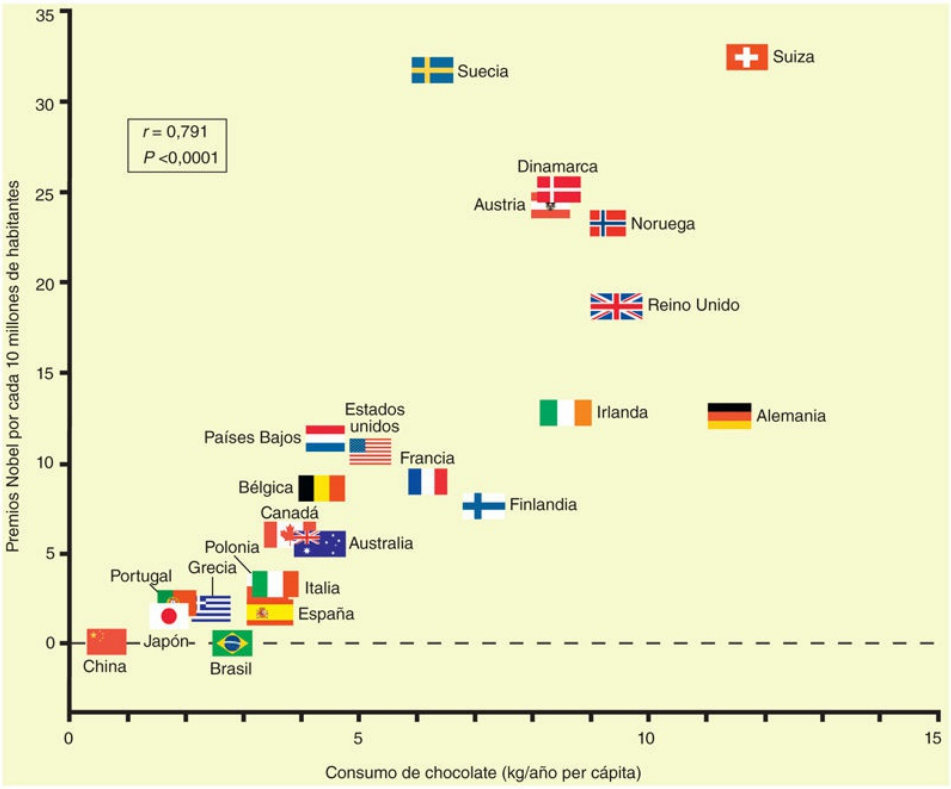


FIG. 7.3 Correlación entre el consumo anual per cápita de chocolate de los países y el número de premios Nobel por cada 10 millones de habitantes. (De Messerli FH. Chocolate consumption, cognitive function, and Nobel laureates. *N Engl J Med*. 2012;367:1562–1564.)

La tabla 7.1 muestra los datos de un estudio en el norte de California que explora la posible relación entre la exposición prenatal a la gripe durante un brote de gripe y el desarrollo posterior de leucemia linfocítica aguda en los niños[6]. La tabla muestra los datos de incidencia en los niños que ya habían nacido durante el brote de gripe y en los niños que estuvieron en el útero en el primer, segundo o tercer trimestre del embarazo durante el brote. Debajo de estas cifras, los datos se presentan como riesgos relativos, con el riesgo establecido en 1,0 para aquellos que no estuvieron en el útero durante el brote y las otras tasas se establecieron en relación con este aspecto. Los datos indican un alto riesgo relativo de leucemia en niños que estuvieron en el útero durante el brote de gripe en el primer trimestre.

Tabla 7.1

Tasas promedio anuales de incidencia bruta y riesgos relativos de leucemia linfocítica aguda por cohorte y trimestre de exposición a la gripe en niños menores de 5 años, San Francisco/Oakland (1969-1973)

	Sin exposición a la gripe	EXPOSICIÓN A LA GRIPE			
		TRIMESTRE			
		1.°	2.°	3.°	Total
Tasas de incidencia por 100.000	3,19	10,32	8,21	2,99	6,94
Riesgos relativos	1,0	3,2	2,6	0,9	2,2

Modificada de Austin DF, Karp S, Dworsky R, et al. Excess leukemia in cohorts of children born following influenza epidemics. *Am J Epidemiol*. 1977;10:77–83.

¿Cuál es el problema? Los propios autores declararon: «La asociación observada es entre el embarazo durante una epidemia de gripe y la leucemia subsiguiente en el fruto de dicha gestación. No se sabe si las madres de alguno de estos niños realmente sufrieron la gripe durante su embarazo»[6]. Lo que nos falta son *datos individuales* sobre la exposición (infección gripal). Uno podría preguntarse, ¿por qué los investigadores no obtuvieron los datos de exposición necesarios? La razón probable es que los investigadores emplearon certificados de nacimiento y datos de un registro de cáncer; ya que ambos tipos de datos son relativamente fáciles de obtener. Este enfoque no requirió seguimiento de los niños y contacto directo con sujetos individuales. Si nos impresionan estos datos ecológicos, podríamos querer realizar un estudio diseñado específicamente para explorar la posible relación entre la gripe prenatal y la leucemia. Sin embargo, tal estudio probablemente sería mucho más difícil y más costoso de realizar.

En vista de estos problemas, ¿son valiosos los estudios ecológicos? Sí, ya que pueden sugerir vías de investigación que pueden ser prometedoras para arrojar luz sobre las relaciones etiológicas. En sí mismos, sin embargo, no demuestran de manera concluyente que exista una verdadera asociación.

Durante muchos años, las preocupaciones legítimas sobre la posibilidad de falacia ecológica dieron mala reputación a los estudios ecológicos y desviaron la atención de la importancia de estudiar relaciones potencialmente significativas que solo pueden estudiarse ecológicamente, como las que existen entre el individuo y la comunidad en la que vive. Por ejemplo, Rose y cols.[7] estudiaron la relación de las características socioeconómicas y raciales de un vecindario y la práctica de angiografías en una muestra comunitaria que sufrió un infarto de miocardio (IM). Entre las 9.941 personas con IM que participaron en el Atherosclerosis Risk in Communities Study, en comparación con los blancos de los vecindarios con ingresos altos, se realizaron menos estudios angiográficos en los negros de los vecindarios con ingresos bajos y medios y en los blancos de las vecindarios con ingresos bajos. Por otro lado, los negros de vecindarios con ingresos altos y los blancos de vecindarios con ingresos medios no tuvieron desventajas con respecto a ser sometidos a estudio mediante angiografía. Por lo tanto, los estudios futuros que aborden tanto los factores de riesgo individuales como los factores de riesgo ecológicos, como las características del vecindario y las posibles

interacciones de ambos tipos de factores, pueden mejorar nuestro conocimiento acerca del acceso a un examen angiográfico.

Otro ejemplo de la importancia de los datos ecológicos es la esquistosomiasis, una enfermedad causada por esquistosomas, un parásito de agua dulce que puede afectar el tracto genitourinario o gastrointestinal, así como el sistema nervioso central, y que también es un factor de riesgo para el cáncer de hígado y vejiga. La población está expuesta mediante contacto con agua infestada. Los que viven en comunidades rurales tienen el mayor riesgo de sufrir esquistosomiasis; la exposición puede presentarse en agricultores o pescadores, mujeres que lavan ropa o niños que juegan en el agua infestada. Egipto tiene la mayor prevalencia endémica de esquistosomiasis de todo el mundo, que se remonta a su periodo dinástico[8]. El uso de la terapia antiesquistosomal parenteral (TAP) con tartrato de antimonio y potasio, comúnmente llamado tártaro emético, se ha utilizado para el tratamiento masivo en Egipto desde la década de 1920, mediante 12 inyecciones intravenosas semanales. Estas inyecciones se realizaban con jeringas de vidrio reutilizables, en general sin los procedimientos de esterilización adecuados, lo que puede haber sido responsable de que Egipto tenga la mayor prevalencia de hepatitis C en el mundo[9]. (El tártaro emético fue el único tratamiento para la esquistosomiasis hasta el praziquantel, un medicamento altamente eficaz por vía oral que se introdujo en la década de 1980). En el año 2000, Frank y cols.[10] estudiaron la asociación ecológica en Egipto entre el uso anual de TAP con tártaro emético y la seroprevalencia de anticuerpos contra el virus de la hepatitis C (VHC) en 8.499 egipcios de 10 a 50 años. En conjunto, se observó que la prevalencia ajustada por edad de los anticuerpos contra el VHC era del 21,9%. La figura 7.4 muestra la asociación entre la prevalencia específica por región de anticuerpos contra el VHC con la exposición a TAP específica por región, lo que sugiere que la variación en la seroprevalencia de los anticuerpos contra el VHC entre regiones puede explicarse por la exposición al TAP (razón de posibilidades 1,31 [intervalo de confianza {CI} 95%: 1,08 − 1,59]; $P = 0,007$). Hasta la fecha, la transmisión masiva del VHC a través del uso del TAP en Egipto se considera la mayor transmisión iatrogénica de un patógeno transmitido por la sangre.

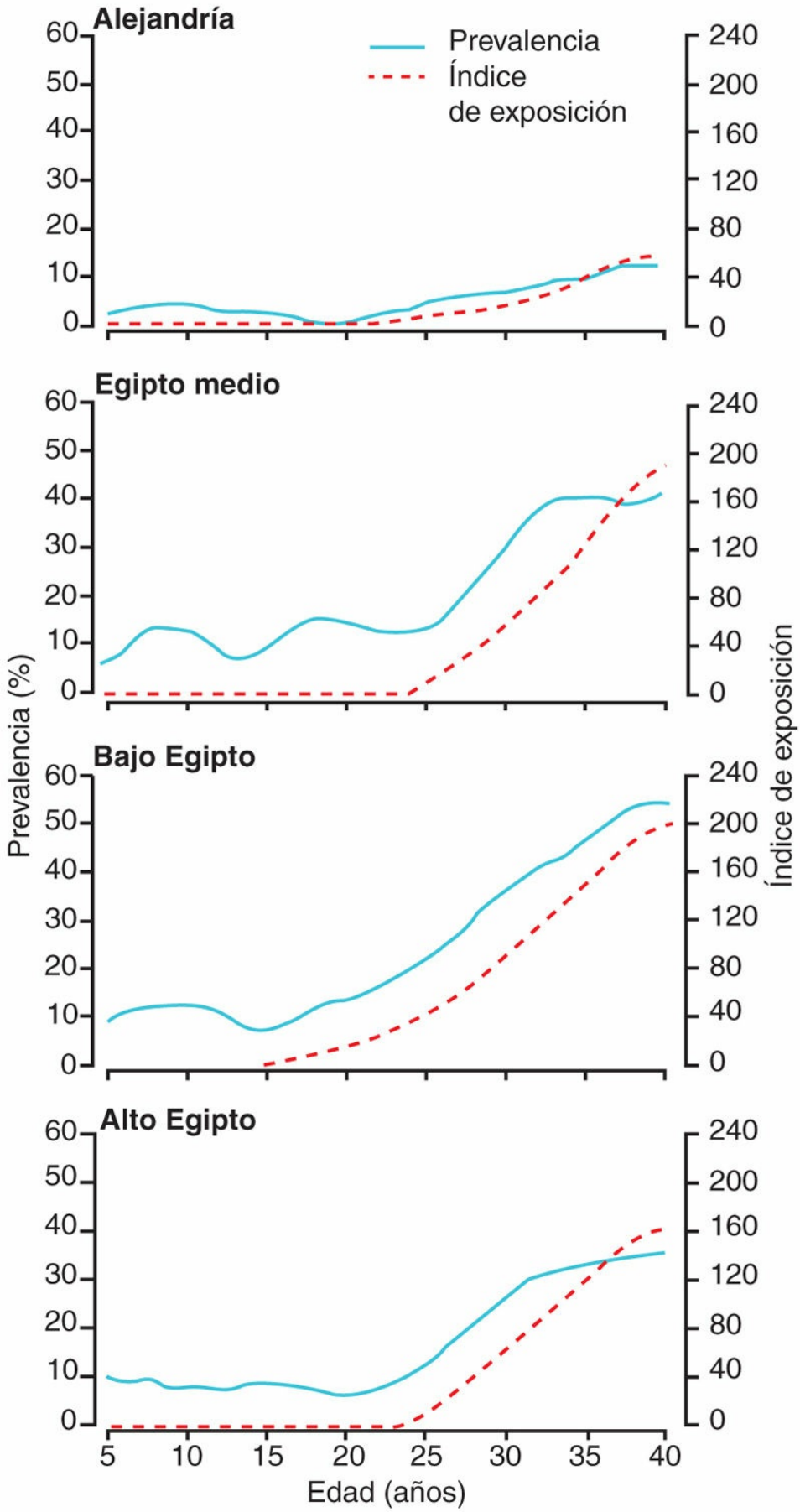


FIG. 7.4 Comparación de la prevalencia del virus de la hepatitis C y el índice de exposición por cohorte y región. (Reproducido con permiso de Frank C, Mohamed MK, Strickland GT, et al. The role of parenteral antischistosomal therapy in the

spread of hepatitis C virus in Egypt. *Lancet*. 2000;11;355:887–891.)

Se ha afirmado que como los epidemiólogos generalmente muestran datos tabulados y se refieren a las características de los grupos, los datos de todos los estudios epidemiológicos son datos de grupos. Esto no es verdad. En los estudios transversales, de casos y controles, de cohortes y en los ensayos aleatorizados, se dispone de datos sobre la exposición y el resultado de la enfermedad para cada individuo en el estudio, aunque estos datos se agrupan generalmente en tablas y figuras. Por otro lado, solo los datos agrupados están disponibles en estudios ecológicos, como, por ejemplo, los datos sobre el consumo medio de sal y la presión arterial sistólica promedio por país.

Curiosamente, cuando la variabilidad de una exposición es limitada, las correlaciones ecológicas pueden proporcionar una respuesta más válida con respecto a la presencia de una asociación que los estudios basados en individuos. Wynder y Stellman han resumido este fenómeno de la siguiente manera: «Si los casos y los controles se extraen de una población en la que el rango de exposiciones es limitado, un estudio puede proporcionar poca información sobre los posibles efectos en la salud»[11].

Un ejemplo es la relación entre la ingesta de sal y la presión arterial, que no se ha encontrado de forma consistente en los estudios de casos y controles y de cohortes; sin embargo, en una correlación ecológica que utiliza las poblaciones de países como unidades analíticas, se ha observado una correlación importante y graduada. Este fenómeno puede explicarse por el reducido rango de la ingesta de sal en los individuos dentro de cada país, pero una variabilidad bastante grande de la ingesta *promedio* de sal entre diferentes países.

Estudios transversales

Otro diseño de estudio común utilizado en la investigación inicial de la asociación entre una exposición específica y una enfermedad de interés es el *estudio transversal*. Supongamos que estamos interesados en la posible relación entre el aumento del nivel de colesterol sérico (la *exposición*) y la evidencia electrocardiográfica (ECG) de arteriopatía coronaria (AC, la *enfermedad*). En cada participante de la población estudiada determinamos el nivel de colesterol sérico y realizamos un ECG para detectar la presencia de AC. La presencia de AC define un caso prevalente. Este tipo de diseño de estudio se denomina *estudio transversal* porque la exposición y el resultado de la enfermedad se determinan *simultáneamente* para cada participante del estudio; es como si estuviéramos viendo una instantánea de la población en un momento determinado. Otra forma de describir un estudio transversal es imaginar que hemos seccionado una porción de la población, capturando niveles de colesterol *y* evidencia de AC al mismo tiempo. Se debe tener en cuenta que en este tipo de enfoque, los casos de enfermedad que identificamos son casos *prevalentes* de la enfermedad en cuestión (razón por la cual un estudio transversal también se denomina «estudio de prevalencia»), porque sabemos que existían en el momento del estudio, pero no sabemos su duración (el intervalo entre el inicio de la enfermedad y el «momento actual»), o si la exposición ocurrió antes del resultado. La imposibilidad de determinar una secuencia temporal de «exposición-enfermedad» puede resultar en un sesgo temporal cuando es la enfermedad la que causa la exposición. Por ejemplo, los casos prevalentes de AC pueden participar en actividades físicas de ocio más a menudo que los sujetos normales, ya que la aparición de un episodio agudo de AC puede hacer que los médicos recomienden el ejercicio físico a sus pacientes con AC, un fenómeno que también se conoce como «causalidad inversa». (Sin embargo, tenga en cuenta que cuando la información sobre la exposición se obtiene mediante un cuestionario, es posible determinar si una exposición determinada [p. ej., hábitos sedentarios, tabaquismo o consumo excesivo de alcohol] estuvo presente antes del inicio de la enfermedad, permitiendo así la identificación de la secuencia temporal entre la exposición y la enfermedad).

Además del sesgo temporal, en un estudio transversal también puede producirse un sesgo de supervivencia/selección cuando la exposición está relacionada con la duración de la enfermedad; así, por ejemplo, si los casos incidentes inducidos por la exposición tienen una supervivencia más corta que los casos incidentes no expuestos, los casos prevalentes, que son por definición supervivientes, pueden tener una proporción menor de exposición pasada que los que se hubieran observado si se hubieran incluido los casos incidentes en el estudio. En otras palabras, identificar solo los casos prevalentes excluiría a aquellos que murieron poco después de desarrollar la enfermedad, pero antes de que se realizara el estudio. Por ejemplo, es un

hecho conocido que una concentración sérica elevada de colesterol causa AC. Sin embargo, cuando se realiza un estudio transversal, la asociación observada puede ser una función tanto del riesgo de *desarrollar* AC como de la *supervivencia* después del inicio de la AC.

Otro ejemplo de sesgo de supervivencia es el enfisema pulmonar inducido por el tabaquismo. Fumar no solo causa enfisema, sino que, además, la supervivencia de los pacientes con enfisema inducido por tabaquismo es peor que la de los pacientes cuyo enfisema se debe a otras causas (p. ej., asma o bronquitis crónica). Como resultado, el antecedente de tabaquismo se observará con menos frecuencia en los casos prevalentes de enfisema que en los incidentes. Este tipo de sesgo de supervivencia también se conoce como *sesgo de prevalencia-incidencia.*

En vista de sus sesgos, los resultados de un estudio transversal deben usarse para generar hipótesis que luego puedan evaluarse utilizando un diseño de estudio que incluya casos incidentes y permita establecer la secuencia temporal de la exposición y el resultado. Sin embargo, los estudios transversales, como las encuestas políticas y las encuestas por muestreo, se utilizan ampliamente y, a menudo, son los primeros estudios realizados antes de pasar a diseños de estudios más válidos.

El diseño general de un estudio transversal (o prevalencia) se expone en la figura 7.5. Definimos una población y determinamos la presencia o ausencia de exposición y la presencia o ausencia de enfermedad para cada individuo al mismo tiempo. Cada individuo puede clasificarse en uno de los cuatro subgrupos posibles.



FIG. 7.5 Diseño de un estudio transversal: I. Identificación de cuatro subgrupos basados en la presencia o ausencia de exposición y presencia o ausencia de enfermedad.

Como se observa en la tabla 2 × 2 de la parte superior de la figura 7.6, habrá

a personas que hayan estado expuestas y tengan la enfermedad; *b* personas, que hayan estado expuestas pero no tengan la enfermedad; *c* personas que tienen la enfermedad pero no han sido expuestas; y *d* personas, que ni han estado expuestas ni tienen la enfermedad.

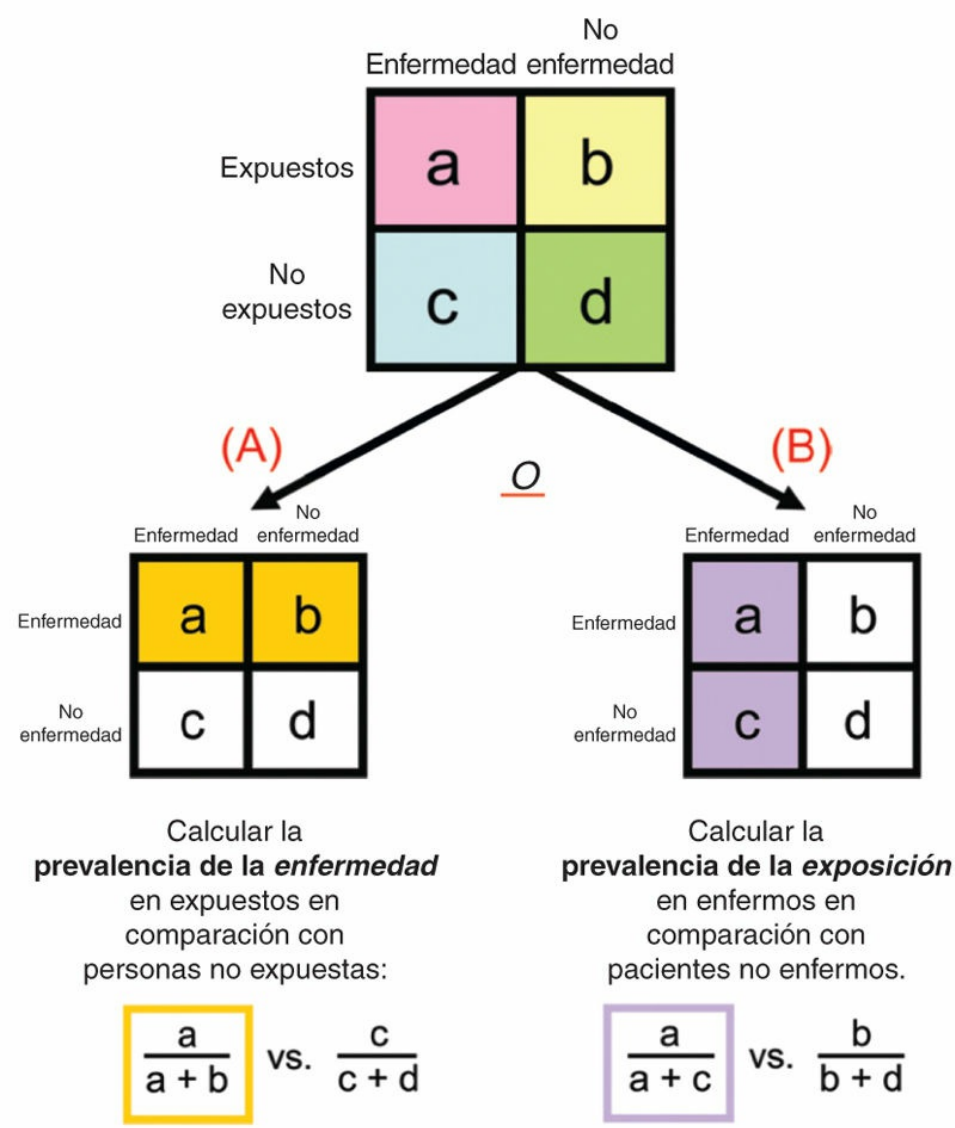


FIG. 7.6 Diseño de un estudio transversal hipotético II: *(arriba)* Tabla 2 × 2 con los hallazgos del estudio; *(abajo)* dos posibles enfoques para el análisis de resultados: (A) Calcular la prevalencia de la enfermedad en las personas expuestas en comparación con la prevalencia de la enfermedad en las personas no expuestas, *o* (B) Calcular la prevalencia de la exposición en personas con la enfermedad en comparación con la prevalencia de la exposición en personas sin la enfermedad.

Para determinar si existe evidencia de una asociación entre la exposición y la enfermedad a partir de un estudio transversal, podemos elegir entre dos enfoques posibles, que en la figura 7.6 se denominan (A) y (B). Si usamos (A), podemos calcular la prevalencia de la enfermedad en personas con la

exposición $\left(\frac{a}{a+b}\right)$ y compararlo con la prevalencia de la enfermedad en personas sin la exposición $\left(\frac{c}{c+d}\right)$. Si usamos (B), podemos comparar la prevalencia de exposición en personas con la enfermedad $\left(\frac{a}{a+c}\right)$ con la prevalencia de exposición en personas sin la enfermedad $\left(\frac{b}{b+d}\right)$. Los detalles de ambos abordajes se muestran en la parte inferior de la figura 7.6.

Si en un estudio de este tipo determinamos que parece existir una asociación entre el aumento de la concentración de colesterol y la AC, debemos considerar varios aspectos. Primero, en este estudio transversal, estamos identificando los casos prevalentes (existentes) de AC en lugar de los incidentes (nuevos); tales casos prevalentes pueden no ser representativos de todos los casos de AC que se han desarrollado en esta población. Por ejemplo, identificar solo los casos prevalentes excluiría a aquellos que murieron después de que se desarrolló la enfermedad, pero antes de que se realizara el estudio. Por lo tanto, incluso si se observa una asociación entre exposición y enfermedad, la asociación puede ser con la *supervivencia* después de una AC y no con el riesgo de *desarrollar* una AC. Segundo, debido a que la presencia o ausencia de exposición y enfermedad se determinó *al mismo tiempo* en cada participante en el estudio, a menudo no es posible establecer una relación temporal entre la exposición y el inicio de la enfermedad. Por lo tanto, del ejemplo presentado al comienzo de esta sección, no es posible inferir si el aumento en el nivel de colesterol *precedió* o no al desarrollo de AC. Sin información sobre las relaciones temporales, cabe la posibilidad de que el aumento de la concentración de colesterol podría haber ocurrido como resultado de la AC, lo que se denomina «causalidad inversa», o tal vez ambos podrían deberse a otro factor. Si resulta que la exposición no precedió al desarrollo de la enfermedad, la asociación no puede reflejar una relación causal.

Farag y cols. emplearon datos del National Health and Nutrition Examination Survey (NHANES), un estudio representativo a nivel nacional de la población civil estadounidense no internada, para examinar una posible asociación entre la vitamina D y la disfunción eréctil en hombres sin enfermedad cardiovascular[12]. Se encontró una relación dosis-respuesta entre la deficiencia de vitamina D y la disfunción eréctil (tasa de prevalencia 1,30; IC 95%: 1,08-1,57; fig. 7.7). A pesar de los sesgos inherentes al diseño transversal, los hallazgos del estudio sugieren la necesidad de realizar un

ensayo aleatorizado sobre la asociación de la deficiencia de vitamina D y la función eréctil.

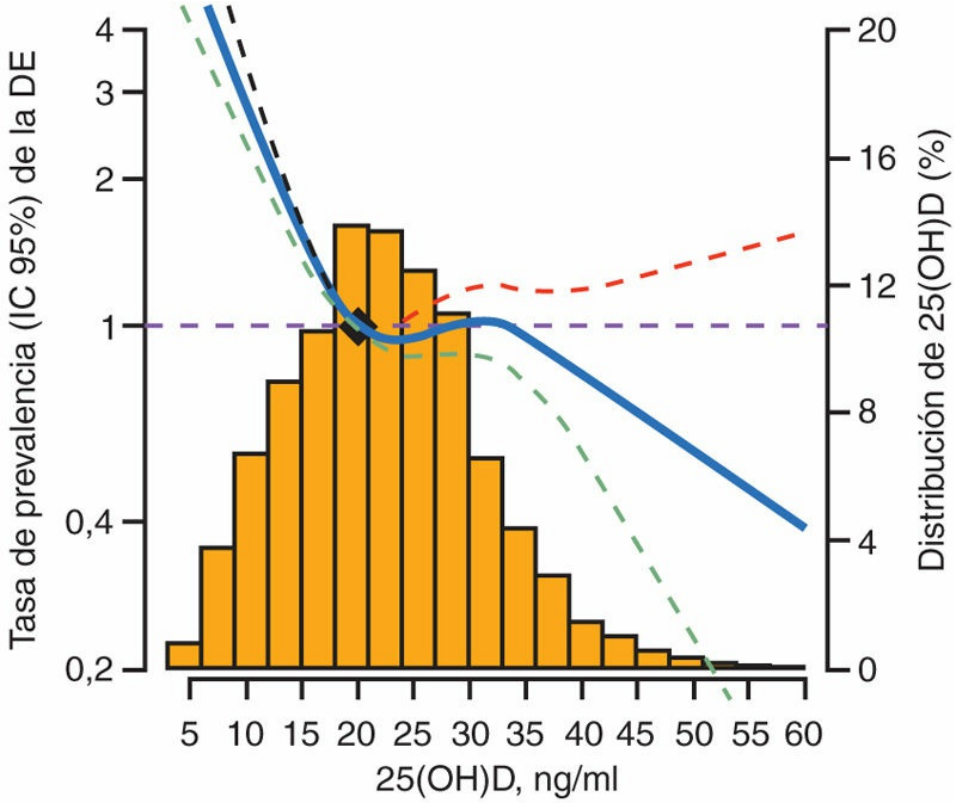


FIG. 7.7 *Spline* cúbico limitado de 25(OH)D y tasa de prevalencia ajustada de la disfunción eréctil *(DE),* NHANES 2001-2004. Las curvas representan la tasa de prevalencia ajustada *(línea continua)* y los intervalos de confianza del 95% *(líneas discontinuas)* basadas en splines cúbicos limitados para la concentración de 25(OH)D con nudos a 10, 20, 30 y 40 ng/ml. Los valores de referencia se fijaron en 20 ng/ml. El modelo se ajusta según la edad, la raza, el tabaquismo, el consumo de alcohol, el índice de masa corporal, la actividad física, la hipertensión, la diabetes, la hipercolesterolemia, la tasa de filtración glomerular estimada, la proteína C reactiva y el uso de antidepresivos y betabloqueantes. (De Farag YM, Guallar E, Zhao D, et al. Vitamin D deficiency is independently associated with greater prevalence of erectile dysfunction: the National Health and Nutrition Examination Survey (NHANES) 2001-2004. *Atherosclerosis.* 2016;252:61–67.)

Los estudios transversales en serie también son útiles para evaluar las tendencias en la prevalencia de la enfermedad a lo largo del tiempo y así poder informar a los responsables de la planificación de las políticas sanitarias. Murphy y cols. utilizaron los datos anuales de NHANES, anualmente de 1988 a 1994 y bianualmente de 1999 a 2012, para examinar las tendencias en la prevalencia de la enfermedad renal crónica (ERC)[13]. La figura 7.8 muestra las tendencias temporales en la prevalencia ajustada de los estadios 3 y 4 de la ERC según el estudio NHANES de 1988-1994 hasta 2011-2012, clasificado por la presencia o ausencia de diabetes. Como se muestra en la figura, se observa un aumento inicial en la prevalencia ajustada de las

etapas 3 y 4 de la ERC que se estabilizó a principios de la década de 2000 entre los pacientes no diabéticos, pero continuó aumentando en los diabéticos.

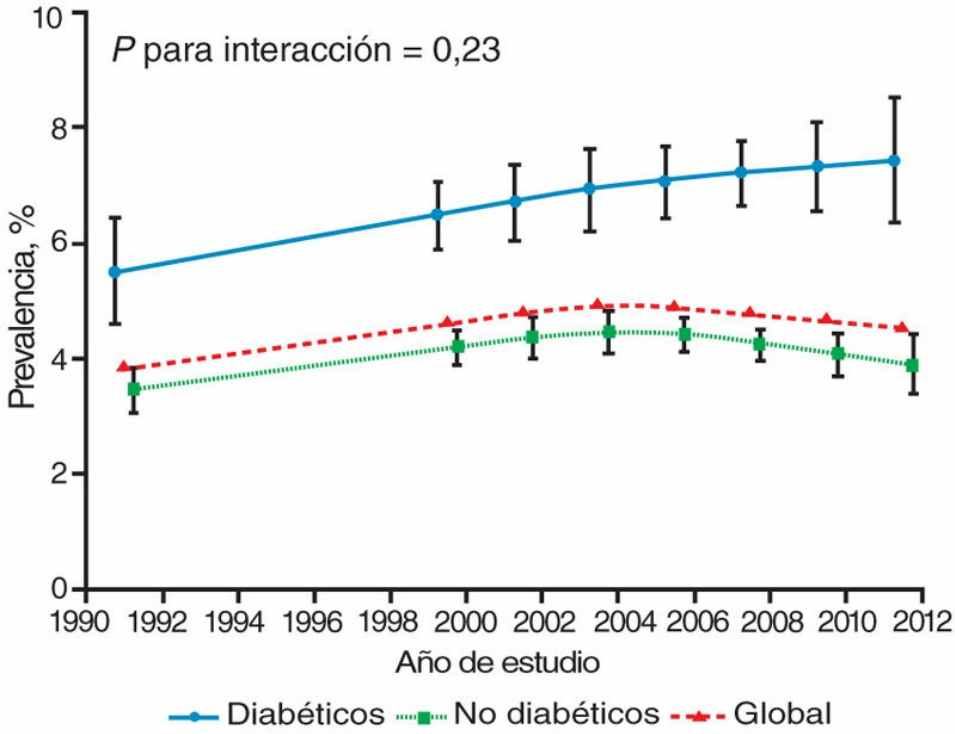


FIG. 7.8 Prevalencia ajustada de enfermedad renal crónica en estadios 3 y 4 (tasa de filtración glomerular estimada de 15 a 59 ml/min/1,73 m^2 calculada con la ecuación de la Chronic Kidney Disease Epidemiology Collaboration) en adultos de EE. UU., NHANES 1988-1994 hasta 2011-2012. (De Murphy D, McCulloch CE, Lin F, et al. Trends in prevalence of chronic kidney disease in the United States. *Ann Intern Med.* 2016;165:473–481.)

Para minimizar los costes de investigación en salud, los investigadores a menudo dependen de datos autoinformados. El peso y la altura son las variables autoinformadas más comunes. Sin embargo, los autoinformes son propensos a una notificación insuficiente o excesiva. Los datos transversales pueden ayudar a validar y corregir errores en el peso y la estatura autoinformados. Por ejemplo, Jain comparó el peso y la talla autoinformados con datos transversales medidos del estudio NHANES en el periodo 1999-2006. Esta comparación le permitió estimar un factor de corrección, que luego se aplicó a la prevalencia de la obesidad en función del peso y la estatura autoinformados por el sistema de vigilancia de factores de riesgo conductuales (*Behavioral Risk Factor Surveillance System*). Jain estimó que el sesgo del autoinforme del peso y la talla resultó en una prevalencia de obesidad aproximadamente un 5% más baja tanto en hombres como en mujeres[14].

Estudios de casos y controles

Suponga que usted es médico y ha visto algunos pacientes con cierta enfermedad. Observa que muchos de ellos han estado expuestos a un agente en particular—biológico o químico. Usted lanza la hipótesis de que su exposición está relacionada con el riesgo de desarrollar esta enfermedad. ¿Cómo haría para confirmar o refutar su hipótesis?

Consideremos un ejemplo de la vida real:

Durante mucho tiempo se pensó que la hiperacidez es la causa de la enfermedad ulcerosa péptica (EUP). En 1982, los médicos australianos Barry Marshall y Robin Warren descubrieron Helicobacter pylori (H. pylori) en los estómagos de pacientes con EUP, y demostraron que H. pylori es capaz de adaptarse al ambiente ácido del estómago. Sin embargo, sus observaciones no fueron suficiente para establecer la asociación causal entre H. pylori y la EUP. Posteriormente, sugirieron que los antibióticos, no los antiácidos, son el tratamiento efectivo para la EUP, una hipótesis que fue muy criticada en ese momento. No fue hasta 1994 cuando los Institutos Nacionales de la Salud llegaron a una opinión de consenso de expertos basada en la evidencia disponible de que la detección y la erradicación de H. pylori son clave en el tratamiento de la EUP. Los Dres. Marshall y Warren recibieron el Premio Nobel de Fisiología o Medicina en 2005[15].

Para determinar la importancia de las observaciones clínicas en un grupo de casos comunicados por los médicos, se necesita un grupo de comparación (a veces llamado de control o referencia). Las observaciones basadas en series de casos habrían sido intrigantes, pero no sería posible llegar a una conclusión firme sin comparar estas observaciones en los casos con las de una serie de controles que son similares en su mayor parte a los casos pero que no presentan la enfermedad en estudio. La comparación es un componente esencial de la investigación epidemiológica y está bien ejemplificada por el diseño de los estudios de casos y controles.

Diseño de un estudio de casos y controles

La figura 7.9 muestra el diseño de un *estudio de casos y controles.* Para examinar la posible relación de una exposición con una determinada enfermedad, identificamos un grupo de individuos con esa enfermedad (llamados *casos*) y, a efectos de comparación, un grupo de personas sin esa enfermedad (llamados *controles*). A continuación, determinamos qué proporción de los casos fue expuesta y qué proporción no fue expuesta. También determinamos qué proporción de los controles fue expuesta y qué

proporción no fue expuesta. En el ejemplo de niños con cataratas, los casos consistirían en niños con cataratas y los controles consistirían en niños sin cataratas. Para cada niño, sería necesario determinar si la madre estuvo expuesta o no a la rubéola durante su embarazo. Prevemos que si la exposición (rubéola) está relacionada con la enfermedad (cataratas), la prevalencia de antecedentes de exposición entre los casos (niños con cataratas) será mayor que entre los controles (niños sin cataratas). Por lo tanto, en un estudio de casos y controles, si existe una asociación entre la exposición y la enfermedad, la prevalencia de antecedentes de exposición debería ser mayor en las personas que tienen la enfermedad (casos) que en las que no la tienen (controles).

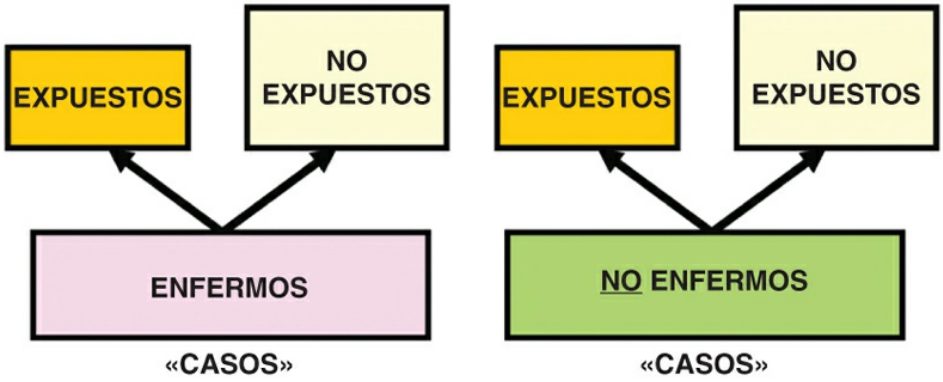


FIG. 7.9 Diseño de un estudio de casos y controles.

La tabla 7.2 presenta un esquema hipotético de cómo se realiza un estudio de casos y controles. Comenzamos seleccionando los casos (con la enfermedad) y los controles (sin la enfermedad), y a continuación determinamos la exposición pasada mediante una entrevista o una revisión de las historias clínicas o los reconocimientos laborales o de los resultados de los análisis químicos o biológicos de sangre, orina o tejidos. Si la exposición es dicotómica, es decir, si la exposición ocurrió (sí) o no (no), es posible establecer una división en cuatro grupos. Existen a casos que fueron expuestos y c casos que no fueron expuestos. De manera similar, hay b controles que fueron expuestos y d controles que no fueron expuestos. Por lo tanto, el número total de casos es $(a + c)$ y el número total de controles es $(b + d)$. Si la exposición se asocia con la enfermedad, esperaríamos que la proporción de los casos que estuvieron expuestos $\left(\frac{a}{a+c}\right)$ fuese mayor que la proporción de los controles que no fueron expuestos $\left(\frac{b}{b+d}\right)$.

Tabla 7.2

Diseño de estudios de casos y controles

	PRIMERO, ELEGIR:	
	Casos (con la enfermedad)	**Controles (sin la enfermedad)**
A continuación, determinar la exposición pasada:		
Expuestos	*a*	*b*
No expuestos	*c*	*d*
Totales	*a* + *c*	*b* + *d*
Proporción de expuestos	$\left(\frac{a}{a+c}\right)$	$\left(\frac{b}{b+d}\right)$

En la tabla 7.3 se muestra un ejemplo hipotético de un estudio de casos y controles. Se está llevando a cabo un estudio de casos y controles sobre si el tabaquismo está relacionado con la AC. Comenzamos con 200 personas con AC (casos) y los comparamos con 400 personas sin AC (controles). Si hubiera una relación entre el antecedente de tabaquismo prolongado y la AC, anticiparíamos que una mayor proporción de los casos con AC que de los controles habrían sido fumadores (expuestos). Digamos que encontramos que de los 200 casos con AC, 112 eran fumadores y 88 no fumadores. De los 400 controles, 176 eran fumadores y 224 no fumadores. Así, el 56% de los casos con AC eran fumadores en comparación con el 44% de los controles. Este cálculo es solo un primer paso. Los cálculos adicionales para determinar si existe o no una asociación de la exposición con la enfermedad se analizarán más adelante. Este capítulo se centra exclusivamente en cuestiones de diseño de los estudios de casos y controles.

Tabla 7.3

Ejemplo hipotético de un estudio de casos y controles sobre AC y tabaquismo

	Casos con AC	**Controles**
Fumadores	112	176
No fumadores	88	224
Totales	200	400
% de fumadores	56	44

AC, arteriopatía coronaria.

De modo adicional, es interesante observar que, si empleamos solo los datos de un estudio de casos y controles, no podemos estimar la prevalencia de la enfermedad. En este ejemplo tuvimos 200 casos y 400 controles, pero esto no implica que la prevalencia sea del 33%, o $\left(\frac{200}{200+400}\right)$. La decisión sobre el número de controles a seleccionar por caso en un estudio de casos y controles está en manos del investigador y no refleja la prevalencia de

la enfermedad en la población. En este ejemplo, el investigador podría haber seleccionado 200 casos y 200 controles (1 control por caso), o 200 casos y 800 controles (4 controles por caso). Debido a que la proporción de la población total del estudio que consiste en casos está determinada por la proporción de controles por caso, y esta proporción está determinada por el investigador, claramente no refleja la verdadera prevalencia de la enfermedad en la población en la que se realiza el estudio.

En este punto, debemos enfatizar que el sello distintivo del estudio de casos y controles es que comienza con personas con la enfermedad (casos) y los compara con personas sin la enfermedad (controles). Esto contrasta con el diseño de un estudio de cohortes que se analizará en el capítulo 8, que comienza con un grupo de personas expuestas y las compara con un grupo no expuesto. Hay quien tiene la impresión errónea de que la diferencia entre los dos tipos de diseño de estudio es que los estudios de cohortes avanzan en el tiempo y los estudios de casos y controles retroceden en el tiempo. Tal distinción no es correcta; de hecho, es desafortunado que el término *retrospectivo* se haya utilizado para los estudios de casos y controles, ya que el término implica incorrectamente que el tiempo del calendario es la característica que distingue los casos y controles del diseño de cohortes. Como se mostrará en un próximo capítulo, un estudio de cohortes retrospectivo también utiliza datos obtenidos en el pasado. Por lo tanto, el tiempo del calendario no es la característica que distingue un estudio de casos y controles de un estudio de cohortes. Lo que distingue los dos diseños de estudio es si el estudio comienza con personas enfermas y no enfermas (estudio de casos y controles) o con personas expuestas y no expuestas (estudio de cohortes). Sir Richard Doll (1912-2005) y Sir Austin Bradford Hill (1897-1991) realizaron uno de los primeros estudios sobre el hábito tabáquico y el cáncer de pulmón. Doll era un epidemiólogo conocido internacionalmente y Hill era un conocido estadístico y epidemiólogo que desarrolló las pautas de «Bradford Hill» para evaluar si una asociación observada es causal[16]. Ambos fueron nombrados caballeros por su trabajo científico en epidemiología y bioestadística.

La tabla 7.4 presenta datos de su estudio frecuentemente citado de 1.357 hombres con cáncer de pulmón y 1.357 controles de acuerdo con el número promedio de cigarrillos fumados al día en los 10 años anteriores a la enfermedad actual[16]. Podemos observar que hay menos fumadores empedernidos entre los controles y muy pocos no fumadores entre los casos de cáncer de pulmón, un hallazgo fuertemente sugerente de una asociación entre el tabaquismo y el cáncer de pulmón. En contraste con el ejemplo anterior, la exposición en este estudio no solo es dicotomizada (expuestos o no expuestos), sino que los datos de exposición están más estratificados en términos de dosis, determinado por el número de cigarrillos fumados habitualmente al día. Debido a que muchas de las exposiciones ambientales que nos preocupan hoy en día no son exposiciones de todo o nada, la posibilidad de realizar un estudio y un análisis que tenga en cuenta la dosis

de la exposición es muy importante.

Tabla 7.4

Distribución de 1.357 pacientes varones con cáncer de pulmón y un grupo de varones control según el número promedio de cigarrillos fumados diariamente durante los 10 años anteriores al inicio de la enfermedad actual

Promedio de cigarrillos diarios	Pacientes con cáncer de pulmón	Grupo control
0	7	61
1-4	55	129
5-14	489	570
15-24	475	431
25-49	293	154
50+	38	12
Total	1.357	1.357

De Doll R, Hill AB. A study of the aetiology of carcinoma of the lung. *BMJ*. 1952;2:1271–1286.

Posibles sesgos en los estudios de casos y controles

Sesgo de selección

Fuentes de casos

En un estudio de casos y controles, los casos se pueden seleccionar de una variedad de fuentes, incluidas los pacientes de hospitales, los pacientes de consultas de médicos o los pacientes de clínicas. Muchas comunidades mantienen registros de pacientes con ciertas enfermedades, como el cáncer, y tales registros pueden servir como fuentes valiosas de casos para tales estudios.

Se deben tener en cuenta varios problemas al seleccionar los casos para un estudio de casos y controles. Si los casos se seleccionan de un solo hospital, cualquier factor de riesgo que se identifique puede ser exclusivo de ese hospital como resultado de patrones de referencia u otros factores, y los resultados pueden no ser generalizables a todos los pacientes con la enfermedad. Por consiguiente, si se van a utilizar casos hospitalizados, es conveniente seleccionar los casos de varios hospitales de la comunidad. Además, si el hospital del cual se extraen los casos es un centro de atención terciaria, que admite selectivamente un gran número de pacientes gravemente enfermos, cualquier factor de riesgo identificado en el estudio puede tratarse de un factor de riesgo solo en personas con formas graves de la enfermedad. En cualquier caso, es esencial que en los estudios de casos y controles, al igual que en los ensayos aleatorizados, los criterios de elegibilidad se especifiquen cuidadosamente por escrito antes de comenzar el

estudio.

Uso de casos incidentes o prevalentes

Una consideración importante en los estudios de casos y controles es si se incluyen los casos incidentes de una enfermedad (casos recién diagnosticados) o los casos prevalentes de la enfermedad (personas que pueden haber tenido la enfermedad durante algún tiempo). El problema con el uso de casos incidentes es que a menudo debemos esperar a que se diagnostiquen nuevos casos; mientras que, si utilizamos los casos prevalentes, que ya se han diagnosticado, a menudo se dispone de un mayor número de casos para el estudio. Sin embargo, a pesar de esta ventaja práctica de usar casos prevalentes, en los estudios de casos y controles sobre la etiología de una enfermedad generalmente es preferible usar casos incidentes de la enfermedad. La razón es que cualquier factor de riesgo que podamos identificar en un estudio con casos prevalentes puede estar más relacionado con la supervivencia de la enfermedad que con el desarrollo de la enfermedad (incidencia). Si, por ejemplo, la mayoría de las personas que desarrollan la enfermedad mueren poco después del diagnóstico, estarán infrarrepresentadas en un estudio que utilice los casos prevalentes, y es más probable que dicho estudio incluya a supervivientes a más largo plazo. Esto constituiría un grupo de casos muy poco representativo, y cualquier factor de riesgo identificado en este grupo no representativo puede no ser una característica general de todos los pacientes con la enfermedad, sino solo de los supervivientes.

Incluso si incluimos solo los casos *incidentes* (pacientes en los que se ha diagnosticado la enfermedad recientemente) en un estudio de casos y controles, por supuesto estaremos excluyendo a los pacientes que pueden haber muerto antes del diagnóstico. No hay una solución fácil para este problema o para otros problemas en la selección de casos, pero es importante que tengamos en cuenta estos aspectos cuando finalmente interpretemos los datos y obtengamos conclusiones del estudio. En ese momento, es fundamental tener en cuenta los posibles sesgos de selección que pueden haber sido introducidos por el diseño del estudio y por la forma en que se realizó el estudio.

Selección de controles

En 1929, Raymond Pearl, profesor de bioestadística en la Universidad Johns Hopkins en Baltimore, Maryland, realizó un estudio para probar la hipótesis de que la tuberculosis protegía contra el cáncer[17]. De 7.500 autopsias consecutivas en el Hospital Johns Hopkins, Pearl identificó 816 casos de cáncer. A continuación, seleccionó un grupo de control de 816 de entre los que se realizaron autopsias en Johns Hopkins y determinó los porcentajes de los casos y de los controles que tuvieron hallazgos de tuberculosis en la autopsia. Las observaciones de Pearl se exponen en la tabla 7.5.

Tabla 7.5

Resumen de los datos del estudio de Pearl sobre cáncer y tuberculosis

	Casos (con cáncer)	Controles (sin cáncer)
N.° total de autopsias	816	816
N.° de autopsias con tuberculosis (%)	54 (6,6)	133 (16,3)

De Pearl R. Cancer and tuberculosis. *Am J Hyg.* 1929;9:97–159.

De las 816 autopsias de pacientes con cáncer, 54 tenían tuberculosis (6,6%), mientras que, de los 816 controles sin cáncer, 133 tenían tuberculosis (16,3%). A partir del hallazgo de que la prevalencia de tuberculosis era considerablemente mayor en el grupo control (sin hallazgos de cáncer) que en el grupo de casos (diagnósticos de cáncer), Pearl concluyó que la tuberculosis tenía un efecto antagónico o protector contra el cáncer.

¿Fue justificada la conclusión de Pearl? La respuesta a esta pregunta depende de la idoneidad de su grupo de control. Si la prevalencia de tuberculosis en los pacientes no cancerosos fuera similar a la de todas las personas sin cáncer, su conclusión sería válida. Pero ese no fue el caso. En el momento del estudio, la tuberculosis era uno de los principales motivos de hospitalización en el Hospital Johns Hopkins. En consecuencia, lo que Pearl había hecho inadvertidamente al elegir el grupo de control libre de cáncer fue seleccionar un grupo en el que muchos de los pacientes habían sido diagnosticados y hospitalizados por tuberculosis. Pearl pensó que la tasa de tuberculosis del grupo de control representaría el nivel de tuberculosis esperado en la población general, pero debido a la forma en que seleccionó los controles, provenían de un grupo que estaba muy ponderado con los pacientes con tuberculosis, que no representaba el nivel en la población general. En efecto, estaba comparando la prevalencia de tuberculosis en un grupo de pacientes con cáncer con la prevalencia de tuberculosis en un grupo de pacientes en los que muchos ya habían sido diagnosticados con tuberculosis. Claramente, su conclusión no estaba justificada sobre la base de estos datos.

¿Cómo podría Pearl haber superado este problema en su estudio? En lugar de comparar a sus pacientes con cáncer con un grupo seleccionado de todos los demás pacientes sometidos a autopsia, podría haber comparado a los pacientes con cáncer con un grupo de pacientes ingresados con otro diagnóstico específico distinto del cáncer (y de la tuberculosis). De hecho, Carlson y Bell18 repitieron el estudio de Pearl, pero compararon a los pacientes que murieron de cáncer con pacientes que fallecieron de una cardiopatía en el Hospital Johns Hopkins. No encontraron diferencias en la prevalencia de tuberculosis en la autopsia entre los dos grupos. (Sin embargo, resulta interesante que a pesar de las limitaciones metodológicas del estudio de Pearl, el bacilo de Calmette-Guérin [BCG], una vacuna contra la tuberculosis, se emplea hoy en día como una forma de inmunoterapia en

varios tipos de cáncer.)

El problema con el estudio de Pearl ejemplifica el desafío de seleccionar controles apropiados como el componente fundamental para sacar conclusiones epidemiológicamente sólidas de los estudios de casos y controles. Sin embargo, sigue siendo uno de los problemas más difíciles que enfrentamos en la realización de estudios epidemiológicos que utilizan el abordaje de casos y controles. El desafío es el siguiente: si realizamos un estudio de casos y controles y encontramos más exposición en los casos que en los controles, nos gustaría ser capaces de concluir que existe una asociación entre la exposición y la enfermedad en cuestión. La forma en que se seleccionan los controles es un determinante importante de si tal conclusión es válida.

Un problema conceptual fundamental relacionado con la selección de controles es si los controles deben ser similares a los casos en todos los aspectos que no sean la enfermedad en cuestión, o si deben ser representativos de todas las personas sin la enfermedad en la población de la cual son seleccionados los casos. Esta pregunta ha suscitado una discusión considerable, pero en realidad, las características de las personas sanas en la población de la cual se seleccionan los casos a menudo no se conocen, porque la población de referencia puede no estar bien definida.

Considere, por ejemplo, un estudio de casos y controles que usa pacientes hospitalizados. Queremos identificar la población de referencia de la que provienen los casos para poder muestrear esta población de referencia para seleccionar controles. Desafortunadamente, por lo general no es fácil o no es posible identificar dicha población de referencia de los pacientes hospitalizados. Los pacientes ingresados en un hospital pueden provenir del vecindario circundante, pueden vivir más lejos en la misma ciudad o, a través de un proceso de referencia, pueden provenir de otra ciudad u otro país. En estas circunstancias, es prácticamente imposible definir una población de referencia específica de la cual surgieron los casos y de la cual podríamos seleccionar controles. Sin embargo, queremos diseñar nuestro estudio de modo que cuando se complete, podamos estar razonablemente seguros de que si encontramos una diferencia en el historial de exposición entre casos y controles, sea probable que no haya ninguna otra diferencia importante entre ellos que pueda limitar la inferencias que podemos derivar.

Fuentes de controles

Los controles pueden seleccionarse de personas no hospitalizadas que viven en la comunidad, de clínicas ambulatorias o de pacientes hospitalizados ingresados por enfermedades distintas de aquellas por las que fueron ingresados los casos.

Uso de personas no hospitalizadas como controles

Los controles no hospitalizados pueden seleccionarse de varias fuentes en la comunidad. Idealmente, podría seleccionarse una muestra probabilística de la

población total, pero en la práctica, esto rara vez es posible. Otras fuentes incluyen listas escolares, listas de votantes registrados y listas de compañías de seguros. Otra opción es seleccionar, como control para cada caso, un residente de un área definida, como el vecindario en el que vive el caso. Tales controles de vecindario han sido utilizados desde hace muchos años. En este enfoque, a los entrevistadores se les indica que identifiquen la casa de un caso como punto de partida, y desde allí pasen por un número específico de casas en una dirección específica y busquen la primera casa que contenga un control elegible. Sin embargo, debido al aumento de los problemas de seguridad en las áreas urbanas de los Estados Unidos, muchas personas ya no abren la puerta a los entrevistadores. Sin embargo, en muchos otros países, particularmente en los países en vías de desarrollo, el método de obtener controles de puerta a puerta para puede ser ideal.

Debido a las dificultades en muchas ciudades de los Estados Unidos para obtener controles de vecindario utilizando el método de puerta a puerta, una alternativa para seleccionar dichos controles es usar métodos de encuesta telefónica. Entre estos se encuentra la marcación de dígitos aleatorios. Debido a que las centrales telefónicas generalmente coinciden con los límites del vecindario (están en el mismo código de área), un número de teléfono de siete dígitos de un caso, de los cuales los tres primeros dígitos son el prefijo de intercambio, se puede usar para seleccionar un número de teléfono de un control, en el que los últimos cuatro dígitos del teléfono se seleccionan aleatoriamente y se utiliza el mismo prefijo de intercambio de tres dígitos. En muchos países en vías de desarrollo, este enfoque no es práctico, ya que es probable que solo las oficinas gubernamentales y los establecimientos comerciales tengan teléfonos. Con la cobertura de teléfono móvil casi universal que ahora existe en casi todo el mundo, el teléfono es un método fascinante de selección de controles. Sin embargo, muchas personas filtran sus llamadas y las tasas de respuesta son tristemente bajas en muchos casos.

Otro enfoque para controlar la selección es utilizar el control de mejor amigo. En este enfoque, a una persona que ha sido seleccionada como un caso se le pide el nombre de su mejor amigo que podría participar en el estudio sabiendo que su mejor amigo también participa. Sin embargo, este método de selección de controles también presenta desventajas. Un control de mejor amigo obtenido de esta manera puede ser similar al caso en edad y en muchas otras características demográficas y sociales. Un problema resultante puede ser que los controles son demasiado similares a los casos con respecto a muchas variables, incluidas las variables que se están investigando en el estudio. A veces, sin embargo, puede ser útil seleccionar como control a un cónyuge o hermano; un hermano puede proporcionar cierto control sobre las diferencias genéticas entre casos y controles.

Uso de pacientes hospitalizados como controles

Los pacientes hospitalizados a menudo se seleccionan como controles debido a que son una «población cautiva», fácilmente accesible y claramente

identificada; por lo tanto, debería ser relativamente más económico realizar un estudio con dichos controles. Sin embargo, como se acaba de exponer, representan una muestra de una población de referencia mal definida que generalmente no se puede caracterizar y, por lo tanto, a la cual no se pueden generalizar los resultados. Además, los pacientes hospitalizados difieren de las personas de la comunidad. Por ejemplo, se conoce que la prevalencia del tabaquismo es mayor en los pacientes hospitalizados que en los residentes de la comunidad; muchos de los diagnósticos por los que los pacientes son ingresados en el hospital están relacionados con fumar.

Dado que generalmente no podemos caracterizar la población de referencia de la que provienen los casos hospitalizados, existe un atractivo conceptual para comparar casos hospitalizados con controles hospitalizados de la misma institución, que probablemente tenderían a provenir de la misma población de referencia[a] (fig. 7.10). Cualquier factor de selección en el sistema de referencia que afectase a la admisión de los casos en un hospital en particular también afectaría a los controles. Sin embargo, los patrones de referencia en el mismo hospital pueden diferir para varios servicios clínicos; tal suposición puede ser cuestionable y, en general, a menudo es imposible saber si se ha cumplido.

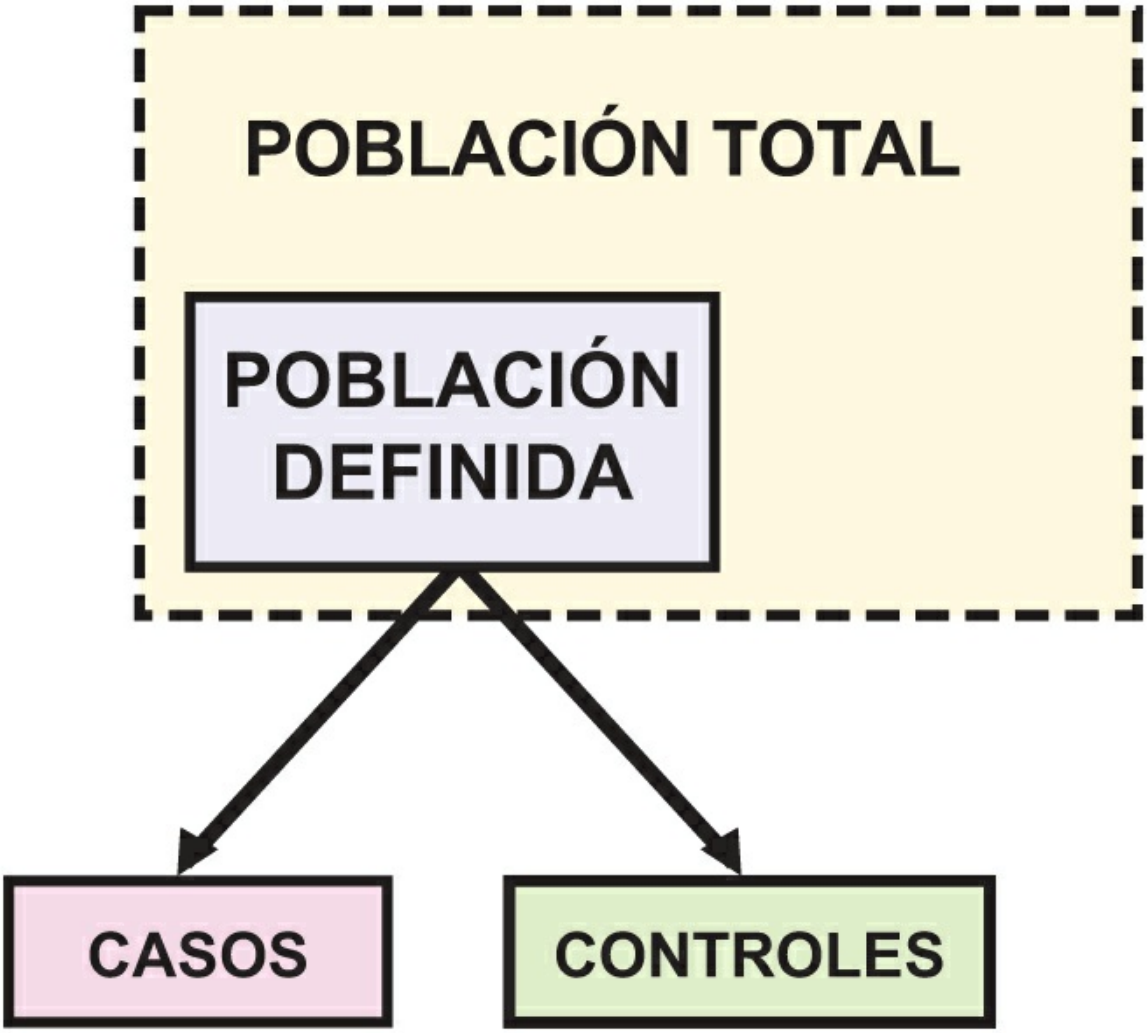


FIG. 7.10 Dado que tanto los casos como los controles del hospital se seleccionan de la población definida, cualquier factor que afecte la admisión de casos en un determinado hospital también afectaría la admisión de los controles del hospital.

Cuando se toma la decisión de usar los controles del hospital, surge la pregunta de si usar una muestra de todos los demás pacientes ingresados en el hospital (que no sean los que tienen el diagnóstico de los casos) o si seleccionar «otro diagnóstico» u «otros diagnósticos» específicos. Si deseamos elegir grupos de diagnóstico específicos, ¿sobre qué base seleccionamos esos grupos y sobre qué base excluimos a otros? El problema es que, aunque es atractivo seleccionar como control hospitalizado un grupo de enfermedades que obviamente no tiene relación con el supuesto factor causal bajo investigación, es poco probable que dichos controles sean representativos de la población de referencia general de no casos. Llevado a su fin lógico, no quedará claro si son los casos o los controles los que difieren de la población general.

El tema de qué grupos de diagnóstico serían elegibles para su uso como controles y cuáles no serían elegibles (y, por lo tanto, excluidos) es muy importante. Supongamos que estamos realizando un estudio de casos y controles sobre cáncer de pulmón y tabaquismo: seleccionamos como casos a

pacientes que han sido hospitalizados con cáncer de pulmón, y como controles seleccionamos a pacientes que han sido hospitalizados con enfisema. ¿Qué problema presentaría este proceder? Como conocemos que existe una estrecha relación entre el tabaquismo y el enfisema, nuestros controles, los pacientes con enfisema, incluirían un alto número de fumadores. En consecuencia, en este estudio no sería fácil detectar cualquier relación entre el tabaquismo y el cáncer de pulmón, porque habríamos seleccionado como controles a un grupo de personas en las que existe una prevalencia de tabaquismo mayor a la esperada de la que existe en la población. Por lo tanto, podríamos querer excluir de nuestro grupo de control a las personas que tienen otros diagnósticos relacionados con el hábito de fumar, como la AC, el cáncer de vejiga, el cáncer de páncreas y el enfisema. Dichas exclusiones podrían generar un grupo de control con una prevalencia de tabaquismo menor a la esperada, y el proceso de exclusión se vuelve demasiado complejo. Una alternativa es no excluir a ningún grupo de la selección como controles en el diseño del estudio, sino analizar los datos del estudio por separado para los diferentes subgrupos de diagnóstico que constituyen el grupo de control. Esto, por supuesto, aumentará la cantidad de controles necesarios y el gasto que acompaña a un tamaño de muestra más grande.

Problemas en la selección de controles

En un estudio clásico publicado en 1981, el reconocido epidemiólogo Brian MacMahon y cols.[19] expuso un estudio de casos y controles sobre el cáncer de páncreas. Los casos fueron pacientes con un diagnóstico histológicamente confirmado de cáncer de páncreas en 11 hospitales de Boston y Rhode Island entre 1974 y 1979. Los controles se seleccionaron de pacientes que fueron hospitalizados al mismo tiempo que los casos; fueron seleccionados de otros pacientes hospitalizados por los médicos que habían hospitalizado los casos. Se excluyeron los no blancos; los mayores de 79 años; los pacientes con enfermedades pancreáticas, del tracto hepatobiliar y relacionadas con el tabaquismo o con el alcohol; y los pacientes con enfermedades cardiovasculares, diabetes, cáncer respiratorio o vesical y úlcera péptica. Sin embargo, los autores no excluyeron a pacientes con otros tipos de enfermedades gastrointestinales, como la hernia diafragmática, el reflujo, la gastritis y la esofagitis.

Un hallazgo en este estudio fue una aparente relación dosis-respuesta entre el consumo de café y el cáncer de páncreas, particularmente en mujeres (tabla 7.6). Cuando se observa una relación de este tipo, es difícil saber si la enfermedad es *causada* por el consumo de café o por algún factor relacionado con dicho consumo. Debido a que fumar es un factor de riesgo conocido para el cáncer de páncreas y porque el consumo de café estaba estrechamente relacionado con fumar cigarrillos en esa época (era raro encontrar un fumador que no bebiera café), ¿MacMahon y otros observaron una asociación entre el consumo de café y el cáncer de páncreas porque el café causa cáncer

de páncreas, o porque el consumo de café está relacionado con fumar cigarrillos, y se conoce que el tabaquismo es un factor de riesgo para el cáncer de páncreas? Reconociendo este problema, los autores analizaron los datos después de la estratificación según los antecedentes de tabaquismo. La relación con el consumo de café se mantuvo tanto para los fumadores actuales como para aquellos que nunca habían fumado (tabla 7.7).

Tabla 7.6

Distribución de casos y controles por hábitos de consumo de café y estimaciones de las razones de riesgo

Sexo	Categoría	CONSUMO DE CAFÉ (TAZAS/DÍA)				Total
		0	1-2	3-4	≥5	
M	N.° de casos	9	94	53	60	216
	N.° de controles	32	119	74	82	307
	Riesgo relativo ajustado[a]	1,0	2,6	2,3	2,6	2,6
	Intervalo de confianza 95%	—	1,2-5,5	1,0-5,3	1,2-5,8	1,2-5,4
F	N.° de casos	11	59	53	28	151
	N.° de controles	56	152	80	48	336
	Riesgo relativo ajustado[a]	1,0	1,6	3,3	3,1	2,3
	Intervalo de confianza 95%	—	0,-8-3,4	1,6-7,0	1,4-7,0	1,2-4,6

[a] Chi cuadrado (extensión de Mantel) con puntuaciones igualmente espaciadas, ajustadas con la edad en décadas: 1,5 para hombres, 13,7 para mujeres. Estimaciones de Mantel-Haenszel de las razones de riesgo, ajustadas por categorías de edad en décadas. En todas las comparaciones, la categoría de referencia fueron los sujetos que nunca bebieron café.

De MacMahon B, Yen S, Trichopoulos D, et al. Coffee and cancer of the pancreas. *N Engl J Med*. 1981;304: 630–633.

Tabla 7.7

Estimaciones del riesgo relativo de cáncer de páncreas asociado con el consumo de café y el tabaquismo

Hábito tabáquico	CONSUMO DE CAFÉ (TAZAS/DíA)			Total[b]
	0	1-2	≥3	
Nunca fumadores	1,0	2,1	3,1	1,0
Ex-fumadores	1,3	4,0	3,0	1,3
Fumadores actuales	1,2	2,2	4,6	1,2 (0,9-1,8)
Total[a]	1,0	1,8 (1,0-3,0)	2,7 (1,6-4,7)	

Los valores entre paréntesis son intervalos de confianza del 95% de las estimaciones ajustadas.

[a] La categoría referente es el grupo que no fuma ni consume café. Las estimaciones se ajustan por sexo y edad en décadas.

[b] Los valores se ajustan para las otras variables, además de la edad y el sexo, y se expresan en relación con la categoría más baja de cada variable.

De MacMahon B, Yen S, Trichopoulos D, et al. Coffee and cancer of the pancreas. *N Engl J Med.* 1981;304:630–633.

Este informe despertó gran interés tanto en las comunidades científicas como en las laicas, particularmente entre los fabricantes de café. Dada la exposición generalizada de los seres humanos al café, si la relación encontrada fuera cierta, tendría importantes implicaciones para la salud pública.

Examinemos el diseño de este estudio. Los casos fueron pacientes blancos con cáncer de páncreas en 11 hospitales de Boston y Rhode Island. Los controles son de particular interés: después de algunas exclusiones, eran pacientes con otras enfermedades que fueron hospitalizados por los mismos médicos que habían ingresado los casos de cáncer de páncreas. Es decir, cuando se identificó un caso, se le preguntó al médico tratante si otro de sus pacientes que fue hospitalizado al mismo tiempo por otra enfermedad podría ser entrevistado para su inclusión como control. Este método inusual de selección de controles tenía una ventaja práctica: uno de los principales obstáculos para obtener la participación de los controles hospitalarios en los estudios de casos y controles es que el permiso para comunicarse con el paciente suele solicitarse al médico que lo atiende. Los médicos a menudo no están motivados para que sus pacientes actúen como controles, porque los pacientes no presentan la enfermedad en la que se centra el estudio. Al pedirles a los médicos que ya habían dado permiso para que participaran pacientes con cáncer de páncreas, se incrementó la probabilidad de que otorgaran permiso para que los pacientes con otras enfermedades participasen como controles.

¿Esa decisión práctica causó algún problema? La pregunta subyacente que los investigadores querían responder era si los pacientes con cáncer de páncreas bebían más café que las personas sin cáncer de páncreas en la misma población (fig. 7.11). Lo que MacMahon y cols. descubrieron fue que el nivel de consumo de café en los casos era mayor que en los controles.

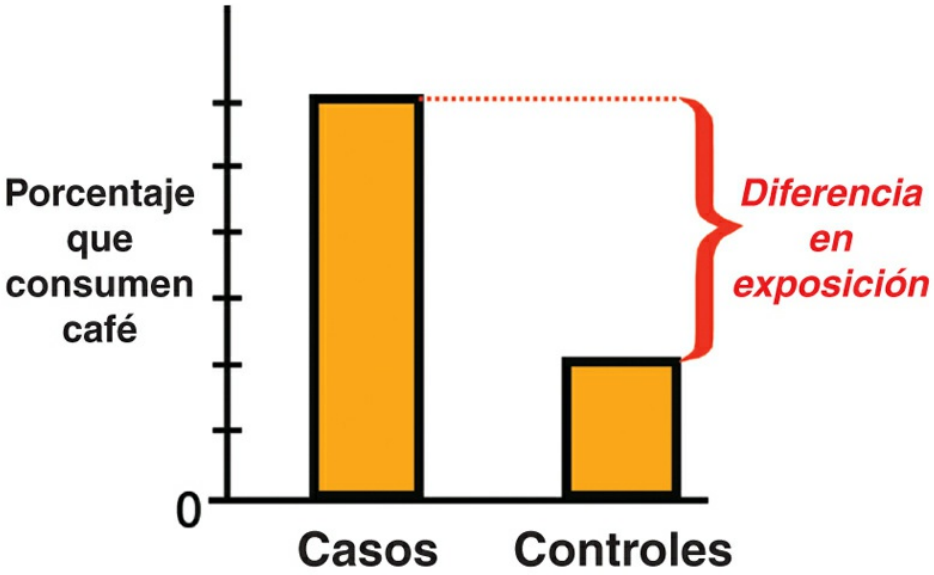


FIG. 7.11 Ejemplo hipotético de un estudio de casos y controles sobre el consumo de café y el cáncer de páncreas: los casos tienen un mayor nivel de consumo de café que los controles.

Los investigadores desearían poder establecer que el nivel de consumo de café observado en los controles es el que se esperaría en la población general sin cáncer de páncreas y que, por lo tanto, los casos demuestran un consumo *excesivo* de café (fig. 7.12A). Pero el problema es el siguiente: ¿qué médicos tienen más probabilidades de ingresar en el hospital a pacientes con cáncer de páncreas? Los gastroenterólogos son a menudo los médicos encargados del ingreso. Muchos de sus otros pacientes hospitalizados (que sirvieron como controles) también tienen problemas gastrointestinales, como esofagitis y gastritis (como se mencionó anteriormente, los pacientes con úlcera péptica se excluyeron del grupo de control). Por lo tanto, en este estudio, es posible que las personas que sirvieron como controles redujeran su consumo de café, ya sea por las instrucciones de un médico o por haberse convencido de que reducir su consumo de café podría aliviar sus síntomas. No podemos asumir que los niveles de consumo de café de los controles sean representativos del nivel de consumo de café esperado en la población general; su tasa de consumo de café puede ser anormalmente baja. Por lo tanto, la diferencia observada en el consumo de café entre los casos de cáncer de páncreas y los controles puede no haber sido necesariamente el resultado de que los casos consuman más café de lo esperado, sino de que el consumo de café de los controles sea menor de lo esperado (v. fig. 7.12B).

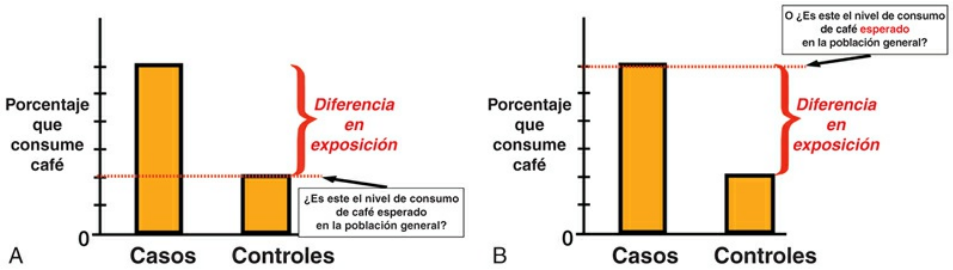


FIG. 7.12 Interpretación de los resultados de un estudio de casos y controles sobre consumo de café y cáncer de páncreas. (A) ¿Es el nivel más bajo de consumo de café en los controles el nivel esperado de consumo de café en la población general? *O* (B) ¿El nivel más alto de consumo de café en los casos es el nivel esperado de consumo de café en la población general?

MacMahon y cols. posteriormente repitieron su análisis, pero separaron los controles con enfermedad gastrointestinal de los controles con otras afecciones. Encontraron que el riesgo asociado con el consumo de café era ciertamente mayor cuando la comparación era con los controles con enfermedad gastrointestinal, pero que la relación entre el consumo de café y el cáncer de páncreas persistió, aunque a un nivel más bajo, incluso cuando la comparación fue con los controles con otras enfermedades. Esto se convirtió en un ejemplo clásico de cómo la selección problemática de controles podría influir en la interpretación de los resultados de un estudio de casos y controles. Varios años después, Hsieh y cols. realizaron un nuevo estudio que intentó replicar estos resultados; pero no apoyó los hallazgos originales[20].

En resumen, cuando se observa una diferencia en la exposición entre los casos y los controles, debemos preguntarnos si el nivel de exposición observado en los controles es realmente el nivel esperado en la población en la que se realizó el estudio o si, quizás, dada la forma de selección, los controles pueden tener un nivel de exposición particularmente alto o bajo que puede no ser representativo del nivel en la población en la que se realizó el estudio.

Sesgo de información

Problemas de recuerdo

Un problema importante en los estudios de casos y controles es el de recordar el historial de exposiciones pasadas. Los problemas de recuerdo son de dos tipos: limitaciones en el recuerdo y el sesgo de recuerdo. El sesgo de recuerdo es la forma principal de sesgo de información en los estudios de casos y controles. El problema del recuerdo no se limita al diseño del estudio de casos y controles. La mayoría de los estudios epidemiológicos indagan sobre los antecedentes vitales y, por lo tanto, están sujetos a sesgos de recuerdo. Las investigaciones han identificado muchas formas de mitigar la cantidad de sesgos asociados con las entrevistas a los participantes sobre acontecimientos de sus vidas. Sin embargo, muchos participantes del estudio se olvidan de las exposiciones u otros eventos, tienden a adelantar los eventos que ocurrieron

hace mucho tiempo («condensación») y pueden ser reticentes a admitir prácticas que podrían considerarse estigmatizantes.

Limitaciones del recuerdo

Gran parte de la información relacionada con la exposición en los estudios de casos y controles a menudo implica la recopilación de datos de los sujetos mediante entrevistas. Debido a que prácticamente todos los seres humanos están limitados en diferente medida en su capacidad para recordar información, las limitaciones en el recuerdo es un tema importante en tales estudios. Un problema relacionado que es algo diferente de las limitaciones en el recuerdo es que las personas entrevistadas pueden simplemente no tener la información solicitada.

Esto se demostró hace años en un estudio histórico realizado por Abraham Lilienfeld y Saxon Graham publicado en 1958[21]. En esa época, se centraba un interés considerable en la observación de que el cáncer de cérvix era altamente inusual en dos grupos de mujeres: mujeres judías y monjas católicas. Esta observación sugirió que un factor de riesgo importante para el cáncer cervical podrían ser las relaciones sexuales con un hombre no circuncidado, y se llevaron a cabo varios estudios para confirmar esta hipótesis. Sin embargo, los autores se mostraron escépticos acerca de la validez de las respuestas con respecto al estado de la circuncisión. Para abordar esta pregunta, preguntaron a un grupo de hombres si habían sido circuncidados o no. Los hombres fueron examinados por un médico. Como se observa en la tabla 7.8, de los 56 hombres que declararon que estaban circuncidados, 19, o el 33.9%, no lo estaban. De los 136 hombres que declararon que no estaban circuncidados, 47, o el 34,6%, sí lo estaban. Estos datos demuestran que los hallazgos de los estudios que utilizan datos de entrevistas no siempre son inequívocos.

Tabla 7.8

Comparación de las declaraciones de los pacientes acerca de la circuncisión con los hallazgos de las exploraciones, Roswell Park Memorial Institute, Buffalo, Nueva York

Hallazgo de la exploración	DECLARACIONES DE LOS PACIENTES ACERCA DE LA CIRCUNCISIÓN			
	SÍ		NO	
	N.°	%	N.°	%
Circunciso	37	66,1	47	34,6
No circunciso	19	33,9	89	65,4
Total	56	100,0	136	100,0

Modificada de Lilienfeld AM, Graham S. Validity of determining circumcision status by questionnaire as related to epidemiologic studies of

cancer of the cervix. *J Natl Cancer Inst*. 1958;21:713–720.

La tabla 7.9 muestra datos más recientes (2002) sobre la relación de los datos autoinformados sobre circuncisión con el estado real de la circuncisión. Estos datos sugieren que los hombres han mejorado su conocimiento y el informe de su estado relacionado con la circuncisión, o las diferencias observadas pueden deberse a que los estudios se han realizado en diferentes países. También puede haber diferencias metodológicas, que podrían justificar los diferentes resultados entre los dos estudios.

Tabla 7.9

Comparación de las declaraciones de los pacientes con los hallazgos de los exámenes médicos sobre el estado de la circuncisión en el estudio de la circuncisión, el virus del papiloma humano del pene y el cáncer cervical

Hallazgos del examen médico	DECLARACIONES DE LOS PACIENTES ACERCA DE LA CIRCUNCISIÓN			
	SÍ		NO	
	N.°	%	N.°	%
Circunciso	282	98,3	37	7,4
No circunciso	5	1,7	466	92,6
Total	287	100,0	503	100,0

Modificado de Castellsague X, Bosch FX, Munoz N, et al. Male circumcision, penile human papillomavirus infection, and cervical cancer in female partners. *N Engl J Med*. 2002;346:1105–1112.

Si la limitación del recuerdo con respecto a la exposición afecta a todos los sujetos en un estudio en la misma medida, independientemente de si son casos o controles, puede producirse una clasificación errónea del estado de exposición. Algunos de los casos o controles que realmente fueron expuestos se clasificarán erróneamente como no expuestos, y algunos que realmente no fueron expuestos se clasificarán erróneamente como expuestos. Para las exposiciones que tienen solo dos categorías (por ejemplo, «sí» frente a «no»), esto lleva a una subestimación del verdadero riesgo de la enfermedad asociada con la exposición (es decir, habrá una tendencia a sesgar los resultados hacia un hallazgo nulo).

Sesgo de recuerdo

Un problema potencial más serio en los estudios de casos y controles es el del sesgo de recuerdo. Supongamos que estamos estudiando la posible relación entre las malformaciones congénitas y las infecciones prenatales. Realizamos un estudio de casos y controles y entrevistamos a madres de niños con

malformaciones congénitas (casos) y madres de niños sin malformaciones (controles). A cada madre se le pregunta acerca de las infecciones que pudo haber tenido durante el embarazo.

Una madre que ha tenido un hijo con un defecto de nacimiento a menudo trata de identificar algún evento inusual que ocurrió durante su embarazo con ese hijo. Quiere saber si la anormalidad fue causada por algo que ella hizo. ¿Por qué sucedió? Tal madre puede incluso recordar un acontecimiento, como una infección respiratoria leve, que la madre de un niño sin un defecto de nacimiento ni siquiera puede haber notado o puede haber olvidado por completo. Este tipo de sesgo se conoce como sesgo de *recuerdo*; Ernst Wynder, un conocido epidemiólogo, también lo llamó «sesgo de rumia».

En el estudio que acabamos de mencionar, supongamos que la verdadera tasa de infección durante el embarazo en madres de bebés con malformaciones y en madres de bebés sanos es del 15%, es decir, no hay diferencias en las tasas de infección. Supongamos que las madres de bebés con malformaciones recuerdan el 60% de las infecciones que tuvieron durante el embarazo, y las madres de bebés sanos recuerdan solo el 10% de las infecciones que tuvieron durante el embarazo. Como se observa en la tabla 7.10, la tasa de infección *aparente* estimada a partir de este estudio de casos y controles mediante entrevistas sería del 9% para las madres de bebés con malformaciones y del 1,5% para las madres de bebés sanos. Por lo tanto, el *recuerdo diferencial* entre los casos y los controles introduce un sesgo de recuerdo en el estudio que podría sugerir artefactualmente una relación entre las malformaciones congénitas y las infecciones prenatales. Aunque un potencial de sesgo de recuerdo es evidente en los estudios de casos y controles, de hecho, pocos ejemplos reales demuestran que el sesgo de recuerdo haya sido un problema importante en los estudios de casos y controles y haya llevado a conclusiones erróneas con respecto a las asociaciones. El pequeño número de ejemplos disponibles podría reflejar la ocurrencia infrecuente de tal sesgo, o el hecho de que los datos necesarios para demostrar claramente la existencia de tal sesgo en un determinado estudio a menudo no están disponibles. Sin embargo, no se puede ignorar el problema potencial, y siempre debe tenerse en cuenta la posibilidad de tal sesgo.

Tabla 7.10

Ejemplo de una asociación artificial resultante de un sesgo de recuerdo: estudio hipotético de infecciones maternas durante el embarazo y malformaciones congénitas

	Casos (con malformaciones congénitas)	Controles (sin malformaciones congénitas)
Si se asume que:		
Incidencia real de infección (%)	15	15
Infecciones recordadas (%)	60	10
El resultado será:		

Tasa de infección según lo determinado por la anamnesis (%)	9,0	1,5

Otras cuestiones en los estudios de casos y controles

Emparejamiento

Una de las principales preocupaciones al realizar un estudio de casos y controles es que los casos y los controles pueden diferir en características o exposiciones distintas de la que fue objeto de estudio. Si se encuentra que han sido expuestos más casos que controles, podemos preguntarnos si la asociación observada podría deberse a diferencias entre los casos y los controles en factores distintos a la exposición que se está estudiando. Por ejemplo, si se encuentran una mayor exposición en los casos que en los controles, y si la mayoría de los casos son de ingresos bajos y la mayoría de los controles son de ingresos altos, no sabríamos si el factor que determina el desarrollo de la enfermedad es la exposición al factor en estudio u otra característica asociada a tener ingresos bajos. Para evitar tal situación, nos gustaría asegurarnos de que la distribución de los casos y controles por estado socioeconómico sea similar, por lo que una diferencia en la exposición probablemente constituya la diferencia crítica, y la presencia o ausencia de enfermedad no es probable que sea atribuible a una diferencia en el estatus socioeconómico.

Un enfoque para abordar este problema en el diseño y la realización del estudio es emparejar los casos y los controles en función de factores que nos pueden interesar, como los ingresos, como en el ejemplo anterior. El emparejamiento se define como el proceso de selección de los controles para que sean similares a los casos en ciertas características, como edad, raza, sexo, estado socioeconómico y ocupación. El emparejamiento puede ser de dos tipos: 1) grupal e 2) individual. Es muy importante distinguir entre los dos tipos, ya que cada uno tiene sus propias implicaciones para el análisis estadístico del estudio de casos y controles, que no se trata en este libro.

Emparejamiento grupal

El *emparejamiento grupal (o emparejamiento de frecuencia)* consiste en seleccionar los controles de tal manera que la *proporción* de controles con cierta característica sea *idéntica* a la proporción de casos con la misma característica. Por lo tanto, si el 25% de los casos están casados, los controles se seleccionarán de manera que el 25% de ese grupo también esté casado. Este tipo de selección generalmente requiere que todos los casos se seleccionen primero. Después de realizar los cálculos de las proporciones de ciertas características en el grupo de casos, se selecciona un grupo de control con las mismas características, en las mismas proporciones. En general, cuando se realiza un emparejamiento grupal, nunca alcanzamos exactamente las

mismas proporciones de la característica clave en los casos y los controles. Cuando el emparejamiento grupal se realiza por edad, por ejemplo, la distribución que es la misma en los casos y controles es la de los grupos de edad (p. ej., de 45 a 49, de 50 a 54); dentro de cada grupo, sin embargo, aún pueden existir diferencias entre los casos y los controles que deben considerarse: por ejemplo, aunque el 10% de los casos y los controles tienen entre 50 y 54 años, puede haber una mayor proporción de casos más cercana a los 54 años que de controles.

Emparejamiento individual

Un segundo tipo de emparejamiento es el emparejamiento individual (o emparejamiento por pares). En este enfoque, para cada caso seleccionado para el estudio, se selecciona un control que es similar al caso en términos de la variable o variables específicas de interés. Por ejemplo, si el primer caso de nuestro estudio es una mujer blanca de 45 años de edad, buscaremos como control a una mujer blanca de 45 años de edad. Si el segundo caso es un hombre negro de 24 años, seleccionaremos un control que también sea un hombre negro de 24 años. Este tipo de selección de controles produce pares de casos y controles coincidentes, es decir, cada caso se empareja individualmente a un control. En nuestro caso hipotético, haríamos coincidir con exactitud los casos por género y raza/etnia, pero podríamos usar un rango de 3 o 5 años para la edad. Por lo tanto, podríamos emparejar a una mujer blanca de 45 años con un control que se tratase de mujer blanca de 42 a 48 años. Las implicaciones de este método de selección de controles para la estimación del riesgo excesivo se analizan en el capítulo 12.

El emparejamiento individual se usa a menudo en los estudios de casos y controles que utilizan controles hospitalarios. La razón de esto es más práctica que conceptual. Supongamos que el sexo y la edad se consideran variables importantes, y se establece que es importante que los casos y los controles sean comparables en términos de estas dos características. En general, no hay una forma práctica de abordar un conjunto de pacientes hospitalizados para seleccionar un grupo con ciertas características de sexo y edad. En su lugar, es más fácil identificar un caso y luego elegir el siguiente ingreso hospitalario que coincida con el caso por sexo y edad. Por lo tanto, el emparejamiento individual es más conveniente en los estudios que utilizan controles hospitalarios.

¿Cuáles son los problemas del emparejamiento? Los problemas del emparejamiento son de dos tipos: prácticos y conceptuales.

Problemas prácticos del emparejamiento

Si se intenta realizar emparejamientos teniendo en cuenta demasiadas características, puede resultar difícil o imposible identificar un control apropiado. Por ejemplo, supongamos que se decide emparejar cada caso por raza, sexo, edad, estado civil, número de hijos, código postal de residencia y ocupación. Si el caso es una mujer negra de 48 años que está casada, tiene

cuatro hijos, vive en el código postal 21209 y trabaja en una planta de procesamiento de fotografías, puede resultar difícil o imposible encontrar un control que sea similar al caso en todas estas características. Por lo tanto, cuantas más variables seleccionemos para emparejar, más difícil será encontrar un control adecuado. El emparejamiento excesivo también se traduce en la incapacidad para analizar estadísticamente las variables utilizadas en la comparación, como abordamos a continuación.

Problemas conceptuales del emparejamiento

Quizás un problema más importante es el conceptual: una vez que hemos emparejado los controles y los casos de acuerdo con una característica dada, no podemos estudiar dicha característica. Por ejemplo, supongamos que estamos interesados en estudiar el estado civil como factor de riesgo del cáncer de mama. Si emparejamos los casos (cáncer de mama) y los controles (sin cáncer de mama) por estado civil, ya no podemos estudiar si el estado civil es o no un factor de riesgo para el cáncer de mama. ¿Por qué no? Porque al realizar el emparejamiento según el estado civil hemos establecido artificialmente una proporción idéntica entre los casos y los controles: si el 35% de los casos están casados, y mediante el emparejamiento hemos creado un grupo de control en el que el 35% también está casado, hemos asegurado artificialmente que la proporción de sujetos casados será idéntica en ambos grupos. Al utilizar la comparación para imponer la comparabilidad para un determinado factor, aseguramos la misma prevalencia de dicho factor en los casos y los controles. Claramente, no podremos preguntar si los casos difieren de los controles en la prevalencia de ese factor. Por lo tanto, en este estudio no querríamos emparejar respecto a la variable del estado civil. De hecho, no deberíamos emparejar respecto a ninguna variable que queramos estudiar en nuestra investigación.

También es importante reconocer que el emparejamiento no planificado puede ocurrir inadvertidamente en los estudios de casos y controles. Por ejemplo, si utilizamos los controles del vecindario, en realidad estamos emparejando en función del estatus socioeconómico, así como de las características culturales y de otro tipo del vecindario. Si utilizamos los controles de mejor amigo, es probable que el caso y su mejor amigo compartan muchas características de estilo de vida, lo que en efecto produce un emparejamiento para estas características. Por ejemplo, en un estudio sobre el uso de anticonceptivos orales y el cáncer de cuello uterino en el que se consideraron los controles de mejor amigo, existía la preocupación de que si el caso usaba anticonceptivos orales bien podría ser que su mejor amiga también pudiera ser usuaria de anticonceptivos orales. El resultado sería un emparejamiento no planificad respecto al uso de anticonceptivos orales, por lo que esta variable ya no podría investigarse en este estudio. Otro ejemplo, y menos sutil, sería emparejar los casos y los controles en función de su lugar de residencia al hacer un estudio de la relación entre la contaminación del aire y la enfermedad respiratoria. El emparejamiento no planificado respecto

a una variable que está estrechamente relacionada con la exposición que se está investigando en el estudio se denomina *sobreemparejamiento*.

Cuando realizamos un estudio de casos y controles, por lo tanto, emparejamos solo respecto a las variables que estamos convencidos de que son factores de riesgo para la enfermedad, por lo que no estamos interesados en investigarlas en este estudio.

Uso de múltiples controles

Al comienzo de este capítulo, mencionamos que el investigador puede determinar cuántos controles se usarán por caso en un estudio de casos y controles y que con frecuencia se emplean múltiples controles para cada caso. El emparejamiento 2:1, 3:1 o 4:1 aumentará el poder estadístico de nuestro estudio. Por lo tanto, muchos estudios de casos y controles tendrán más controles que casos. Estos controles pueden ser: 1) controles del mismo tipo o 2) controles de diferentes tipos, como controles hospitalarios y de vecindarios o controles con diferentes enfermedades.

Controles del mismo tipo

Para aumentar la potencia del estudio se utilizan múltiples controles del mismo tipo, como dos o tres controles por cada caso. En la práctica, un aumento notable en la potencia se obtiene solo hasta una proporción de aproximadamente 4 controles por cada caso. Podríamos preguntarnos: «¿Por qué usar múltiples controles para cada caso? ¿Por qué no mantener la proporción de controles a casos en 1:1 y solo aumentar el número de casos?». La respuesta es que para muchas de las enfermedades relativamente infrecuentes que estudiamos (que se estudian mejor con diseños de casos y controles), puede haber un límite en la cantidad de casos potenciales disponibles para el estudio. Una clínica puede ver solo un cierto número de pacientes con un cáncer determinado o con un cierto trastorno del tejido conectivo cada año. Debido a que el número de casos no se puede aumentar sin extender el estudio en el tiempo para incluir más casos o realizar un estudio multicéntrico colaborativo, a menudo se elige la opción de aumentar el número de controles por caso. Estos controles son del mismo tipo (p. ej., controles de vecindario); solo ha cambiado la relación de controles a casos.

Múltiples controles de diferentes tipos

Por el contrario, podemos optar por utilizar *múltiples controles de diferentes tipos.* Por ejemplo, es posible que nos preocupe que la exposición de los controles hospitalarios utilizados en nuestro estudio no represente la tasa de exposición «esperada» en una población de personas sin enfermedad, es decir, los controles pueden ser un subconjunto altamente seleccionado de personas no enfermas y pueden tener una experiencia de exposición diferente. Anteriormente mencionamos que los pacientes hospitalizados fuman más que las personas que viven en la comunidad, y estamos

preocupados porque no sabemos cuál es el nivel de prevalencia del hábito de fumar en los controles hospitalizados o cómo interpretar una comparación de estas tasas con las de los casos. Para abordar este problema, podemos elegir utilizar un grupo de control adicional, como los controles de vecindario. La esperanza es que los resultados obtenidos cuando se comparan los casos con los controles hospitalarios sean similares a los resultados obtenidos cuando los casos se comparan con los controles del vecindario. Si los resultados difieren, se debe buscar el motivo de la discrepancia. Al usar múltiples controles de diferentes tipos, el investigador idealmente debería decidir qué comparación se considerará el «referente de la verdad» antes de embarcarse en el estudio.

En 1979, Ellen Gold y cols. publicaron un estudio de casos y controles sobre tumores cerebrales en niños[22]. Usaron dos tipos de controles: niños sin cáncer (llamados *controles normales*) y niños con cánceres diferentes a los tumores cerebrales (llamados *controles con cáncer*; fig. 7.13). ¿Cuál fue la razón para emplear estos dos grupos de control?



FIG. 7.13 Grupos de estudio de Gold y cols. para tumores cerebrales en niños. (Datos de Gold EB, Gordis L, Tonascia J, et al. Risk factors for brain tumors in children. *Am J Epidemiol*. 1979;109:309–319.)

Consideremos la pregunta «¿Las madres de niños con tumores cerebrales sufrieron más exposición a radiación prenatal que las madres de controles?"» En la figura 7.14A se exponen algunos resultados posibles.

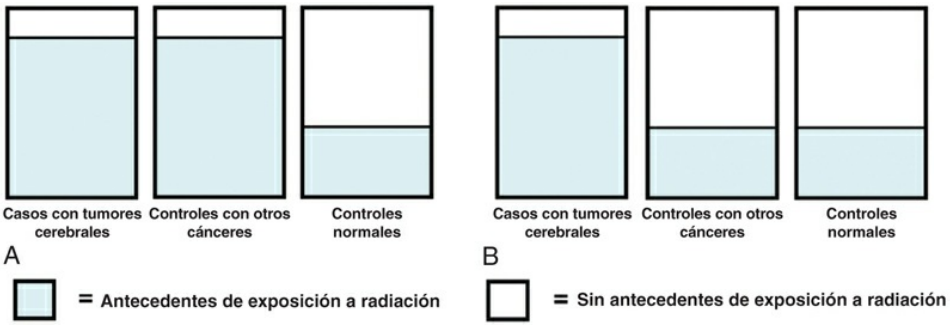


FIG. 7.14 Razón para usar dos grupos de control: (A) La exposición a la radiación es la misma tanto en los casos con tumores cerebrales como en los controles con otros cánceres, pero es mayor en ambos grupos que en los

controles normales: ¿esto podría deberse a un sesgo de recuerdo? (B) La exposición a la radiación en otros controles con cáncer es la misma que en los controles normales, pero es más baja que en los casos con tumores cerebrales: el sesgo de recuerdo es improbable. (Datos de Gold EB, Gordis L, Tonascia J, et al. Risk factors for brain tumors in children. *Am J Epidemiol*. 1979;109:309–319.)

Si la exposición a la radiación de las madres de niños con tumores cerebrales es mayor que la de las madres de los controles normales, y la exposición a la radiación de las madres de niños con otros tipos de cáncer también es mayor que la de las madres de niños normales, ¿cuáles son las explicaciones posibles? Una conclusión podría ser que la radiación prenatal es un factor de riesgo tanto para los tumores cerebrales como para otros cánceres, es decir, su efecto es el de un carcinógeno que no es específico del sitio. Otra explicación que considerar es que los hallazgos podrían haber resultado del sesgo de recuerdo y que las madres de niños con cualquier tipo de cáncer recuerdan mejor la exposición a radiación prenatal que las madres de niños normales.

Considere otros posibles hallazgos, que se muestran en la figura 7.14B. Si las madres de niños con tumores cerebrales tienen un historial de exposición a radiación mayor que las madres de controles normales y las madres de niños con otros tipos de cáncer, los hallazgos podrían sugerir que la radiación prenatal es un carcinógeno específico para el cerebro. Estos hallazgos también reducirían la probabilidad de que el sesgo de recuerdo estuviera jugando un papel, ya que parecería improbable que las madres de niños con tumores cerebrales recordasen la radiación prenatal mejor que las madres de niños con otros tipos de cáncer. Por lo tanto, múltiples controles de diferentes tipos pueden ser valiosos para explorar hipótesis alternativas y para tener en cuenta posibles sesgos potenciales, como el sesgo de recuerdo.

A pesar de las cuestiones planteadas en este capítulo, los estudios de casos y controles son inestimables para explorar la etiología de la enfermedad. Las publicaciones recientes en la literatura demuestran la utilidad del diseño del estudio de casos y controles en la investigación contemporánea.

Kristian Filion y cols. abordaron en Canadá la preocupación de que una clase común de medicamentos antidiabéticos (medicamentos basados en incretina utilizados en la práctica clínica) se asocie con un mayor riesgo de insuficiencia cardiaca[23]. Los informes anteriores resultantes de ensayos clínicos habían sido contradictorios. Los investigadores combinaron los datos de atención médica de cuatro provincias canadienses, Estados Unidos y el Reino Unido, y realizaron un estudio de casos y controles en el que cada paciente que fue hospitalizado por insuficiencia cardiaca fue emparejado con 20 controles. Los criterios de emparejamiento incluyeron la edad, el sexo, el tiempo que llevaban incorporados en el estudio, la duración del tratamiento de la diabetes y el tiempo de seguimiento de los pacientes diabéticos. Casi 30.000 pacientes fueron hospitalizados por insuficiencia cardiaca de casi 1.5 millones de pacientes en total. No se encontró que los medicamentos basados en incretina aumenten la hospitalización por insuficiencia cardiaca en

comparación con los fármacos antidiabéticos orales. Otro ejemplo de la utilidad del estudio de casos y controles lo brindan Su y cols., de la Universidad de Michigan, quienes evaluaron la asociación entre las exposiciones laborales y ambientales con el riesgo de desarrollar esclerosis lateral amiotrófica (ELA, comúnmente conocida como enfermedad de Lou Gehrig, una enfermedad neurológica progresiva que afecta a las neuronas del cerebro y la médula espinal responsables de controlar el movimiento muscular voluntario)[24]. Se identificaron casos en un centro de referencia terciario para ELA entre 2011 y 2014. Los casos consistieron en 156 pacientes con ELA; se seleccionaron 128 controles de voluntarios que respondieron a encuestas en línea. Los controles, que se emparejaron en función de la frecuencia a los casos por edad, sexo y educación, autoinformaron que estaban libres de enfermedad neurodegenerativa y no tenían parientes de primer o segundo grado con ELA. Un cuestionario comprobó las exposiciones ocupacionales y ambientales. Se evaluaron las concentraciones en sangre para 122 contaminantes comunes. En total, se completaron los datos demográficos y de contaminantes en 101 casos y 110 controles. De la historia ocupacional se observó que el servicio militar se asoció con ELA. La exposición a pesticidas autoinformada se asoció con cinco veces mayores probabilidades de ELA. Al estudiar que otros posibles factores podrían estar asociados con la ELA, se identificaron tres exposiciones medidas en la sangre: exposiciones ocupacionales a pesticidas y policlorobifenilos (PCB) en las industrias agrícolas y pesqueras. Los autores concluyeron que los contaminantes ambientales persistentes, medidos en la sangre, se asociaron significativamente con la ELA y sugirieron que reducir la exposición a estos agentes podría reducir la incidencia de la ELA a nivel de la población.

Un último ejemplo de la utilidad del estudio de casos y controles se relaciona con su uso durante un brote de enfermedad. En un estudio que abordó la asociación del síndrome de Guillain-Barré con la infección por el virus del Zika en la Polinesia Francesa en 2013-2014, Cao-Lormeau y cols. observaron que durante el brote de Zika, hubo un aumento en los casos de síndrome de Guillain-Barré, lo que sugiere una posible relación[25]. Cuarenta y dos pacientes ingresados en el principal hospital de referencia en Papeete, Tahití, que cumplían con los criterios diagnósticos de síndrome de Guillain-Barré, fueron emparejados con dos tipos de controles: 1) pacientes sin fiebre atendidos en la instalación, emparejados por edad y residencia (n = 98) y 2) pacientes de la misma edad con Zika agudo sin síntomas neurológicos (n = 70). De los 42 pacientes con síndrome de Guillain-Barré, el 98% (41/42) tenía anticuerpos contra el virus Zika, en comparación con el 56% de los controles. Todos los pacientes en el grupo de control 2 tuvieron una confirmación positiva del virus Zika. Los autores concluyeron que su estudio proporciona evidencia de que la infección por el virus del Zika «causa» síndrome de Guillain-Barré. Esta afirmación parece ir un poco más allá de la evidencia, como veremos en la siguiente sección y se reiterará en los capítulos siguientes.

¿Cuándo está indicado realizar un estudio de casos y controles?

Un estudio de casos y controles es útil como primer paso cuando se busca una causa de un resultado adverso para la salud, como se observa en los ejemplos al comienzo de este capítulo y en los que se acaban de presentar. En una etapa temprana en nuestra búsqueda de una etiología, podemos sospechar cualquiera de varias exposiciones, pero es posible que no tengamos pruebas, y ciertamente ninguna evidencia sólida, para sugerir una asociación de cualquiera de las exposiciones sospechosas con la enfermedad en cuestión. Usando el diseño de casos y controles, comparamos personas con la enfermedad (casos) y personas sin la enfermedad (controles; fig. 7.15A). A continuación, podemos explorar los posibles papeles de una variedad de exposiciones o características para causar la enfermedad (v. fig. 7.15B). Si la exposición está asociada con la enfermedad, esperaríamos que la proporción de casos que han estado expuestos sea mayor que la proporción de controles expuestos (v. fig. 7.15C). Cuando tal asociación se documenta en un estudio de casos y controles, el siguiente paso a menudo es llevar a cabo un estudio de cohortes para dilucidar más la relación. Debido a que los estudios de casos y controles son generalmente menos costosos que los estudios de cohortes y se pueden llevar a cabo más rápidamente, a menudo son el primer paso para determinar si una exposición está relacionada con un mayor riesgo de enfermedad.

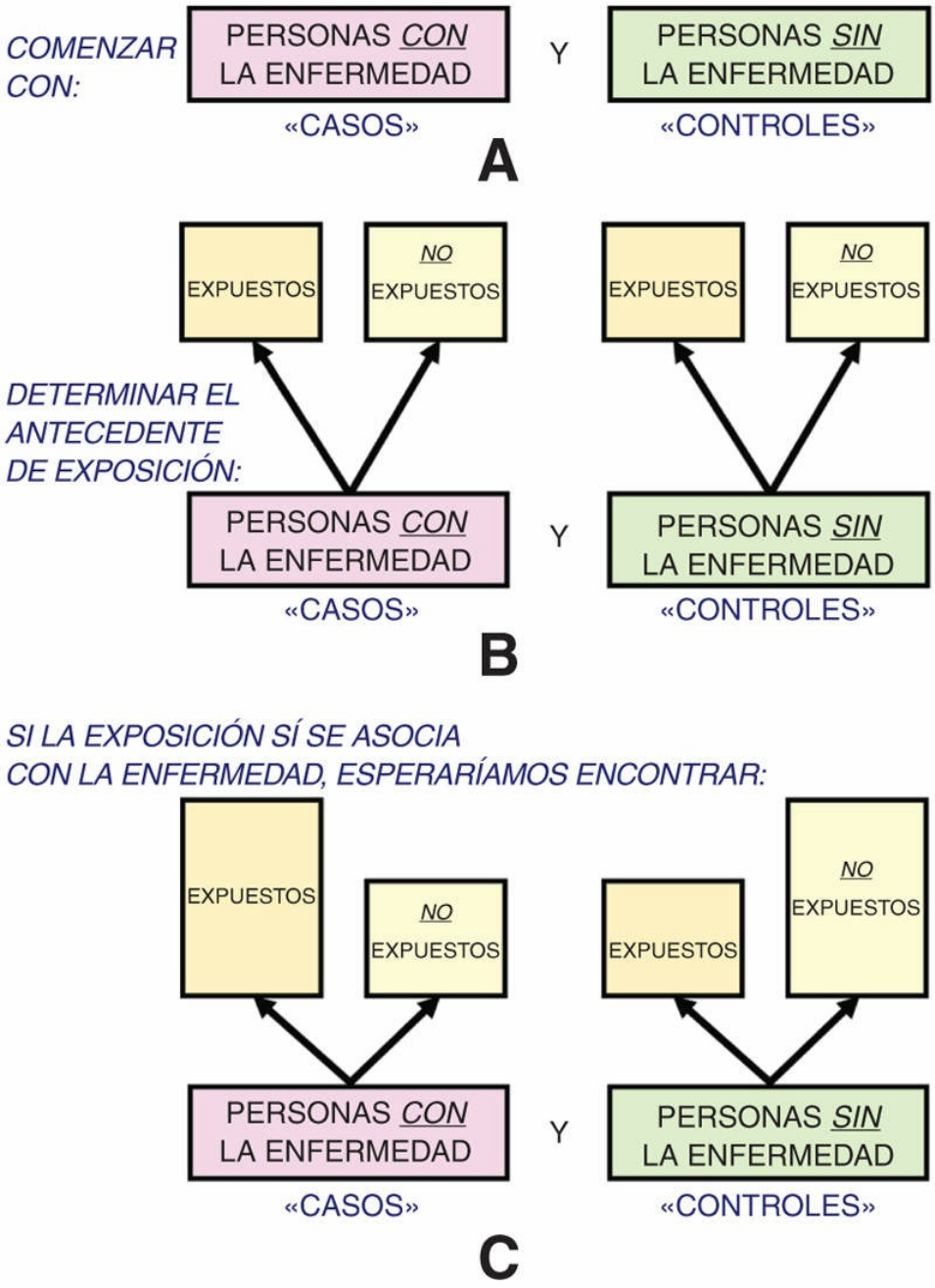


FIG. 7.15 Diseño de un estudio de casos y controles. (A) Comenzar con los casos y los controles. (B) Determinar la exposición pasada en ambos grupos. (C) Resultados esperados si la exposición está asociada con la enfermedad.

Los estudios de casos y controles también son valiosos cuando la enfermedad que se investiga es rara. A menudo es posible identificar casos para estudio a partir de registros de enfermedades, registros hospitalarios u otras fuentes. En contraste, si realizamos un estudio de cohortes para una enfermedad rara, es posible que se necesite una población de estudio extremadamente grande para observar que un número suficiente de individuos en la cohorte desarrollan la enfermedad en cuestión. Además, dependiendo de la duración del intervalo entre la exposición y el desarrollo de la enfermedad, el diseño de cohortes puede implicar muchos años de seguimiento de la cohorte y una considerable dificultad y gasto logístico para

mantener y seguir a la cohorte durante el periodo de estudio.

Diseño de casos cruzados

El *diseño de casos cruzados* se usa principalmente para estudiar la etiología de patologías agudas, como los IM o las muertes por eventos agudos en situaciones en las que la exposición presunta es transitoria y su efecto se produce en poco tiempo. Este tipo de diseño se ha utilizado en el estudio de exposiciones como la contaminación del aire caracterizada por aumentos rápidos y transitorios de partículas. En este tipo de estudio, se identifica un caso (p. ej., una persona que ha sufrido un IM) y se determina el nivel de exposición ambiental, como el nivel de partículas, durante un breve periodo de tiempo que precede al evento (el periodo de riesgo). Este nivel se compara con el nivel de exposición en un periodo de tiempo control que está más alejado del evento. Por lo tanto, cada persona que es un caso sirve como su propio control, y el periodo inmediatamente anterior a su resultado adverso se compara con un periodo de «control» en un momento anterior cuando no se había producido ningún resultado adverso. Es importante destacar que, en este tipo de estudio, existe un emparejamiento inherente por variables que no cambian (p. ej., los factores genéticos) o variables que solo cambian dentro de un periodo razonablemente largo (p. ej., la altura). La pregunta que se plantea es: ¿Hubo alguna diferencia en la exposición entre el periodo de tiempo inmediatamente anterior al resultado y un periodo de tiempo en el pasado más remoto que no fue seguido inmediatamente por ningún efecto adverso para la salud?

Veamos un estudio hipotético de casos cruzados muy pequeño de 4 meses sobre la contaminación del aire y el IM (fig. 7.16A a E).

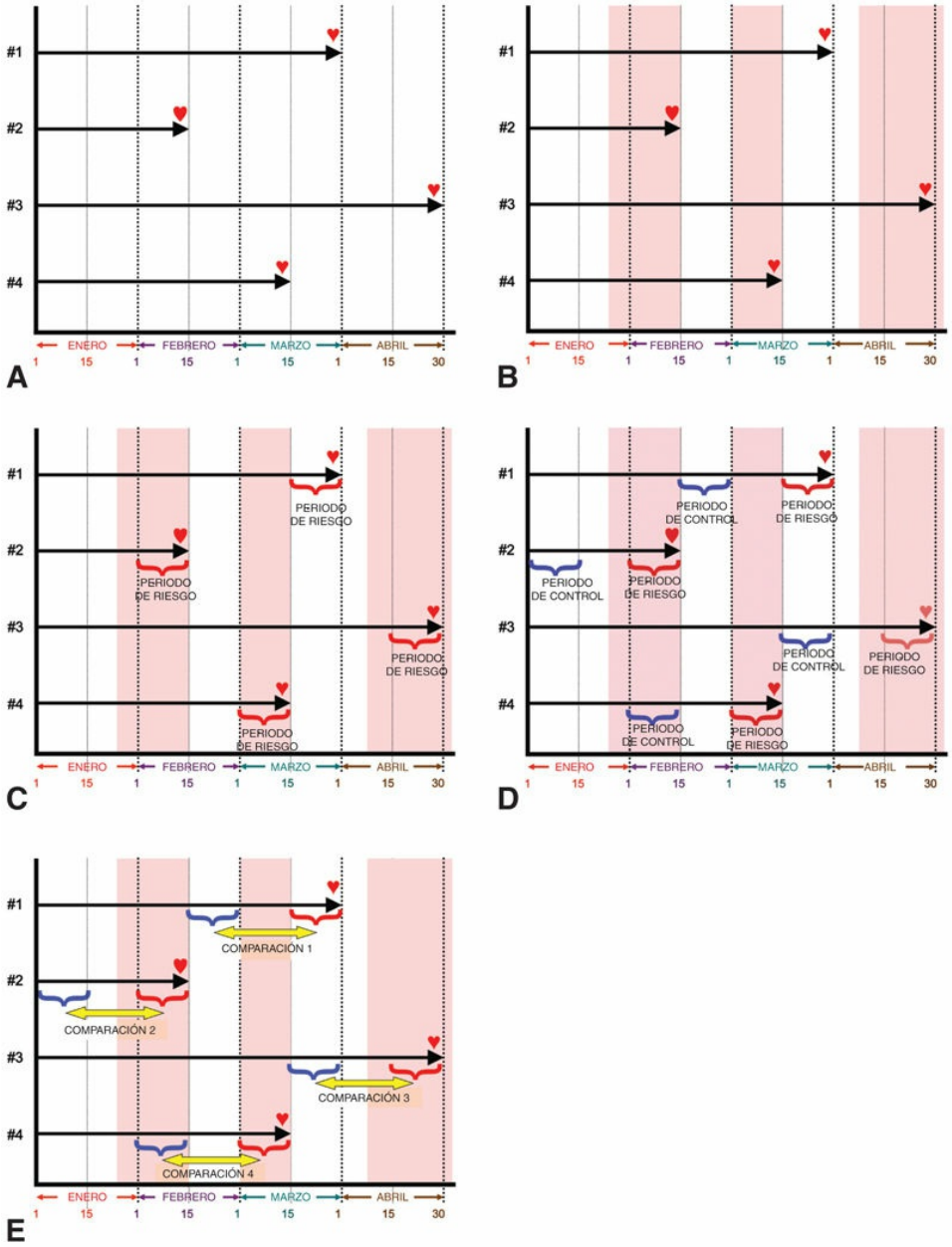


FIG. 7.16 Diseño y hallazgos de un hipotético estudio de casos cruzados de 4 meses sobre la contaminación del aire y el infarto de miocardio (IM; v. discusión en el texto en la pág. 172). (A) Momentos de aparición de casos de IM. (B) Periodos de alta contaminación del aire (mostrados por las *bandas coloreadas*). (C) Definición de los periodos de riesgo *(llaves rojas)*. (D) Definición de los periodos de control *(llaves azules)*. (E) Comparaciones hechas de los niveles de contaminación del aire en los periodos de riesgo y de control para cada caso de IM en el estudio *(flechas amarillas)*.

La figura 7.16A muestra que, durante un periodo de 4 meses, de enero a abril, se identificaron cuatro casos de IM simbolizados por los pequeños corazones rojos en los diagramas. Las líneas de puntos verticales delinean intervalos de 2 semanas durante el periodo de 4 meses. Durante el mismo periodo de 4 meses, se midieron los niveles de contaminación del aire. Se identificaron tres periodos de altos niveles de contaminación del aire en diferentes periodos de tiempo y se muestran en las áreas rosas de la figura 7.16B.

Para cada paciente con IM en este estudio, el periodo de «riesgo» (también denominado «periodo de peligro») se definió como las 2 semanas inmediatamente anteriores al evento. Estos periodos de riesgo están indicados por los corchetes rojos en la figura 7.16C. Si una exposición tiene un efecto a corto plazo sobre el riesgo de un IM, esperaríamos que la exposición se haya producido durante dicho periodo de riesgo de 2 semanas. Sin embargo, el elemento crítico en un diseño de casos cruzados es que para cada sujeto del estudio, comparamos el nivel de exposición en ese periodo de riesgo con un periodo de control (también llamado «periodo de referencia») que es poco probable que sea relevante para la ocurrencia del evento (el IM) porque está demasiado alejado en el tiempo de la ocurrencia. En este ejemplo, el periodo de control seleccionado para cada sujeto es un periodo de 2 semanas que comienza 1 mes antes del periodo de riesgo, y estos periodos de control se indican con corchetes azules en la figura 7.16D. Por lo tanto, como muestran las flechas amarillas en la figura 7.16E, para cada sujeto, estamos comparando el nivel de contaminación del aire en el periodo de riesgo con el nivel de contaminación del aire en el periodo de control. Para demostrar una asociación entre el IM y la contaminación del aire, esperaríamos ver una mayor exposición a altos niveles de contaminación del aire durante el periodo de riesgo que durante el periodo de control.

En este ejemplo, vemos que para el sujeto 1, tanto el periodo de riesgo como el periodo de control se encontraban en momentos de baja contaminación. Para los sujetos 2 y 3, los periodos de riesgo se encontraban en momentos de alta contaminación y los periodos de control en momentos de baja contaminación. Para el individuo 4, tanto los periodos de riesgo como de control se encontraban en momentos de alta contaminación.

Por lo tanto, en el diseño de casos cruzados, cada individuo sirve como su propio control. En este sentido, el diseño de casos cruzados es similar al diseño cruzado planeado presentado en el capítulo 10. En este tipo de diseño, no nos preocupan otras diferencias entre las características de los casos y las de un grupo de controles separado. Este diseño también elimina el costo adicional que se asociaría con la identificación y entrevista de una población de control separada.

Por más atractivo que sea este diseño, quedan preguntas sin responder. Por ejemplo, el diseño de casos cruzados se puede usar para estudiar a las personas con ataques cardiacos respecto a si hubo un episodio de gran pesadumbre o ira durante el periodo inmediatamente anterior al ataque. En el diseño de este estudio, la frecuencia de tales eventos con carga emocional durante ese intervalo de tiempo se compararía, por ejemplo, con la frecuencia de dichos eventos durante un periodo del mes anterior, que no se asoció con ningún evento de salud adverso. La información sobre tales eventos en ambos periodos a menudo se obtiene al entrevistar al sujeto. Sin embargo, surge la pregunta de si podría haber un sesgo de recuerdo, ya que una persona puede recordar un episodio de carga emocional que ocurrió poco antes de un evento coronario, mientras que un episodio comparable un mes

antes, en ausencia de cualquier evento de salud adverso, puede permanecer olvidado. Por lo tanto, el sesgo de recuerdo puede suponer un problema no solo cuando comparamos casos y controles, como se explicó anteriormente en este capítulo, sino también cuando comparamos al mismo individuo en dos periodos de tiempo diferentes. Los trabajos de Maclure y Mittleman proporcionan más información sobre los casos cruzados[26].

Conclusión

Acabamos de revisar los diseños más básicos de los estudios observacionales utilizados en investigaciones epidemiológicas e investigaciones clínicas. Desafortunadamente, en la literatura se utilizan una variedad de términos diferentes para describir diferentes diseños de estudios, y es importante estar familiarizado con ellos. La tabla 7.11 está diseñada para guiarlo a través de la terminología a menudo confusa. El siguiente diseño de estudio es el «estudio de cohortes», que se presenta en el capítulo 8, y se basa en lo que hemos aprendido de los diseños de estudios observacionales iniciales presentados en este capítulo. A continuación, siguen dos capítulos sobre ensayos aleatorizados, que no son estudios observacionales «estrictamente». En estudios los observacionales, el investigador simplemente sigue a aquellos que están enfermos o no, o expuestos y no expuestos. En el diseño del estudio de ensayo aleatorizado, el investigador utiliza un programa de asignación aleatoria para determinar qué participantes están expuestos o no. Por lo tanto, el ensayo aleatorizado es similar a un experimento y también se conoce como «estudio experimental». Sin embargo, se diferencia de los estudios observacionales solo en que la exposición es asignada experimentalmente (de modo aleatorizado) por el investigador del estudio.

Tabla 7.11

Aproximación a la jungla terminológica

Estudio de casos y controles		=		Estudio retrospectivo
Estudio de cohortes	=	Estudio longitudinal	=	Estudio prospectivo
Estudio de cohortes prospectivo	=	Estudio de cohortes concurrente	=	Estudio prospectivo concurrente
Estudio de cohortes retrospectivo	=	Estudio de cohortes histórico	=	Estudio prospectivo no concurrente
Ensayo aleatorizado		=		Estudio experimental
Estudio transversal		=		Estudio de prevalencia

El propósito de todos estos tipos de estudios es identificar asociaciones entre exposiciones y enfermedades. Si se encuentran dichas asociaciones, el siguiente paso es determinar si es probable que las asociaciones sean causales. Estos temas, comenzando con la estimación del riesgo y la determinación de si la exposición a un determinado factor se asocia con un riesgo excesivo de la enfermedad, se tratan más adelante.

Bibliografía

1. Dick GW, Kitchen SF, Haddow AJ. Zika virus. I. Isolations and serological specificity. *Trans R Soc Trop Med Hyg*. 1952;46:509–520.
2. Schuler-Faccini L, Ribeiro EM, Feitosa IM, et al. Possible association between Zika virus infection and microcephaly—Brazil, 2015. *MMWR Morb Mortal Wkly Rep*. 2016;65:59–62.
3. Mlakar J, Korva M, Tul N, et al. Zika virus associated with microcephaly. *N Engl J Med*. 2016;374(10):951–958.
4. Fleming-Dutra KE, Nelson JM, Fischer M, et al. Update: interim guidelines for health care providers caring for infants and children with possible Zika virus infection—United States, February 2016. *MMWR Morb Mortal Wkly Rep*. 2016;65:182–187.
5. Messerli FH. Chocolate consumption, cognitive function, and Nobel laureates. *N Engl J Med*. 2012;367(16):1562–1564.
6. Austin DF, Karp S, Dworsky R, et al. Excess leukemia in cohorts of children born following influenza epidemics. *Am J Epidemiol*. 1975;101:77–83.
7. Rose KM, Suchindran CM, Foraker RE, et al. Neighborhood disparities in incident hospitalized myocardial infarction in four U.S. communities: the ARIC surveillance study. *Ann Epidemiol*. 2009;19(12):867–874.
8. Deelder AM, Miller RL, de Jonge N, et al. Detection of schistosome antigen in mummies. *Lancet*. 1990;335:724–725.
9. El-Sayed NM, Gomatos PJ, et al. Seroprevalence survey of Egyptian tourism workers for hepatitis B virus, hepatitis C virus, HIV and Treponema pallidium infections: association of hepatitis C virus infections with specific regions of Egypt. *Am J Trop Med Hyg*. 1996;55:179–184.
10. Frank C, Mohamed MK, Strickland GT, et al. The role of parenteral antischistosomal therapy in the spread of hepatitis C virus in Egypt. *Lancet*. 2000;355(9207):887–891.
11. Wynder EL, Stellman SD. The "over-exposed" control group. *Am J Epidemiol*. 1992;135:459–461.
12. Farag YM, Guallar E, Zhao D, et al. Vitamin D deficiency is independently associated with greater prevalence of erectile dysfunction: the National Health and Nutrition Examination

Survey (NHANES) 2001-2004. *Atherosclerosis*. 2016;252:61–67.
13. Murphy D, McCulloch CE, Lin F, et al. Trends in prevalence of chronic kidney disease in the United States. *Ann Intern Med*. 2016;165(7):473–481.
14. Jain RB. Regression models to predict corrected weight, height and obesity prevalence from self-reported data: data from BRFSS 1999-2007. *Int J Obes (Lond)*. 2010;34(11):1655–1664.
15. Barry M. "Autobiography." Nobel Foundation; 2005 [fecha de última consulta: 25 de julio de 2017]. http://www.nobelprize.org/nobel_prizes/medicine/laureates/2 bio.html.
16. Hill AB. The environment and disease: association or causation? *Proc R Soc Med*. 1965;58:295–300.
17. Pearl R. Cancer and tuberculosis. *Am J Hyg*. 1929;9:97–159.
18. Carlson HA, Bell ET. Statistical study of occurrence of cancer and tuberculosis in 11,195 postmortem examinations. *J Cancer Res*. 1929;13:126–135.
19. MacMahon B, Yen S, Trichopoulos D, et al. Coffee and cancer of the pancreas. *N Engl J Med*. 1981;304:630–633.
20. Hsieh CC, MacMahon B, Yen S, et al. Coffee and pancreatic cancer (chapter 2) [letter]. *N Engl J Med*. 1986;315:587–589.
21. Lilienfeld AM, Graham S. Validity of determining circumcision status by questionnaire as related to epidemiologic studies of cancer of the cervix. *J Natl Cancer Inst*. 1958;21:713–720.
22. Gold EB, Gordis L, Tonascia J, et al. Risk factors for brain tumors in children. *Am J Epidemiol*. 1979;109:309–319.
23. Filion KB, Azoulay L, Platt RW, et al. A multicenter observational study of incretin-based drugs and heart failure. *N Engl J Med*. 2016;374:1145.
24. Su F-C, Goutman SA, Chernyak S, et al. Association of environmental toxins with amyotrophic lateral sclerosis. *JAMA Neurol*. 2016;73:803.
25. Cao-Lormeau V-M, Blake A, Mons S, et al. Guillain-Barre syndrome outbreak associated with Zika virus infection in French Polynesia: a case-control study. *Lancet*. 1531;397:2016.
26. Maclure M, Mittleman MA. Should we use a case-crossover design? *Annu Rev Public Health*. 2000;21:193–221.

Preguntas de repaso del capítulo 7

1. *Un estudio de casos y controles se caracteriza por todos los siguientes, excepto:*

a. Es relativamente económico en comparación con la mayoría de los diseños de estudios epidemiológicos.

b. Los pacientes con la enfermedad (casos) se comparan con personas sin la enfermedad (controles).

c. Las tasas de incidencia se pueden calcular directamente.

d. La evaluación de la exposición pasada puede estar sesgada.

e. La definición de los casos puede ser difícil.

2. *Se pidió a los residentes de tres aldeas con tres tipos diferentes de suministro de agua que participaran en una encuesta para identificar a los portadores de cólera. Debido a que recientemente se produjeron varias muertes por cólera, prácticamente todas las personas presentes en ese momento se sometieron a un examen. Se calculó y comparó la proporción de residentes en cada aldea que eran portadores. ¿Cuál es la clasificación adecuada de este estudio?*

a. Estudio transversal.

b. Estudio de casos y controles.

c. Estudio de cohortes prospectivo.

d. Estudio de cohortes retrospectivo.

e. Estudio experimental.

3. *¿Cuál de los siguientes es un estudio de casos y controles?*

a. Estudio de tendencias de mortalidad o morbilidad pasadas para permitir estimaciones de la aparición de enfermedades en el futuro.

b. Análisis de investigaciones anteriores en diferentes lugares y en diferentes circunstancias para permitir el establecimiento de hipótesis basadas en el conocimiento acumulado de todos los factores conocidos.

c. Obtención de historias y otra información de un grupo de casos conocidos y de un grupo de comparación para determinar la frecuencia relativa de una característica o exposición en estudio.

d. Estudio de la incidencia de cáncer en hombres que han dejado de fumar.

e. Tanto *a* como *c.*

4. *En un estudio iniciado en 1965, se preguntó a un grupo de 3.000 adultos en Baltimore sobre el consumo de alcohol. La aparición de casos de cáncer entre 1981 y 1995 fue estudiada en este grupo. Esto es un ejemplo de:*

a. Estudio transversal.

b. Estudio de cohortes prospectivo.

c. Estudio de cohortes retrospectivo.

d. Ensayo clínico

e. Estudio de casos y controles.

5. *En un pequeño estudio piloto, 12 mujeres con cáncer de endometrio (cáncer del útero) y 12 mujeres sin enfermedad aparente fueron contactadas y se les preguntó si alguna vez habían usado estrógenos. Cada mujer con cáncer fue emparejada por edad, raza, peso y la cantidad de veces que habían dado a luz con una mujer sin enfermedad. ¿Qué tipo de diseño de estudio es este?*

a. Estudio de cohortes prospectivo.

b. Estudio de cohortes retrospectivo.

c. Estudio de casos y controles.

d. Estudio transversal.

e. Estudio experimental.

6. *Los registros de exámenes físicos de toda la clase de primer año de 1935 en la Universidad de Minnesota fueron examinados en 1977 para ver si la altura y peso registrados en el momento de entrada en la universidad estaban relacionados con el desarrollo de arteriopatía coronaria (AC) en 1986. Este es un ejemplo de:*

a. Estudio transversal.

b. Estudio de casos y controles.

c. Estudio de cohortes prospectivo.

d. Estudio de cohortes retrospectivo.

e. Estudio experimental.

7. *En un estudio de casos y controles, ¿cuál de las siguientes afirmaciones es verdadera?*

a. La proporción de casos con la exposición se compara con la proporción de controles con la exposición.

b. Las tasas de enfermedad se comparan en las personas con el factor de interés y en las personas sin el factor de interés.

c. El investigador puede optar por tener múltiples grupos de comparación.

d. El sesgo de recuerdo es un problema potencial.

e. *a, c* y *d.*

8. *¿En cuál de los siguientes tipos de diseños de estudio un individuo sirve como su propio control?*

a. Estudio de cohortes prospectivo.

b. Estudio de cohortes retrospectivo.

c. Estudio de casos y cohortes.

d. Estudio de casos cruzados.

e. Estudio de casos y controles.

9. *La falacia ecológica se refiere a:*

a. La evaluación de la exposición en grupos grandes en lugar de en muchos grupos pequeños.

b. La evaluación del resultado en grupos grandes en lugar de en muchos grupos pequeños.

c. La asignación de las características de un grupo a cada individuo de ese grupo.

d. El examen de las correlaciones entre la exposición y los resultados en lugar de las tendencias temporales.

e. El fracaso en el examen de las relaciones temporales entre exposiciones y resultados.

10. *Un investigador desea investigar si el consumo de té (evaluado por un biomarcador para el metabolismo del té) aumenta el riesgo de AC. Para responder esta pregunta utiliza un estudio de casos y controles. La AC es rara en jóvenes. ¿Qué dos grupos son los mejores para incluir en el estudio y comparar para este propósito?*

a. Un grupo de casos con AC y un grupo de aquellos que no tienen AC emparejados individualmente a los casos en función del biomarcador del metabolismo del té.

b. Un grupo de casos con AC y un grupo de aquellos que no tienen AC emparejados en función de la frecuencia a los casos por el biomarcador del metabolismo del té.

c. Un grupo de casos con AC y un grupo de aquellos que no desarrollaron AC, emparejados por edad.

d. Una muestra aleatoria de consumidores de té y una muestra aleatoria de no consumidores de té, emparejados por edad.

e. Una muestra aleatoria de consumidores de té y una muestra aleatoria de no consumidores de té, no emparejados por edad.

11. *¿Cuál de las siguientes es una afirmación verdadera con respecto al emparejamiento?*

a. Una vez que hemos emparejado los controles con los casos de acuerdo con una característica dada, solo podemos estudiar dicha característica cuando la prevalencia de la enfermedad es baja.

b. Si se intenta emparejar teniendo en cuenta demasiadas características, puede resultar difícil o imposible ajustar todas las características durante el análisis de datos.

c. El emparejamiento teniendo en cuenta muchas variables puede dificultar la búsqueda de un control adecuado.

d. El emparejamiento individual difiere del emparejamiento en función de la frecuencia porque los controles se seleccionan de los hospitales en lugar de la población general.

e. Ninguna de las anteriores.

---

[a] En la jerga actual, la población de referencia también se conoce como la «base del estudio».

CAPÍTULO 8

Estudios de cohortes

Palabras clave

incidencia
estudio de cohortes concurrente y no concurrente (histórico/retrospectivo)
sesgo de selección
sesgo de información
estudio de casos y controles anidado
estudio de cohortes de casos

Objetivos de aprendizaje

- Describir el diseño de un estudio de cohortes y las opciones para realizar estudios longitudinales.
- Ilustrar el diseño de un estudio de cohortes con dos ejemplos históricos destacados.
- Comentar algunos sesgos que pueden producirse en los estudios de cohortes.

En este capítulo y en los siguientes de la sección 2, la atención se centra en el uso de la epidemiología para dilucidar las relaciones etiológicas o causales. Los dos pasos que subyacen a los diseños de los estudios se comentan en este capítulo y en los capítulos sobre ensayos clínicos. En la figura 8.1 se representan esquemáticamente estos dos pasos conceptuales:

1. En primer lugar, hay que determinar si existe una asociación entre un factor o una característica y el desarrollo de una enfermedad. Esto se puede lograr mediante el estudio de las características de los grupos, de las características de los individuos o de ambos factores.
2. En segundo lugar, se deducen inferencias apropiadas respecto a una posible relación causal a partir de los patrones de asociación que se han encontrado (v. caps. 14 y 15).

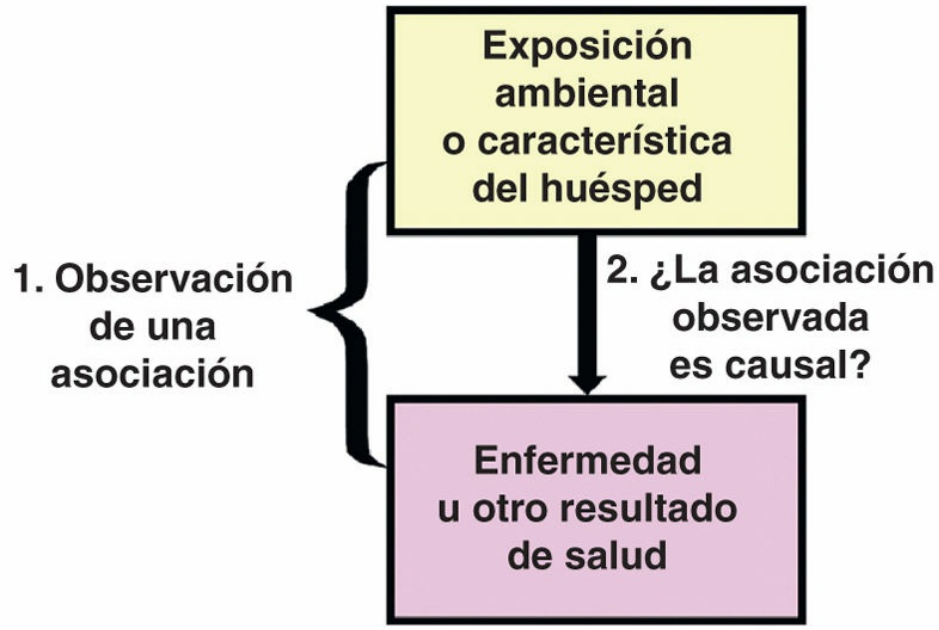


FIG. 8.1 Si se observa una asociación entre una exposición y una enfermedad u otro resultado (1), surge la pregunta de si la asociación es causal (2).

Anteriormente hemos descrito los diseños de los estudios utilizados para el paso 1. En este capítulo se analizan los estudios de cohortes. Los ensayos controlados aleatorizados (experimentos) se comentan en los capítulos 10 y 11. Los estudios de cohortes, junto con los estudios de casos-controles, los transversales, y los ecológicos, a diferencia de los ensayos controlados aleatorizados, se denominan en conjunto estudios *observacionales,* es decir, no hay manipulación experimental involucrada. Investigamos las exposiciones entre los participantes del estudio (en un momento dado o a lo largo del tiempo) y observamos sus resultados en el mismo momento o en algún momento posterior.

Diseño de un estudio de cohortes

En un estudio de cohortes, el investigador selecciona un grupo de individuos expuestos y un grupo de individuos no expuestos y realiza un seguimiento de ambos a lo largo del tiempo para comparar la incidencia de la enfermedad (o la mortalidad por enfermedad) en ambos grupos (fig. 8.2). El diseño puede incluir más de dos grupos (como no expuestos, nivel de exposición baja o alta), aunque aquí solo se muestran dos grupos para fines esquemáticos.



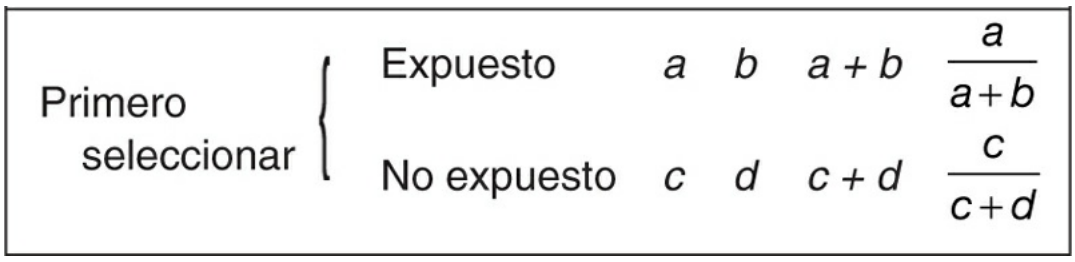
FIG. 8.2 Diseño de un estudio de cohortes.

Si existe una asociación positiva entre la exposición y la enfermedad, sería de esperar que la proporción de personas del grupo expuesto que desarrollan la enfermedad (incidencia en el grupo expuesto) fuese mayor que la proporción de personas del grupo no expuesto que desarrollan la enfermedad (incidencia en el grupo no expuesto).

Los cálculos correspondientes se muestran en la tabla 8.1. Se comienza con un grupo expuesto y un grupo no expuesto. De las $(a + b)$ personas expuestas, la enfermedad se desarrolla en a pero no en b. Por tanto, la incidencia de la enfermedad entre las personas expuestas es $\left(\frac{a}{a+b}\right)$. De forma similar, en las $(c + d)$ personas no expuestas del estudio, la enfermedad se desarrolla en c pero no en d, por lo que la incidencia de la enfermedad entre los no expuestos es $\left(\frac{c}{c+d}\right)$.

Tabla 8.1

Diseño de un estudio de cohortes



El uso de estos cálculos se aprecia en un ejemplo hipotético de un estudio de cohortes que se muestra en la tabla 8.2. En este estudio de cohortes, la asociación del tabaquismo con la arteriopatía coronaria (AC) se investiga seleccionando para el estudio un grupo de 3.000 fumadores (expuestos) y un grupo de 5.000 no fumadores (no expuestos) que no presentan cardiopatía al inicio del estudio. En ambos grupos se realiza el seguimiento de la aparición de AC y se compara la *incidencia* de AC en ambos. La AC se desarrolla en 84 de los fumadores y en 87 de los no fumadores. El resultado es una incidencia de AC de 28,0/1.000 en los fumadores y 17,4/1.000 en los no fumadores.

Tabla 8.2

Resultados de un hipotético estudio de cohortes sobre el tabaquismo y la arteriopatía coronaria (AC)

		DESPUÉS, SEGUIMIENTO PARA VER SI			Incidencia por 1.000 por año
		Se desarrolla AC	No se desarrolla AC	Total	
Primero seleccionar	Fumadores de cigarrillos	84	2.916	3.000	28,0
	No fumadores de cigarrillos	87	4.913	5.000	17,4

Se debe tener en cuenta que, debido a que se están identificando casos *nuevos* (incidentes) de la enfermedad a medida que ocurren, se puede determinar si existe una relación temporal entre la exposición y la enfermedad (es decir, si la exposición precedió a la aparición de la enfermedad). Es evidente que esta relación temporal se debe establecer si hay que considerar que la exposición es una posible causa de la enfermedad en cuestión.

Selección de las poblaciones de estudio

La característica esencial en el diseño de los estudios de cohortes es la comparación de los resultados en un grupo expuesto y en un grupo no expuesto (o en un grupo con una característica determinada y en otro sin esa característica; p. ej., participantes más jóvenes o de edad más avanzada). Hay dos formas básicas para generar tales grupos:

1. Se puede crear una población de estudio mediante la selección de grupos para la inclusión en este, basándose en si han sido expuestos o no (p. ej., cohortes que presentan una exposición laboral comparadas con residentes en la comunidad de edad similar que no poseen la misma ocupación laboral) (fig. 8.3).
2. Se puede seleccionar una población definida antes de que cualquiera de sus miembros se exponga o antes de identificar sus exposiciones. Se podría seleccionar una población basándose en algún factor no relacionado con la exposición (p. ej., la comunidad de residencia) (fig. 8.4) y realizar la anamnesis, o llevar a cabo análisis de sangre o de otro tipo, en toda la población. Utilizando los resultados de la anamnesis o de las pruebas analíticas, se puede separar a la población en los grupos *expuesto* y *no expuesto* (o en aquellos que tienen ciertas características biológicas y los que no), como se hizo en el estudio de Framingham, que se describe más adelante en este capítulo.

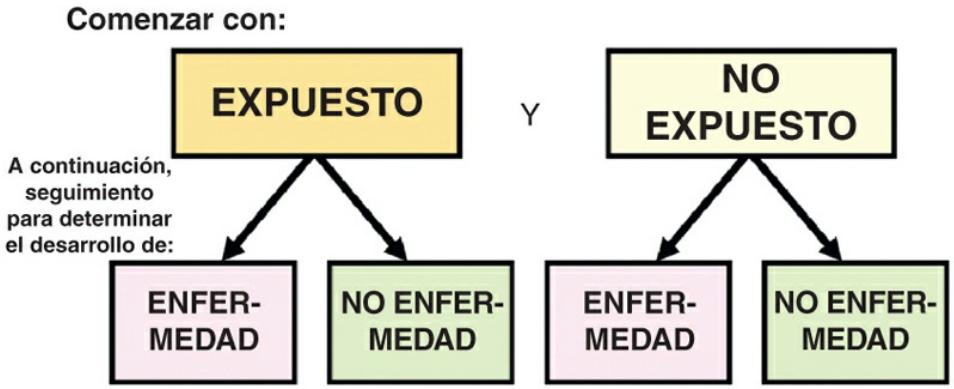


FIG. 8.3 Diseño de un estudio de cohortes comenzando con los grupos expuesto y no expuesto.



FIG. 8.4 Diseño de un estudio de cohortes comenzando con una población definida.

Los estudios de cohortes, en los que se espera que se produzca un resultado en una población, a menudo requieren un periodo de seguimiento prolongado, que dura hasta se han producido bastantes fenómenos (resultados). Cuando se emplea la segunda estrategia (en la que se identifica una población para el estudio basándose en alguna característica no relacionada con la exposición en cuestión), la exposición de interés puede que no tenga lugar durante un cierto tiempo, incluso durante muchos años después de que la población se haya definido. En consecuencia, la duración del seguimiento requerido es aún mayor con la segunda estrategia que con la primera. Hay que tener en cuenta que, con cualquiera de las estrategias, el diseño del estudio de cohortes es fundamentalmente el mismo: *se comparan personas expuestas y no expuestas.* Esta comparación es el sello distintivo del diseño de cohortes.

Tipos de estudios de cohortes

Un problema fundamental con el diseño de cohortes que se acaba de describir es que la población de estudio a menudo debe seguirse durante un periodo prolongado para determinar si se ha producido el resultado de interés. Tomemos como ejemplo un estudio hipotético de la relación del tabaquismo con el cáncer de pulmón. Se identifica una población de estudiantes de primaria y se siguen; diez años después, cuando son adolescentes, se identifican los que fuman y los que no lo hacen. Después, se continúa el seguimiento de ambos grupos (fumadores y no fumadores) para ver quién desarrolla cáncer de pulmón y quién no. Pongamos por caso que el estudio comienza en 2012 (fig. 8.5) y supongamos que muchos niños que se convertirán en fumadores lo harán en el plazo de 10 años. Por tanto, el estatus de la exposición (fumador o no fumador) se determinará 10 años más tarde, en 2022. Para los fines de este ejemplo, se supondrá que el periodo de latencia medio desde que se empieza a fumar hasta que se desarrolla la enfermedad pulmonar es de 20 años. Por tanto, el desarrollo de cáncer de pulmón se determinará, de media, 20 años después, en 2042.



FIG. 8.5 Cronología para un hipotético estudio de cohortes prospectivo iniciado en 2012.

Este tipo de diseño del estudio se denomina *estudio de cohortes prospectivo* (también denominado por otros *estudio de cohortes concurrente* o *longitudinal).* Es *concurrente* (sucede o se hace a la vez) porque el investigador identifica la población original al comienzo del estudio y, en efecto, sigue a los sujetos al mismo tiempo a lo largo del tiempo hasta el punto en el que la enfermedad se desarrolla o no se desarrolla.

Esta estrategia tiene una serie de problemas. Una dificultad es que, tal y como se acaba de describir, el estudio requerirá al menos 30 años para completarse, lo que puede conllevar varios problemas. Si se tiene la suerte de obtener una beca de investigación, la financiación suele limitarse a un máximo de tan solo 5 años. Además, con un estudio de esta duración, existe el riesgo de que los sujetos de estudio sobrevivan al investigador o de que el investigador muera antes del final del estudio. Teniendo en cuenta estas cuestiones, el estudio de cohortes prospectivo a menudo resulta poco atractivo para los investigadores que están pensando en nuevos temas que evaluar.

Hay que dilucidar si estos problemas significan que el diseño de cohortes no es práctico y si hay alguna manera de acortar el periodo de tiempo necesario para llevar a cabo un estudio de cohortes. A continuación, se considerará una estrategia alternativa usando el diseño de cohortes (fig. 8.6). Supongamos que de nuevo se comienza el estudio en 2012, pero ahora se dispone en la comunidad de una antigua lista de los escolares de primaria elaborada en 1982, y además se les había encuestado con respecto a su hábito de fumar en el instituto, en 1992. Gracias al uso de estos recursos de datos en 2012, se puede empezar a identificar qué personas de esta población han desarrollado el cáncer de pulmón y cuáles no. Este tipo de estudio se denomina *estudio de cohortes retrospectivo* o *estudio de cohortes histórico* (o también *estudio prospectivo no concurrente).* Sin embargo, se debe tener en cuenta que el diseño del estudio no difiere del diseño de cohortes prospectivo (todavía se está comparando un grupo expuesto con uno no expuesto); lo que se ha hecho en el diseño de cohortes retrospectivo ha sido utilizar datos históricos del pasado para poder acortar (reducir) el marco temporal para el estudio y obtener los resultados antes. Ya no es un diseño prospectivo, porque se está comenzando el estudio con una población preexistente para reducir su duración. Sin embargo, como se muestra en la figura 8.7, los *diseños, tanto para el estudio de cohortes prospectivo como para el estudio de cohortes retrospectivo* o *histórico, son idénticos: se comparan las poblaciones expuesta y no expuesta.* La única diferencia entre ellos es el tiempo. En un *diseño de cohortes prospectivo,* la exposición y la no exposición se determinan a medida que ocurren durante el estudio, y los grupos se siguen a continuación durante varios años en el futuro y se mide la incidencia. En un *diseño de cohortes retrospectivo,* la exposición se determina a partir de los registros anteriores y el resultado (desarrollo o no desarrollo de la enfermedad) se determina en el momento de iniciar el estudio.

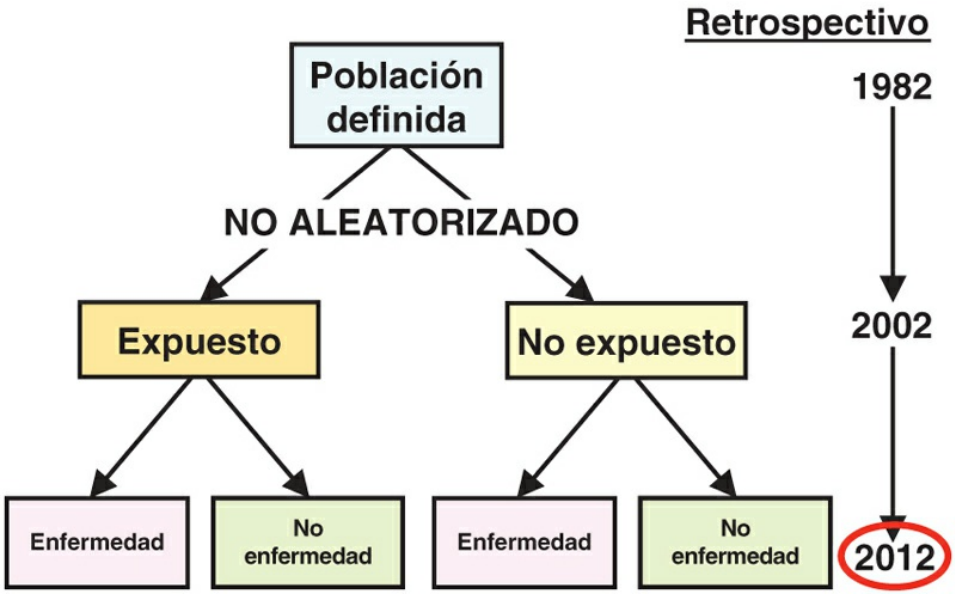


FIG. 8.6 Cronología para un hipotético estudio de cohortes retrospectivo iniciado en 2012.

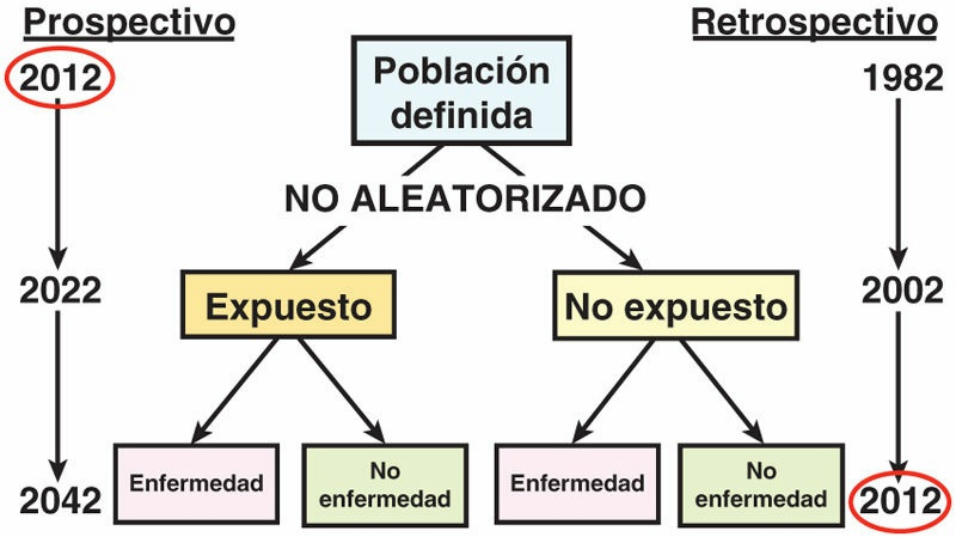


FIG. 8.7 Cronología para un hipotético estudio de cohortes prospectivo y un hipotético estudio de cohortes retrospectivo iniciados en 2012.

También es posible llevar a cabo un estudio que sea una combinación de un diseño de cohortes prospectivo y un diseño de cohortes retrospectivo. Con esta estrategia, la exposición se determina a partir de registros objetivos en el pasado (como en un estudio de cohortes histórico), y el seguimiento y la medición de resultados continúan en el futuro.

Ejemplos de estudios de cohortes

Ejemplo 1: estudio Framingham

El estudio Framingham de enfermedades cardiovasculares, que comenzó en 1948[1], es uno de los primeros estudios de cohortes, de los más importantes y mejor conocidos. Framingham es una ciudad de Massachusetts, aproximadamente a unos 32 kilómetros al oeste de Boston. Se pensó que las características de su población (algo menos de 30.000 habitantes) serían apropiadas para un estudio de este tipo y facilitarían el seguimiento de los participantes, ya que la emigración era baja (es decir, la población era estable).

Los residentes se consideraron elegibles si tenían entre 30 y 62 años de edad al inicio del estudio. La justificación para usar este rango de edad fue que es poco probable que las personas menores de 30 años manifiesten los criterios de valoración cardiovasculares que se evalúan durante el periodo de seguimiento propuesto de 20 años. Muchas personas mayores de 62 años ya tienen una enfermedad coronaria establecida, por lo que no merecería la pena estudiar la incidencia de enfermedad coronaria en las personas de este grupo de edad.

Los investigadores buscaron un tamaño muestral de 5.000 personas. En la tabla 8.3 se muestra cómo se obtuvo la población final del estudio. Constaba de 5.127 varones y mujeres de entre 30 y 62 años de edad en el momento de la inclusión en el estudio, sin enfermedad cardiovascular en dicho momento. En este estudio se definieron muchas «exposiciones» propuestas, como la edad y el sexo, el tabaquismo, el peso, la presión arterial, la concentración de colesterol, el nivel de actividad física y otros factores.

Tabla 8.3

Constitución de la población del estudio Framingham

	N.° de varones	N.° de mujeres	Total
Muestra aleatoria	3.074	3.433	6.507
Personas que respondieron	2.024	2.445	4.469
Voluntarios	312	428	740
Personas que respondieron sin AC	1.975	2.418	4.393
Voluntarios sin AC	307	427	734
Total de personas sin AC grupo del estudio Framingham	2.282	2.845	5.127

AC, arteriopatía coronaria.

De Dawber TR, Kannel WB, Lyell LP. An approach to longitudinal studies in a community: The Framingham Study. *Ann NY Acad Sci* 1993;107:539–556.

Los nuevos episodios coronarios (incidencia) se identificaron evaluando a la población del estudio cada 2 años y controlando a diario los ingresos en el único hospital de Framingham.

El estudio fue diseñado para comprobar las siguientes hipótesis:

- La incidencia de AC aumenta con la edad. Se produce antes y con más frecuencia en los varones.
- Las personas con hipertensión desarrollan AC a un ritmo mayor que las que son normotensas.
- La hipercolesterolemia se asocia con un riesgo mayor de AC.
- El tabaquismo y el consumo habitual de alcohol se asocian con una mayor incidencia de AC.
- El aumento de la actividad física se asocia con una disminución del desarrollo de AC.
- El aumento del peso corporal predispone a una persona a desarrollar AC.
- Los pacientes con diabetes mellitus tienen una mayor incidencia de AC.

Cuando se analiza esta lista en la actualidad, es posible preguntarse por qué este tipo de relaciones tan obvias y bien conocidas deberían haberse evaluado en un estudio tan extenso. No debe olvidarse el peligro de este enfoque «retrospectivo»; es sobre todo *gracias* al estudio Framingham, un estudio de cohortes clásico que hizo contribuciones fundamentales a nuestra comprensión de la epidemiología de las enfermedades cardiovasculares, por lo que estas relaciones son bien conocidas en la actualidad.

En este estudio se utilizó el segundo método descrito anteriormente en este capítulo para seleccionar una población para un estudio de cohortes: se seleccionó una población definida en función de la ubicación de la residencia o de otros factores no relacionados con la exposición o exposiciones en cuestión. Después, la población se observó a lo largo del tiempo para determinar qué personas desarrollaron o ya tenían la «exposición o exposiciones» de interés y, más adelante, para determinar qué participantes del estudio desarrollaron el resultado o los resultados cardiovasculares de interés. Esta estrategia proporcionó una ventaja importante: permitió a los investigadores estudiar múltiples «exposiciones», como la hipertensión, el tabaquismo, la obesidad, los niveles de colesterol y otros factores, así como las complejas interacciones entre las exposiciones, mediante el uso de técnicas multifactoriales. Por tanto, aunque un estudio de cohortes que comienza con un grupo expuesto y otro no expuesto se centra a menudo en solo una exposición específica, un estudio de cohortes que se inicia con una población definida puede explorar los papeles de muchas exposiciones sobre el resultado medido por el estudio.

Ejemplo 2: incidencia de cáncer de mama y deficiencia de progesterona

Se sabe desde hace mucho tiempo que el cáncer de mama es más frecuente en mujeres que son mayores en el momento de su primer embarazo. Esta observación suscita una pregunta difícil de responder: ¿la asociación entre la edad avanzada en el momento del primer embarazo y el mayor riesgo de cáncer de mama se relacionan con el hallazgo de que un primer embarazo precoz protege contra el cáncer de mama (y, por tanto, esa protección no existe en las mujeres que tienen un embarazo más tardío o ningún embarazo), o tanto un primer embarazo tardío como el mayor riesgo de cáncer de mama se deben a un tercer factor, como una anomalía hormonal subyacente?

Es difícil disociar estas dos interpretaciones. Sin embargo, en 1978, Linda Cowan y cols.[2] realizaron un estudio diseñado para determinar cuál de estas dos hipótesis, era probable que fuese la correcta (fig. 8.8). Los investigadores identificaron una población de mujeres que eran pacientes de la clínica de infertilidad del Johns Hopkins Hospital en Baltimore, Maryland, de 1945 a 1965. Debido a que eran pacientes de esta clínica, todas las mujeres, por definición, tenían una edad tardía en el momento del primer embarazo. En el transcurso de sus evaluaciones diagnósticas se elaboraron unos perfiles hormonales detallados para cada mujer. Por tanto, los investigadores fueron capaces de separar a las mujeres que tenían una anomalía hormonal subyacente, como la deficiencia de progesterona (expuestas), de las que no tenían dicha anomalía hormonal (no expuestas) que presentaban otra causa de infertilidad, como un problema de permeabilidad tubárica o un recuento de espermatozoides bajo del marido. A continuación, los dos grupos de mujeres se sometieron a seguimiento para detectar el desarrollo de cáncer de mama con posterioridad.

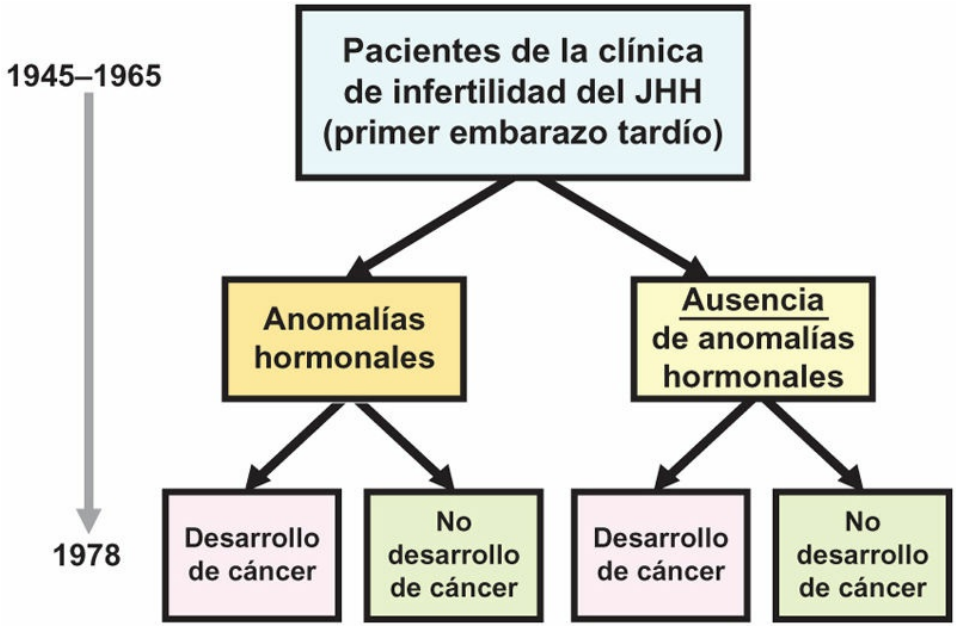


FIG. 8.8 Diseño del estudio de cohortes retrospectivo de Cowan del cáncer de mama. *JHH,* Johns Hopkins Hospital. (Datos de Cowan LD, Gordis L, Tonascia JA, Jones GS. Breast cancer incidence in women with progesterone deficiency. *Am J Epidemiol* 1981;114:209–217.)

¿Cómo podrían los resultados del diseño de este estudio aclarar la relación entre la edad tardía en el momento del primer embarazo y el mayor riesgo de cáncer de mama? Si la hipótesis para la asociación de una edad tardía en el primer embarazo y el mayor riesgo de cáncer de mama fuese que un primer embarazo precoz protege contra el cáncer de mama, no sería de esperar que existiese ninguna diferencia en cuanto a la incidencia de cáncer de mama entre las mujeres que tienen una anomalía hormonal y las que no la tienen (y ninguna de las mujeres habría tenido un primer embarazo temprano). Sin embargo, si la hipótesis del mayor riesgo de cáncer de mama es que la anomalía hormonal subyacente predispone a estas mujeres a desarrollar un cáncer de mama, sería de esperar encontrar una mayor incidencia de cáncer de mama en las mujeres con la anomalía hormonal que en aquellas sin dicha anomalía.

En el estudio se observó que, cuando se consideraba la aparición de cáncer de mama para todo el grupo, la incidencia era 1,8 veces mayor en las mujeres con anomalías hormonales que en aquellas sin tales anomalías, pero el hallazgo no era significativo desde el punto de vista estadístico. Sin embargo, cuando la aparición de cáncer de mama se dividía en las categorías de incidencia premenopáusica y posmenopáusica, las mujeres con anomalías hormonales tenían un riesgo 5,4 veces mayor de aparición de cáncer de mama antes de la menopausia (desarrollaron cáncer de mama antes); no se observaron diferencias en cuanto a la aparición de cáncer de mama después de la menopausia. No está claro si esta falta de diferencia en cuanto a la incidencia de cáncer de mama después de la menopausia representa la verdadera ausencia de una diferencia o si se puede atribuir a la pequeña cantidad de mujeres de esta población que habían llegado a la menopausia en

el momento en que se realizó el estudio.

¿Qué tipo de diseño de estudio es este? Es evidente que se trata de un diseño de cohortes, porque compara las personas expuestas y no expuestas. Por otra parte, debido a que el estudio se llevó a cabo en 1978 y a que el investigador utilizó una lista de pacientes que habían sido atendidas en la clínica de infertilidad de 1945 a 1965, es un diseño de cohortes retrospectivo.

Estudios de cohortes para investigar la salud y las enfermedades infantiles

Un uso especialmente atractivo del diseño de cohortes es para los estudios de cohortes a largo plazo sobre la salud y las enfermedades infantiles. En los últimos años, cada vez se acepta más que las experiencias y exposiciones durante la vida fetal pueden tener efectos a largo plazo, incluso en la vida adulta. Las infecciones durante el embarazo, así como la exposición a toxinas ambientales, las anomalías hormonales o el uso de fármacos o drogas (ya sean medicamentos o sustancias consumidos durante el embarazo), pueden tener efectos potencialmente dañinos en el feto y el niño, y estas sustancias podrían tener efectos que durasen incluso hasta la vida adulta. David Barker y cols. llegaron a la conclusión a partir de sus estudios de que las enfermedades crónicas en la vida adulta se programan biológicamente en la vida intrauterina o en la primera infancia[3]. Se ha hecho hincapié en la importancia de incluir un enfoque del ciclo de vida en el estudio epidemiológico de las enfermedades crónicas a lo largo de la vida.

En este capítulo se han comentado dos tipos de estudios de cohortes; ambos pueden aplicarse al estudio de la salud infantil. En el primer tipo de estudio de cohortes se comienza con los grupos expuesto y no expuesto. Por ejemplo, los estudios de seguimiento de los fetos expuestos a la radiación de las bombas atómicas en Hiroshima y Nagasaki durante la Segunda Guerra Mundial han proporcionado mucha información sobre el cáncer y otros problemas de salud derivados de la exposición intrauterina a la radiación[4]. La dosis de exposición se calibró para los supervivientes basándose en la distancia a la que estaba la mujer embarazada del punto de la caída de la bomba en el momento de su lanzamiento y en las características de las barreras existentes entre dicha persona y el lugar de la caída de la bomba. A continuación, fue posible relacionar el riesgo de resultados adversos con la dosis de radiación que recibió cada persona. Otro ejemplo es la cohorte de embarazos durante la hambruna holandesa en la Segunda Guerra Mundial[5]. Gracias a que en Holanda se conservaron unos registros excelentes, fue posible identificar las cohortes que estuvieron expuestas a la inanición en diferentes momentos de la gestación y compararlas entre sí y con un grupo no expuesto.

Como se ha comentado anteriormente en este capítulo, en el segundo tipo de estudio de cohortes se identifica un grupo antes de que cualquiera de sus miembros resulte expuesto o antes de que se haya identificado la exposición. Por ejemplo, los lactantes nacidos durante una única semana de 1946 en Gran Bretaña fueron seguidos hasta la infancia y más tarde hasta la vida adulta. El Collaborative Perinatal Study, que se inició en Estados Unidos en la década de 1950, fue un estudio de cohortes multicéntrico en el que se siguieron más de 58.000 niños desde el nacimiento hasta los 7 años[6].

Aunque los conocimientos que pueden adquirirse con este tipo de estudios son muy atractivos, surgen varias preguntas de difícil respuesta cuando se conciben unos estudios de cohortes infantiles tan amplios y cuando se planifica un seguimiento tan prolongado. Algunas de las preguntas son las siguientes:

1. ¿En qué momento se deben identificar los individuos de la cohorte por primera vez? Cuando se inicia una cohorte al nacer y luego se sigue (fig. 8.9), los datos sobre las exposiciones prenatales se pueden obtener de forma retrospectiva mediante la anamnesis y los registros pertinentes. Por tanto, algunos estudios de cohortes se han comenzado en el periodo prenatal, la primera vez que se diagnostica el embarazo. Sin embargo, incluso de este modo es posible que los datos preconcepcionales y periconcepcionales que pueden ser necesarios para responder a ciertas preguntas solo se logren obtener de forma retrospectiva. Por tanto, es deseable contar con una cohorte iniciada antes del momento de la concepción (fig. 8.10) para responder a muchas preguntas, porque permite la recopilación simultánea de datos sobre exposiciones en el momento de la concepción o anteriores a esta y en los periodos prenatal y perinatal. Sin embargo, esto suele ser un desafío difícil desde el punto de vista logístico, y muy caro.
2. ¿La cohorte debería constituirse a partir de un centro o de varios centros, o debería ser una muestra nacional elaborada con la intención de que fuese representativa de la población nacional? ¿Los resultados de los estudios basados en la cohorte podrán generalizarse ampliamente solo si la cohorte se obtiene a partir de una muestra nacional? El National Children's Study (NCS) fue un estudio planificado a largo plazo de 100.000 niños y sus padres en los Estados Unidos, que fue diseñado para investigar las influencias ambientales en la salud y el desarrollo del niño. El estudio piloto se inició en 2009 y, para 2013, solo se habían reclutado 5.000 niños de 40 centros en todo Estados Unidos. Basado en las recomendaciones de un panel de expertos, el director de los National Institutes of Health (NIH) dio por finalizado el NCS en 2014. En 2016, el NIH inició un estudio de 7 años llamado Environmental Influences on Child Health Outcomes (ECHO) que utilizó cohortes existentes de niños (y padres en algunos casos) que fueron seguidas utilizando la recopilación de datos armonizada. La «cohorte sintética» resultante (o una cohorte de cohortes) debería ser mucho más eficiente que la cohorte propuesta planeada por el NCS.
3. ¿Durante cuánto tiempo se debe seguir una cohorte? Eaton instó a que una cohorte se debería establecer en el momento de la concepción y seguirse hasta la vida adulta o hasta el fallecimiento[7]. Este enfoque ayudaría a probar la hipótesis de Barker referente a los orígenes

precoces de muchas enfermedades crónicas. Teniendo en cuenta que la financiación federal se limita generalmente a 5 años, esto es un impedimento para el seguimiento a largo plazo.

4. ¿Qué y cuántas hipótesis deben probarse en la cohorte que se constituirá? Un problema principal asociado con el seguimiento a largo plazo de grandes cohortes es que, cuando la cohorte se ha constituido y seguido durante varios años, las hipótesis que dieron lugar originariamente a la creación de la cohorte puede que ya no tengan suficiente interés o relevancia porque el conocimiento científico y sanitario ha cambiado con el tiempo. Por otra parte, a medida que los nuevos conocimientos dan lugar a nuevas hipótesis y a preguntas que no se habían previsto inicialmente cuando se inició el estudio, puede que los datos sobre las variables necesarias para poner a prueba estas nuevas hipótesis y para responder a estas nuevas preguntas no estén disponibles en los datos recopilados en un principio. Un ejemplo de la investigación del VIH/SIDA ilustra estos problemas. A principios de la década de 1980, cuando se identificaron grupos de hombres que sufrían neoplasias malignas raras asociadas con una función inmune comprometida, lo que luego se definiría como VIH/SIDA, los NIH iniciaron el Multicenter AIDS Cohort Study, en 1983, e incorporaron a los primeros participantes de cuatro ciudades de Estados Unidos en 1984[8]. El objetivo era identificar los factores de riesgo de esta enfermedad viral y dilucidar la historia natural de la enfermedad. Con el advenimiento de la terapia antirretroviral altamente activa en 1996, prácticamente todos los participantes del estudio que ya estaban infectados fueron sometidos a tratamiento, y sus sistemas inmunitarios fueron reconstituidos. ¿Cómo podría entonces ser relevante la historia natural de una infección tratada? ¿Qué utilidad tenía seguir a esta cohorte? De hecho, se plantearon una gran cantidad de preguntas nuevas y relevantes, entre las que destaca cuál es el impacto de los tratamientos antirretrovirales a largo plazo sobre el envejecimiento natural y la incidencia de enfermedades crónicas (cáncer, enfermedades cardiovasculares y diabetes, entre otras)[9]. Además, en los últimos 15 años se han descubierto nuevas pruebas genéticas que brindan nuevas perspectivas sobre por qué algunos participantes evolucionan mejor que otros con el tratamiento[10]. Se debe enfatizar que los estudios de cohortes cuyos participantes se examinan periódicamente, como el estudio Atherosclerosis Risk in Communities (ARIC)[11] permiten la evaluación de nuevas hipótesis basadas en la información que se recopila en los exámenes de seguimiento.

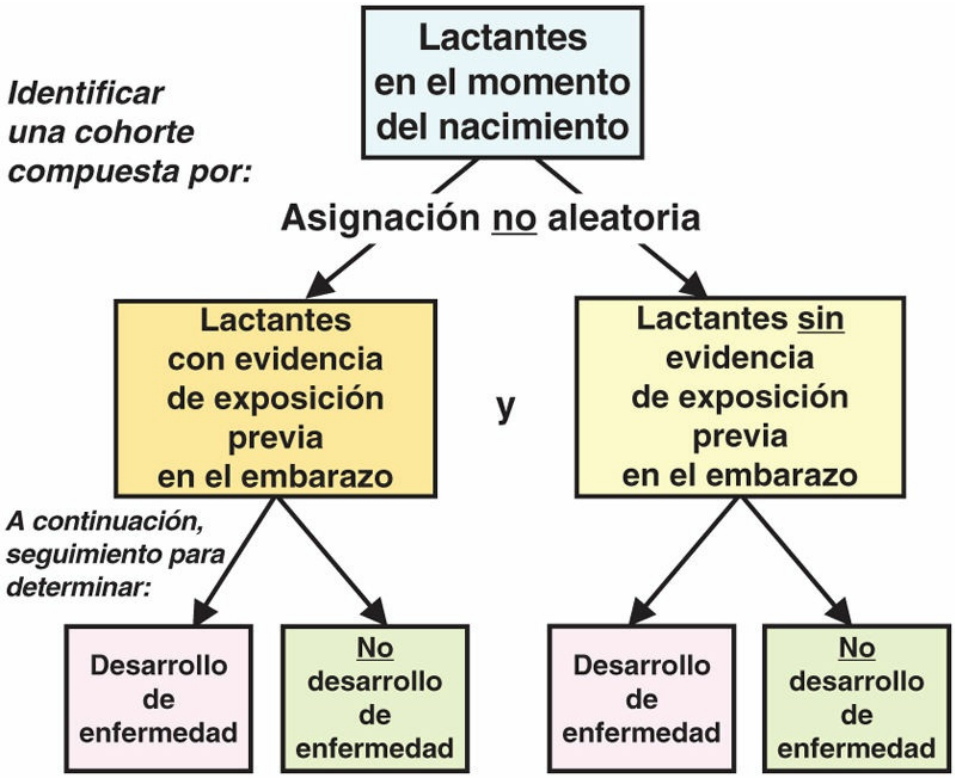


FIG. 8.9 Diseño de un estudio de cohortes para investigar los efectos de exposiciones durante el embarazo sobre las enfermedades a lo largo de la vida: estudio iniciado al nacer.

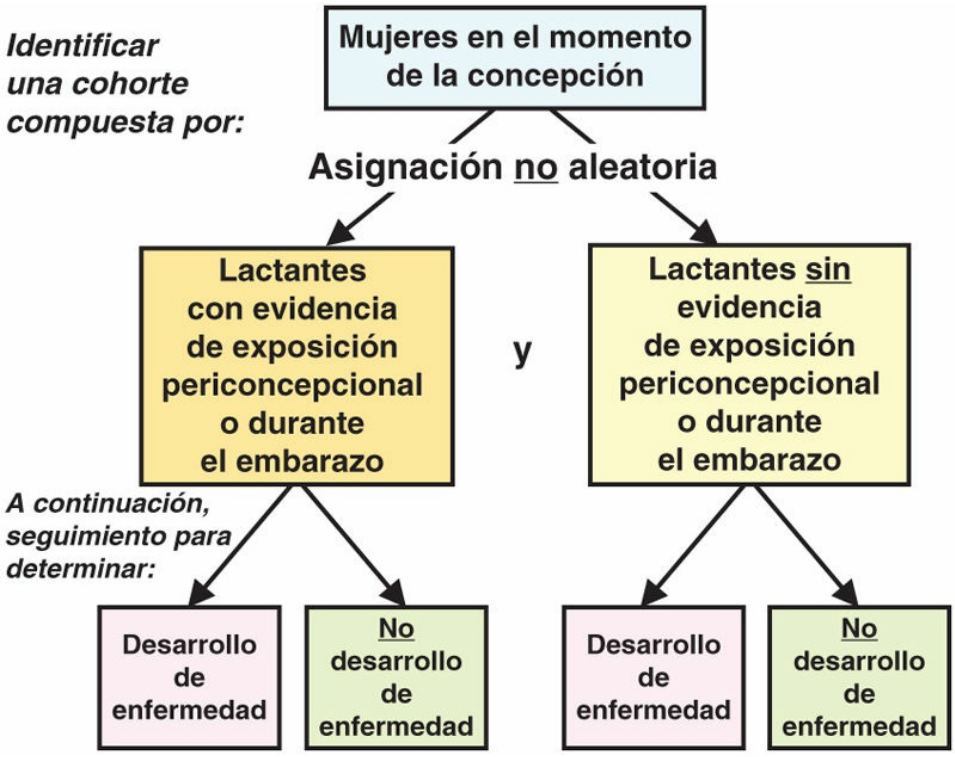


FIG. 8.10 Diseño de un estudio de cohortes para investigar los efectos de exposiciones durante el embarazo sobre las enfermedades a lo largo de la vida: estudio iniciado en el momento de la concepción.

Posibles sesgos en los estudios de cohortes

Al realizar estudios de cohortes se deben evitar o tener en cuenta los sesgos que pueda haber. La exposición sobre los sesgos en relación con los estudios de casos y controles se presentó anteriormente; el sesgo en relación con las inferencias causales será presentado más adelante. Las definiciones utilizadas para muchos tipos de sesgos a menudo se superponen; y, en aras de la claridad, hay dos categorías principales que suelen usarse: sesgo de selección y sesgo de información.

Sesgos de selección

La ausencia de participación y la falta de respuesta pueden introducir sesgos significativos que pueden complicar la interpretación de los hallazgos del estudio. Si los participantes se niegan a incorporarse a una cohorte, ¿podrían sus características diferir suficientemente de aquellos que consienten en inscribirse, y estas diferencias pueden conducir a inferencias erróneas con respecto a las exposiciones a los resultados? Por ejemplo, si los que se niegan a participar en un estudio tienen más probabilidades de fumar que los que aceptan participar, ¿nuestra estimación del efecto del hábito de fumar en el resultado de la enfermedad estaría sesgada? Si los fumadores que se niegan a participar son más propensos a desarrollar la enfermedad que los que participan, el impacto sería disminuir la asociación hacia el nulo. Del mismo modo, la pérdida durante el seguimiento puede ser un grave problema: si las personas con la enfermedad se pierden de forma selectiva durante el seguimiento, y las que se pierden durante el seguimiento difieren de las que no se pierden, las tasas de incidencia calculadas en los grupos expuesto y no expuesto serán claramente difíciles de interpretar.

Sesgos de información

1. Si la calidad y la cantidad de la información obtenida son diferentes para las personas expuestas y para las no expuestas, se puede introducir un sesgo significativo. Esto es muy probable que ocurra en los estudios de cohortes históricos, en los que la información se obtiene a partir de los registros previos. Como se abordará a continuación en relación con los ensayos clínicos, en cualquier estudio de cohortes es esencial que la calidad de la información obtenida sea comparable tanto en las personas expuestas como en las no expuestas.
2. Si la persona que decide si la enfermedad se ha desarrollado en cada

sujeto también sabe si dicho sujeto estuvo expuesto, y si esa persona conoce la hipótesis que se está probando, el criterio de esa persona en cuanto a si la enfermedad se ha desarrollado puede estar sesgado por ese conocimiento. Este problema puede resolverse mediante el enmascaramiento de la persona que realiza la evaluación de la enfermedad y también determinando si esta persona conocía, en realidad, el estatus de exposición de cada sujeto.

3. Al igual que en cualquier estudio, si los epidemiólogos y estadísticos que analizan los datos tienen prejuicios firmes, pueden introducir involuntariamente sus sesgos en el análisis de los datos y en la interpretación de los hallazgos del estudio.

¿Cuándo se requiere un estudio de cohortes?

En la figura 8.11A-C se revisan los pasos básicos de un estudio de cohortes, comenzando con la identificación de un grupo expuesto y de otro no expuesto (v. fig. 8.11A). A continuación, se determina la incidencia, tanto en el grupo expuesto como en el no expuesto (v. fig. 8.11B). Si la exposición se asocia con la enfermedad, sería previsible encontrar una mayor incidencia de la enfermedad en el grupo expuesto que en el no expuesto, como se muestra esquemáticamente en la figura 8.11C.

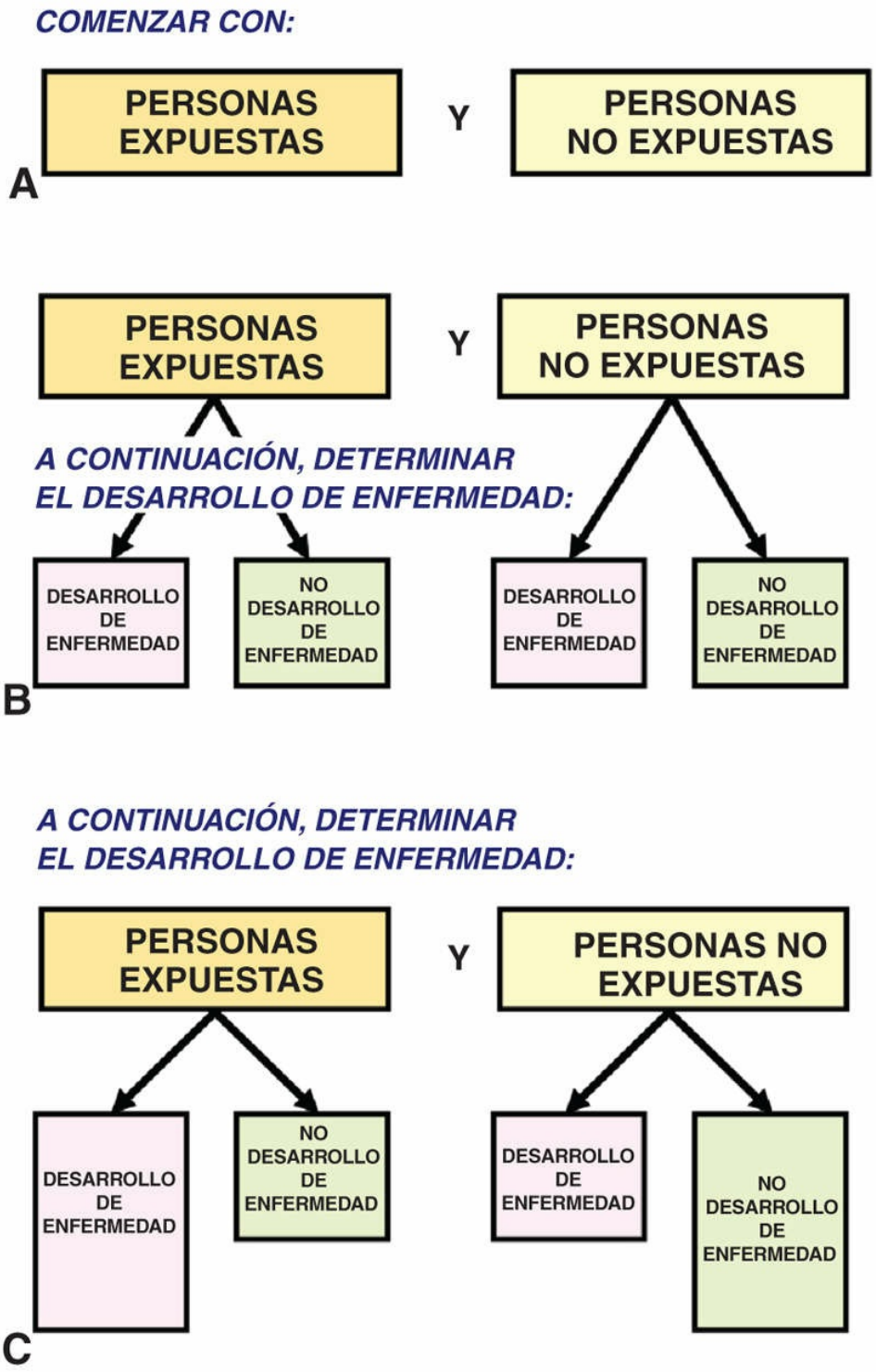


FIG. 8.11 Diseño de un estudio de cohortes. (A) Inicio con los grupos expuesto y no expuesto. (B) Determinación del desarrollo de enfermedad en ambos grupos. (C) Hallazgos previstos si la exposición se asocia a enfermedad.

Es evidente que para llevar a cabo un estudio de cohortes se debe tener una cierta idea de qué exposiciones se sospechan a priori como posibles causas de una enfermedad y, por tanto, vale la pena investigar. En consecuencia, un estudio de cohortes está indicado cuando una evidencia adecuada sugiere una asociación de la enfermedad con una cierta exposición o exposiciones (evidencia obtenida a partir de observaciones clínicas o bien de estudios de

casos y controles o de otro tipo). A menudo obtenemos muestras biológicas al inicio del estudio (reclutamiento), lo que permite el análisis de estas muestras en el futuro, frecuentemente cuando se desarrollan nuevos métodos de prueba y/o se generan nuevas hipótesis. Como ejemplo, George Comstock recogió muestras de suero durante una evaluación comunitaria en la década de 1960 en el condado de Washington, Maryland. Décadas más tarde, estas muestras se analizaron en busca de «pistas» para el desarrollo del cáncer. Los resultados del estudio de cohortes Campaign Against Cancer and Heart Disease (CLUE II) iniciado por el Dr. Comstock demostraron que el colesterol sérico alto aumenta el riesgo de cáncer de próstata de alto grado y, posteriormente, apoyaron la hipótesis de que la reducción del colesterol es un mecanismo potencial por el cual las estatinas, un medicamento para bajar el colesterol, podrían tener efectos anticancerígenos[12].

Debido a que los estudios de cohortes a menudo implican el seguimiento de las poblaciones durante un periodo prolongado, el enfoque de cohortes es especialmente atractivo cuando se puede minimizar la deserción (pérdidas durante el seguimiento) de la población estudiada. En consecuencia, tales estudios suelen ser más fáciles de llevar a cabo cuando el intervalo entre la exposición y el desarrollo de la enfermedad es corto. Un ejemplo de una asociación en la que el intervalo entre la exposición y el resultado es corto es la relación entre la infección por rubéola durante el embarazo y el desarrollo de malformaciones congénitas en la descendencia.

Estudios de casos y controles basados en una cohorte definida

En los últimos años, se ha prestado mucha atención a si es posible aprovechar los beneficios de los diseños de estudio de casos y controles y de cohortes combinando algunos elementos de ambos en un solo estudio. El estudio combinado resultante es en efecto un diseño híbrido en el que se inicia un estudio de casos y controles dentro de un estudio de cohortes. El diseño general se muestra esquemáticamente en la figura 8.12.

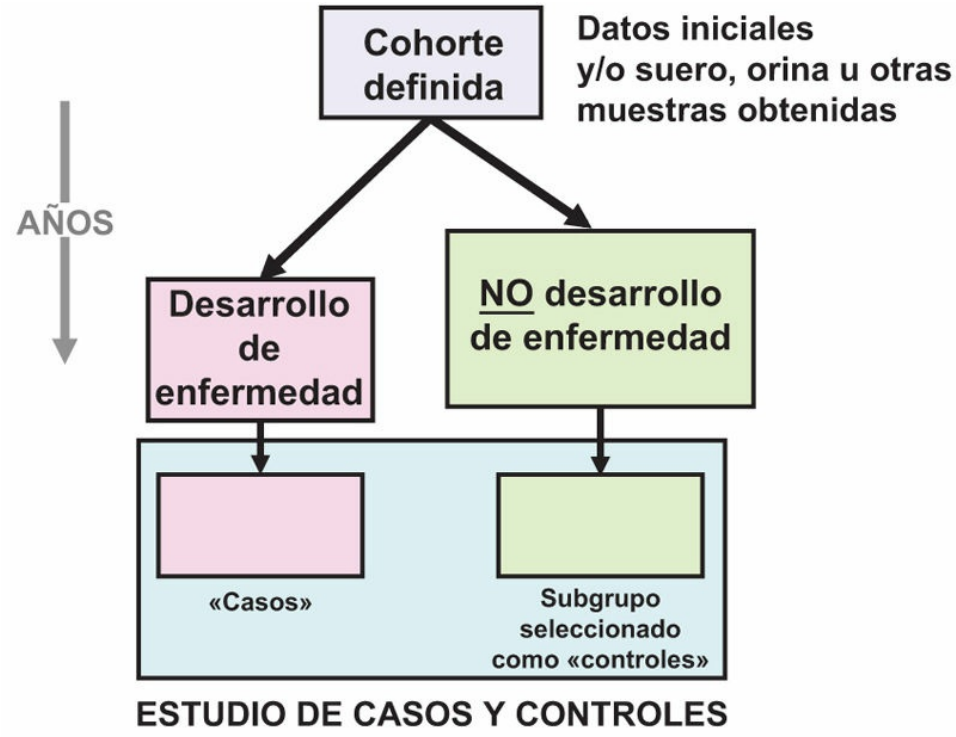


FIG. 8.12 Diseño de un estudio de casos y controles iniciado dentro de una cohorte.

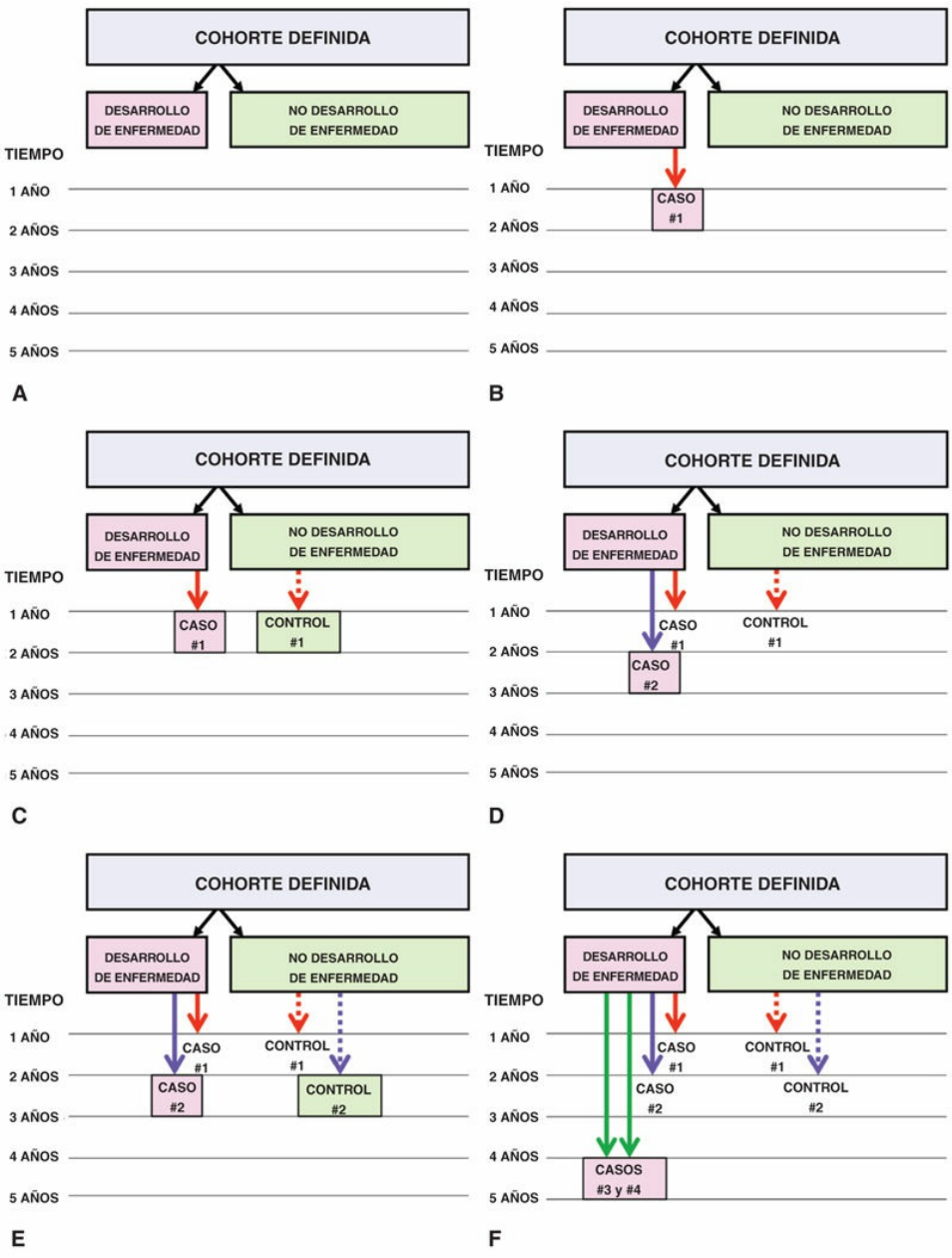
En este tipo de estudio, una población es identificada y seguida en el tiempo. En el momento en que se identifica a la población, los datos de referencia se obtienen de registros o entrevistas, análisis de sangre u orina y de otras maneras. La población es seguida por un periodo de años. Para la mayoría de las enfermedades que se estudian, un pequeño porcentaje de los participantes del estudio manifiesta la enfermedad, mientras que la mayoría no lo hace. Como se observa en la figura 8.12, a continuación, se lleva a cabo un estudio de casos y controles utilizando como casos a las personas en las que se desarrolló la enfermedad y usando como controles una muestra de aquellos en quienes no se desarrolló la enfermedad.

Dichos estudios de casos y controles basados en cohortes se pueden dividir

en dos tipos, en gran medida en función del enfoque utilizado para seleccionar los controles. Estos dos tipos de estudios se denominan *estudios de casos y controles anidados* y *estudios de cohortes de casos.*

Estudios de casos y controles anidados

En los *estudios de casos y controles anidados*, los controles son una muestra de individuos que están en riesgo de contraer la enfermedad *en el momento en que se desarrolla cada caso de la enfermedad.* Esto se muestra esquemáticamente en la figura 8.13A a I.



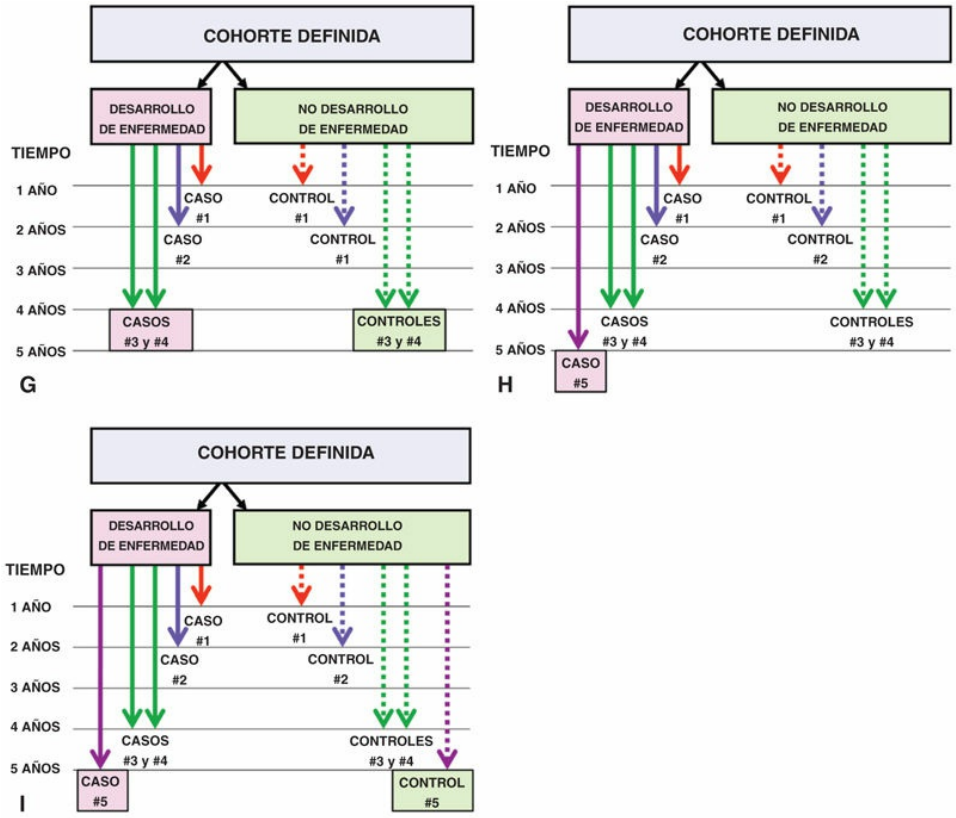


FIG. 8.13 (A–I) Diseño de un estudio hipotético de casos y controles anidado: pasos para seleccionar casos y controles.

La figura 8.13A muestra el punto de partida como una cohorte definida de individuos. Algunos de ellos desarrollan la enfermedad en cuestión, pero la mayoría no lo hacen. En este ejemplo hipotético, la cohorte es observada durante un periodo de 5 años. Durante este tiempo se desarrollan cinco casos: un caso después de 1 año, uno después de 2 años, dos después de 4 años y uno después de 5 años.

Sigamos la secuencia de pasos a lo largo del tiempo. Las figuras 8.13B a I muestran la secuencia temporal en la que se desarrollan los casos después del inicio de las observaciones. En el momento en que se desarrolla cada caso, o casos, se selecciona el mismo número de controles. Las flechas continuas en el lado izquierdo de la figura indican la aparición de casos de la enfermedad, y las flechas de puntos en el lado derecho indican la selección de controles que están libres de enfermedad, pero que están en riesgo de desarrollar dicha enfermedad en el momento en el que el caso desarrolla la enfermedad. La figura 8.13B muestra el caso #1 que apareció después de 1 año, y la figura 8.13C muestra el control #1 seleccionado en ese momento. La figura 8.13D muestra el caso #2 que apareció después de 2 años, y la figura 8.13E muestra el control #2 seleccionado en ese momento. La figura 8.13F muestra los casos #3 y #4 de aparición después de 4 años, y la figura 8.13G muestra los controles #3 y #4 seleccionados en ese momento. Por último, la figura 8.13H muestra el último caso (#5) que apareció después de 5 años, y la figura 8.13I muestra el control #5 que se está seleccionando en este punto.

La figura 8.13I es también un resumen del diseño y las poblaciones de estudio finales utilizadas en el estudio de casos y controles anidado. Al final de los 5 años aparecieron cinco casos y en el momento en que aparecieron, se seleccionaron un total de cinco controles para el estudio. De esta manera, los casos y los controles se corresponden, en efecto, temporalmente y respecto a la duración del seguimiento. Debido a que cada vez que aparece un caso se selecciona un control, en el mismo estudio un control seleccionado al principio de este podría desarrollar la enfermedad y convertirse en un caso.

Estudios de casos y cohortes

El segundo tipo de estudio de casos y controles basado en cohortes es el *diseño de casos y cohortes* que se muestra en la figura 8.14. En el hipotético estudio de casos y cohortes que abordamos aquí, los casos se desarrollan en los mismos momentos que se vieron en el diseño de casos y controles anidado que acabamos de exponer, pero los controles se seleccionan aleatoriamente de la cohorte definida con la que comenzó el estudio. Este subconjunto de la cohorte completa se llama subcohorte. Una ventaja de este diseño es que, debido a que los controles no se corresponden individualmente con cada caso, es posible estudiar diferentes enfermedades (diferentes conjuntos de casos) en el mismo estudio de cohortes de casos utilizando la misma cohorte para los controles. En este diseño, a diferencia del diseño de casos y controles anidado, los casos y los controles no coinciden en el tiempo ni en la duración del seguimiento; en cambio, la exposición se caracteriza por la subcohorte. Esta diferencia en el diseño del estudio debe tenerse en cuenta al analizar los resultados de este.

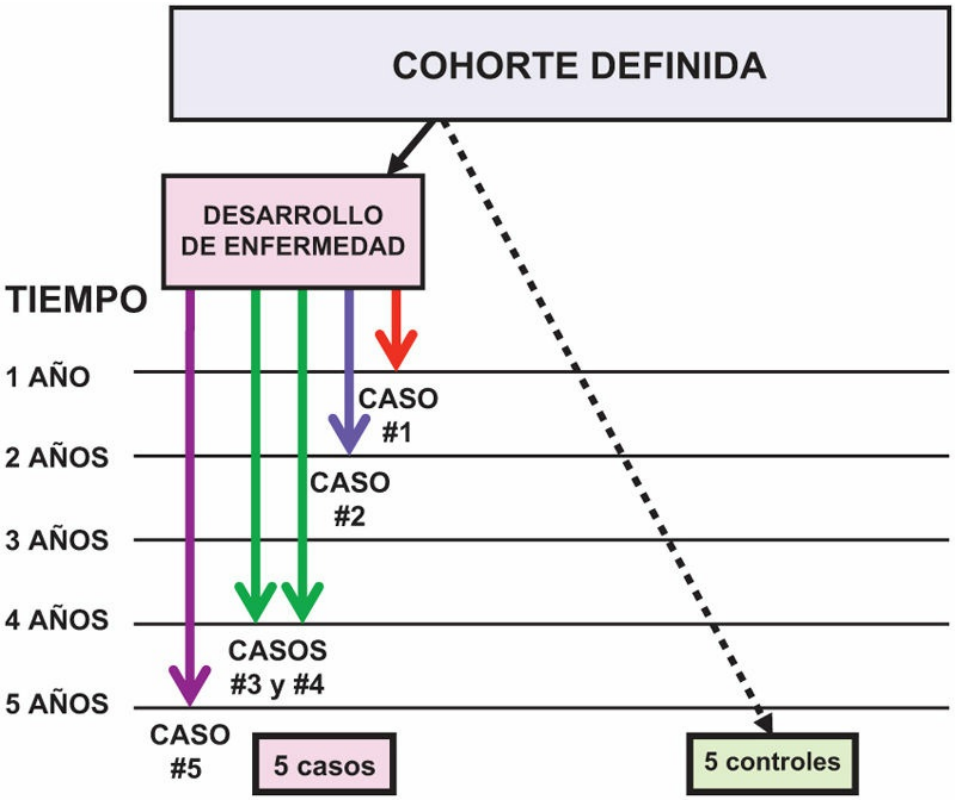


FIG. 8.14 Diseño de un estudio hipotético de casos y cohortes: pasos para seleccionar casos y controles.

Ventajas de insertar un estudio de casos y controles en una cohorte definida

¿Cuáles son las ventajas de realizar un estudio de casos y controles en una cohorte definida? Primero, debido a que las entrevistas o las muestras de sangre u orina se obtuvieron al inicio del estudio (en la línea de base), los datos se consiguen antes de que se haya desarrollado cualquier enfermedad. En consecuencia, se elimina el problema del posible sesgo de recuerdo discutido anteriormente en este capítulo. En segundo lugar, si se encuentran anomalías en las características biológicas, como en los valores de laboratorio, como las muestras se obtuvieron años antes del desarrollo de la enfermedad clínica, es más probable que estos hallazgos representen factores de riesgo u otras características premórbidas que una manifestación de enfermedad subclínica temprana. Cuando se encuentran dichas anomalías en el estudio tradicional de casos y controles, no sabemos si precedieron a la enfermedad o si fueron el resultado de esta, particularmente cuando la enfermedad tiene una fase subclínica (asintomática) larga, como el cáncer de próstata y la leucemia linfocítica crónica. Tercero, tal estudio es a menudo más económico de realizar. Uno podría preguntarse, ¿por qué realizar un estudio de casos y controles dentro de una cohorte definida? ¿Por qué no realizar un estudio de cohorte prospectivo regular? La respuesta es que en un estudio de cohorte de, digamos, 10.000 personas, deben realizarse análisis de laboratorio de todas las

muestras obtenidas, a menudo con un coste elevado, para definir los grupos *expuestos y no expuestos.* Sin embargo, en un estudio de casos y controles dentro de la misma cohorte, las muestras obtenidas inicialmente se congelan o se almacenan de otro modo. Solo después de que la enfermedad se ha desarrollado en algunos sujetos, se inicia un estudio de casos y controles y se descongelan y analizan las muestras del número relativamente pequeño de personas que están incluidas en el estudio de casos y controles. Las pruebas de laboratorio no deben realizarse en las 10.000 personas de la cohorte original, por lo que la carga al laboratorio y los costos se reducen drásticamente.

Finalmente, tanto en los diseños de casos y controles anidados como en los de casos y cohortes, los casos y los controles derivan de la misma cohorte original, por lo que es probable que haya una mayor comparabilidad entre los casos y los controles de lo que normalmente se puede encontrar en un estudio de casos-controles tradicional. Por todas estas razones, el estudio de casos y controles basado en cohortes es un tipo de diseño de estudio extremadamente valioso.

Conclusión

Hay varios aspectos que pueden hacer que el diseño de cohortes sea poco práctico. A menudo no existe una evidencia sólida para justificar la elaboración de un estudio extenso y costoso que evalúe a fondo el papel de un factor de riesgo específico en la etiología de una enfermedad. Incluso si se dispone de tal evidencia, a menudo no se puede identificar con facilidad una cohorte de personas expuestas y no expuestas. En general, no existen registros anteriores apropiados u otras fuentes de datos que permitan llevar a cabo un estudio de cohortes retrospectivo; como resultado, se requiere un estudio prolongado debido a la necesidad de un seguimiento largo de la población después de la exposición. Además, muchas de las enfermedades que son de interés hoy en día tienen una incidencia muy baja. Por tanto, se deben reclutar cohortes muy extensas en un estudio para garantizar que aparezcan suficientes casos al final del periodo de estudio con el fin de lograr que el análisis y las conclusiones tengan validez.

A la vista de estos aspectos, a menudo se necesita un planteamiento distinto a un diseño de cohortes y que pueda superar muchas de estas dificultades. Como expusimos previamente, estos diseños de estudio (estudio de casos y controles y diseños de estudios transversales) se utilizan cada vez más. Más adelante se comenta el uso de estos diseños de estudios en la estimación del aumento del riesgo asociado a una exposición.

Bibliografía

1. Kannel WB. CHD risk factors: a Framingham Study update. *Hosp Pract*. 1990;25:93–104.
2. Cowan LD, Gordis L, Tonascia JA, et al. Breast cancer incidence in women with progesterone deficiency. *Am J Epidemiol*. 1981;114:209–217.
3. Barker DJP, ed. Fetal and Infant Origins of Adult Disease. London: BMJ Books; 1992.
4. Yoshimoto Y, Kato H, Schull WJ. Cancer risk among in utero exposed survivors: a review of 45 years of study of Hiroshima and Nagasaki atomic bomb survivors. *J Radiat Res (Tokyo)*. 1991;32(suppl):231–238: Also available as RERF Technical Report 4–88, and in *Lancet*. 1988;2:665–669.
5. Susser E, Hoek HW, Brown A. Neurodevelopmental disorders after prenatal famine: the story of the Dutch Famine Study. *Am J Epidemiol*. 1998;147:213–216.
6. Broman S, The collaborative perinatal project: an overview. Mednick SA, Harway M, Pinello KM, eds. Handbook of Longitudinal Research, Vol I. New York: Praeger; 1984.
7. Eaton WW. The logic for a conception-to-death cohort study. *Ann Epidemiol*. 2002;12:445–451.
8. Kaslow R, Ostrow DG, Detels R, et al. for the Multicenter AIDS Cohort Study. The Multicenter AIDS Cohort Study: rationale, organization, and selected characteristics of the participants. *Am J Epidemiol*. 1987;126:310–318.
9. Brown TT, Cole SR, Li X, et al. Antiretroviral therapy and the prevalence and incidence of diabetes mellitus in the Multicenter AIDS Cohort Study. *Arch Intern Med*. 2005;165:1179–1184.
10. Winkler C, Modi W, Smith MW, et al. Genetic restriction of AIDS pathogenesis by an SDF-1 chemokine gene variant. *Science*. 1998;279:389–393.
11. ARIC investigators The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. *Am J Epidemiol*. 1989;129:687–702.
12. Mondul AM, Clipp SL, Helzlsouer KJ, et al. Association between plasma total cholesterol concentration and incident prostate cancer in the CLUE II cohort. *Cancer Causes Control*.

2010;21(1):61–68.

Preguntas de repaso del capítulo 8

1. *En los estudios de cohortes sobre el papel de un factor del que se sospecha su implicación en la etiología de una enfermedad, es esencial que:*

a. Haya el mismo número de personas en los dos grupos del estudio.

b. Al comienzo del estudio, las personas con la enfermedad y las personas sin ella presenten el mismo riesgo de tener el factor.

c. El grupo de estudio con el factor y el grupo de estudio sin él sean representativos de la población general.

d. Los grupos expuesto y no expuesto sometidos a estudio sean lo más similares posible en lo que respecta a los posibles factores de confusión.

e. *b* y *c*.

2. *¿Cuál de los siguientes aspectos no es una ventaja de un estudio de cohortes prospectivo?*

a. Suele ser más barato que un estudio de casos y controles.

b. Puede realizarse una medición precisa de la exposición.

c. Se pueden calcular las tasas de incidencia.

d. El sesgo de recuerdo se minimiza en comparación con un estudio de casos y controles.

e. Se pueden estudiar simultáneamente muchos resultados de la enfermedad.

3. *Los estudios de cohortes retrospectivos se caracterizan por todo lo siguiente excepto:*

a. Los grupos de estudio son el grupo expuesto y el no expuesto.

b. Se pueden calcular las tasas de incidencia.

c. El tamaño muestral requerido es menor que el necesario para un estudio de cohortes prospectivo.

d. El tamaño muestral requerido es similar al necesario para un estudio de cohortes prospectivo.

e. Son útiles para exposiciones raras.

4. *Un problema principal debido a la falta de asignación aleatoria en un*

estudio de cohortes es:

a. La posibilidad de que la enfermedad esté causada por un factor que condujo a la exposición en lugar de por la propia exposición.

b. La posibilidad de que una mayor proporción de personas del estudio puede haber estado expuesta.

c. La posibilidad de que una menor proporción de personas del estudio puede haber estado expuesta.

d. El hecho de que, sin una asignación aleatoria, el estudio puede requerir más tiempo para llevarse a cabo.

e. El cruce de grupos planificado es más probable.

5. *En un estudio de cohortes, la ventaja de comenzar seleccionando una población definida para su estudio antes de que cualquiera de sus miembros se exponga, en lugar de comenzar por seleccionar a los individuos expuestos y no expuestos, es que:*

a. El estudio se puede completar con mayor rapidez.

b. Se pueden estudiar varios resultados de forma simultánea.

c. Se pueden estudiar varias exposiciones de forma simultánea.

d. La realización del estudio será más barata.

e. *a* y *d*.

6. *En 2010, los investigadores estaban interesados en estudiar la obesidad en adultos jóvenes como factor de riesgo de mortalidad por cáncer. Los investigadores obtuvieron informes médicos de alumnos que estudiaron en la Universidad de Glasgow entre 1948 y 1968. Estos informes incluían registros de las alturas y pesos de los estudiantes en el momento en que asistían a la universidad. Los estudiantes fueron seguidos hasta el año 2010. La información acerca de la mortalidad se obtuvo utilizando certificados de defunción. Este estudio se puede describir mejor como:*

a. Casos-controles anidado

b. Transversal

c. Cohorte prospectiva

d. Cohorte retrospectiva

e. Casos-controles basado en la población

7. *De 1983 a 1988, se obtuvieron muestras de sangre de 3.450 hombres VIH-negativos en el Multicenter AIDS Cohort Study (MACS) y se almacenaron en un depósito nacional. En 2010, un investigador estaba*

interesado en examinar la asociación entre los niveles de inflamación y la infección por el VIH. De los 3.450 hombres, 660 hombres fueron identificados como casos de infección por VIH. El investigador estudió la asociación entre la proteína C reactiva (PCR) y la infección por VIH entre estos 660 casos y 660 controles, ajustados con los casos por edad y raza, que no se infectaron con el VIH. El investigador utilizó las muestras de sangre almacenadas para medir el nivel sérico de PCR, un marcador de inflamación sistémica. El estudio iniciado en 2010 es un ejemplo de:

a. Estudio de casos y cohortes anidado

b. Estudio de casos-controles anidado

c. Estudio de cohortes retrospectivo

d. Estudio transversal

e. Ensayo clínico aleatorizado

CAPÍTULO 9

Comparación de los estudios de casos y controles y de cohortes

Palabras clave

comparación de expuestos y no expuestos
comparación de enfermos (casos) y no enfermos (controles)
comparación de los estudios de casos y controles y de cohortes
temporalidad

Llegados a este punto de nuestra exposición, revisaremos una parte del material que se ha expuesto hasta este punto en la sección 2. Debido a que los temas se están describiendo de una manera escalonada, es importante comprender lo que se ha presentado hasta ahora.

En primer lugar, se compararán los diseños de los estudios de cohortes y de casos y controles, como se ve en la figura 9.1. El aspecto principal que distingue ambos tipos de diseño de estudio es que, en un estudio de cohortes, se comparan personas expuestas y no expuestas, y en un estudio de casos y controles, se comparan personas con la enfermedad (casos) y sin la enfermedad (controles) (fig. 9.2A). En los estudios de cohortes se compara la incidencia de la enfermedad en los individuos expuestos y no expuestos, mientras que en los estudios de casos y controles se comparan las proporciones de quienes tienen la exposición de interés en las personas con la enfermedad y en las personas sin la enfermedad (v. fig. 9.2B).

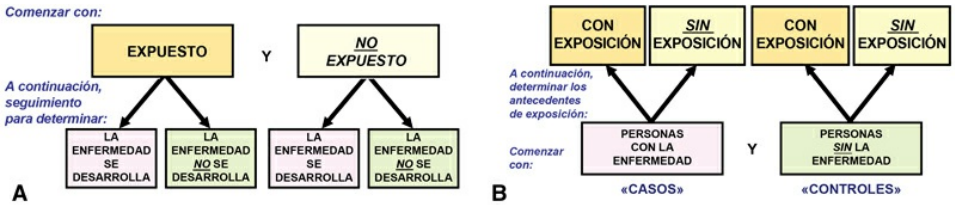


FIG. 9.1 Diseño de estudios de cohortes y de casos y controles. (A) Estudio de cohortes. (B) Estudio de casos y controles.

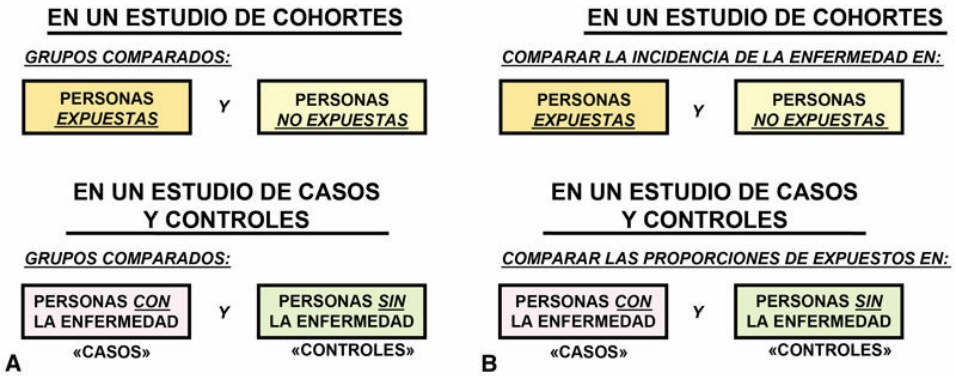


FIG. 9.2 Comparación de los diseños de los estudios de cohortes y de casos y

controles. (A) Grupos comparados. (B) Medidas de resultado.

En la tabla 9.1 se muestra una comparación detallada de los diseños de los estudios de cohortes prospectivo, de cohortes retrospectivo (histórico) y de casos y controles. Si el lector ha seguido la exposición de la sección 2 hasta este punto, las entradas de la tabla deberían ser fáciles de entender.

Tabla 9.1

Comparación de los estudios de casos y controles y de cohortes

	ESTUDIOS DE COHORTES		Estudios de casos y controles
	Prospectivo	**Retrospectivo**	
A. Grupo de estudio	Personas expuestas: $(a + b)$	Personas expuestas: $(a + b)$	Personas con la enfermedad (casos): $(a + c)$
B. Grupo de comparación	Personas no expuestas: $(c + d)$	Personas no expuestas: $(c + d)$	Personas sin la enfermedad (controles): $(b + d)$
C. Medidas de resultado	Incidencia en expuestos $\left(\frac{a}{a+b}\right)$ e Incidencia en no expuestos $\left(\frac{c}{c+d}\right)$	Incidencia en expuestos $\left(\frac{a}{a+b}\right)$ e Incidencia en no expuestos $\left(\frac{c}{c+d}\right)$	Proporción de casos expuestos $\left(\frac{a}{a+c}\right)$ y Proporción de controles expuestos $\left(\frac{b}{b+d}\right)$
D. Medidas de riesgo	Riesgo absoluto	Riesgo absoluto	—
	Riesgo relativo	Riesgo relativo	—
	Odds ratio (cociente de posibilidades)	*Odds ratio* (cociente de posibilidades)	*Odds ratio* (cociente de posibilidades)
	Riesgo atribuible	Riesgo atribuible	Riesgo atribuible[a]
E. Relación temporal entre la exposición y la enfermedad	Fácil de establecer	A veces es difícil de establecer	A veces es difícil de establecer
F. Asociaciones múltiples	Es posible estudiar las asociaciones de una exposición con varias enfermedades[b]	Es posible estudiar las asociaciones de una exposición con varias enfermedades[b]	Es posible estudiar las asociaciones de una enfermedad con varias exposiciones o factores
G. Tiempo requerido para el estudio	Por lo general largo, debido a la necesidad de seguir a los sujetos	Puede ser corto	Relativamente corto
H. Coste del estudio	Caro	Generalmente más barato que un estudio prospectivo	Relativamente barato
I. Tamaño de la población necesario	Relativamente grande	Relativamente grande	Relativamente pequeño
J. Posibles sesgos	Evaluación del resultado	Susceptible de sesgos tanto en la evaluación de la exposición como en la evaluación de los	Evaluación de la exposición

		resultados	
K. Idóneo cuando	La exposición es infrecuente	La exposición es infrecuente	La enfermedad es infrecuente
	La enfermedad es frecuente en los expuestos	La enfermedad es frecuente en los expuestos	La exposición es frecuente en las personas con enfermedad
L. Problemas	La selección del grupo de comparación no expuesto suele ser difícil Cambios de los criterios y los métodos a lo largo del tiempo	La selección del grupo de comparación no expuesto suele ser difícil Cambios de los criterios y los métodos a lo largo del tiempo	La selección de los controles adecuados suele ser difícil Información incompleta sobre la exposición

[a] Se debe disponer de información adicional.

[b] También es posible estudiar exposiciones múltiples cuando la población de estudio se selecciona basándose en un factor no relacionado con la exposición.

Cuando se comienza un estudio de cohortes con los grupos expuesto y no expuesto, solo se puede estudiar la exposición específica que distingue a un grupo de otro. Sin embargo, como se muestra en la figura 9.3, es posible estudiar varios resultados o enfermedades en relación con la exposición de interés. La mayoría de los estudios de cohortes comienzan con individuos expuestos y no expuestos. Menos común es la situación en la que se parte de una población definida en la que se selecciona la población de estudio basándose en un factor no relacionado con la exposición, como el lugar de residencia, y algunos miembros de la cohorte se convierten en expuestos y otros en no expuestos a lo largo del tiempo (fig. 9.4). En un estudio de cohortes que se inicia con una población definida, es posible estudiar varias exposiciones. Así, por ejemplo, en el estudio Framingham, fue posible estudiar muchas exposiciones, como el peso, la presión arterial, la concentración de colesterol, el tabaquismo y la actividad física en las personas participantes que residían en Framingham, Massachusetts.

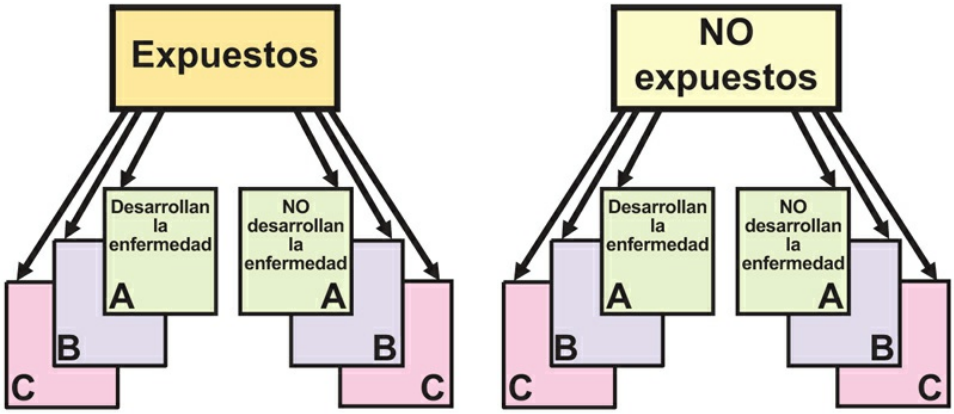


FIG. 9.3 En un estudio de cohortes que se inicia con un grupo expuesto y un grupo no expuesto se pueden estudiar múltiples resultados, pero solo una exposición.

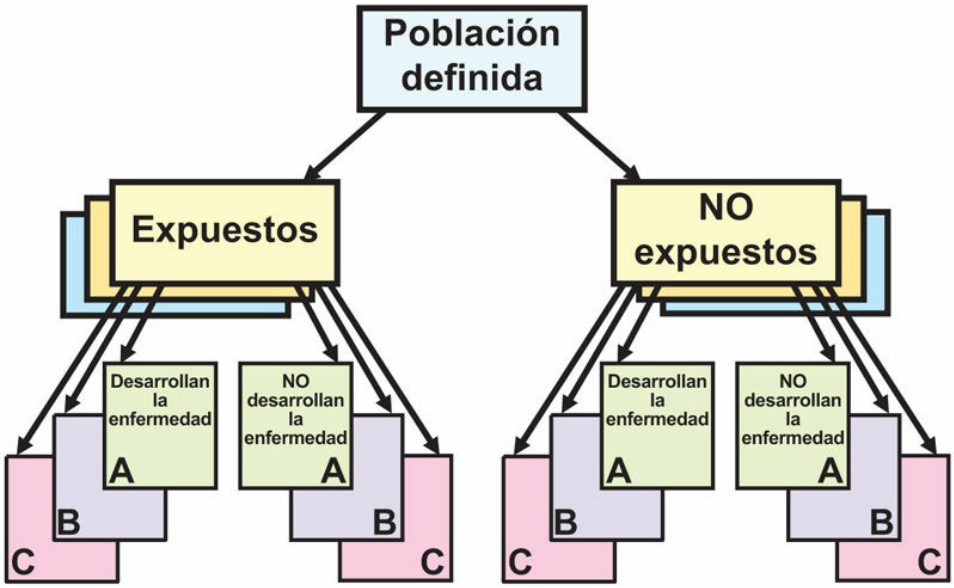


FIG. 9.4 En un estudio de cohortes que se inicia con una población definida se pueden estudiar tanto varias exposiciones como varios resultados.

En los estudios de cohortes, *se puede* calcular la incidencia tanto en el grupo expuesto como en el no expuesto y, por tanto, es posible calcular directamente el riesgo relativo. Los estudios de cohortes prospectivos minimizan la posibilidad de que se produzcan sesgos de recuerdo y de otro tipo a la hora de analizar la exposición y tienen mayor validez de las evaluaciones de la exposición. Sin embargo, en los estudios de cohortes retrospectivos, que requieren datos del pasado, estos problemas pueden ser significativos. Los estudios de cohortes son recomendables cuando la exposición de interés es infrecuente. En un diseño de casos y controles, es poco probable que se identifique un número suficiente de personas expuestas cuando se está analizando una exposición infrecuente. En los estudios de cohortes prospectivos en particular, es probable que se obtengan mejores datos sobre la relación temporal entre la exposición y el resultado (es decir,

¿la exposición precedió al resultado?). Uno de los inconvenientes de los estudios de cohortes es que suelen requerir grandes poblaciones y, en general, los estudios de cohortes prospectivos son especialmente caros de realizar, ya que se requiere un seguimiento de una población extensa a lo largo del tiempo. En los estudios de cohortes hay mayores posibilidades de que se produzcan sesgos a la hora de evaluar el resultado presente que en los estudios de casos y controles. Por último, los estudios de cohortes a menudo se vuelven inviables cuando la enfermedad que se está estudiando es infrecuente.

Como se observa en la tabla 9.1, los estudios de casos y controles tienen una serie de ventajas. Son relativamente baratos y requieren un número relativamente pequeño de sujetos para el estudio. Son recomendables cuando la incidencia de la enfermedad es baja, ya que, si se realizase un estudio de cohortes en esta circunstancia, se requeriría seguir a un número muy elevado de personas para obtener una cantidad suficiente de sujetos con la enfermedad para el estudio. Como se ve en la figura 9.5, en un estudio de casos y controles, gracias a que se empieza con casos y controles, es posible estudiar más de un posible factor etiológico y analizar las interacciones entre los factores.

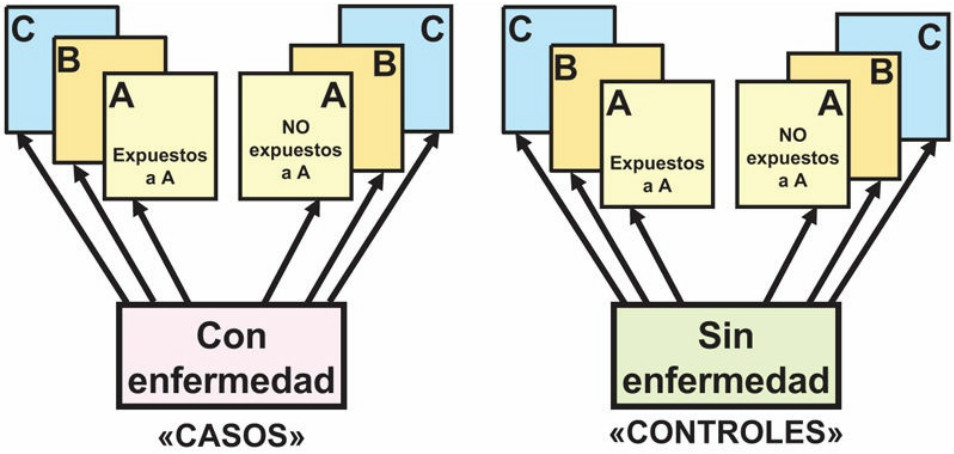


FIG. 9.5 En un estudio de casos y controles que se inicia con la identificación de los casos y los controles se pueden estudiar varias exposiciones, pero solo un resultado.

Dado que los estudios de casos y controles a menudo requieren datos sobre eventos o exposiciones previas, suelen enfrentarse a las dificultades que se encuentran al utilizar estos datos (como un posible sesgo de recuerdo). Además, como se ha descrito con cierto detalle, la selección de un grupo control apropiado es uno de los problemas metodológicos más difíciles que se encuentran en epidemiología. Además, en la mayoría de los estudios de casos y controles no es posible calcular la incidencia de la enfermedad, ya sea en la población total, ya sea en los grupos expuesto y no expuesto, sin alguna información suplementaria.

El diseño de casos y controles anidado combina elementos tanto de los estudios de cohortes como de los de casos y controles y ofrece una serie de ventajas. La posibilidad de sesgo de recuerdo se elimina debido a que los datos sobre la exposición se obtienen antes de que aparezca la enfermedad. Es más probable que los datos de exposición representen el estado previo a la enfermedad, ya que se obtienen años antes del diagnóstico de la enfermedad clínica. Por último, los costes son más bajos que con un estudio de cohortes, ya que las pruebas de laboratorio se deben hacer solo en muestras de sujetos que se eligen posteriormente como casos o controles; es decir, solo se deben realizar pruebas de laboratorio de modo selectivo a un subgrupo de la cohorte total, lo que supone un ahorro de costes considerable.

Además de los diseños de los estudios de cohortes y de casos y controles, se ha descrito el diseño de estudio transversal, en el que los datos sobre la exposición y los resultados de la enfermedad de cada sujeto se recogen de forma simultánea. Los datos de un estudio transversal se pueden analizar mediante la comparación de la prevalencia de la enfermedad en los individuos expuestos con la de los individuos no expuestos, o comparando la prevalencia de la exposición en las personas con la enfermedad con la de las personas sin la enfermedad. Aunque los datos transversales se obtienen a menudo mediante encuestas representativas y pueden ser muy útiles, por lo general no permiten que el investigador determine la relación temporal entre la exposición y el desarrollo de la enfermedad. Como resultado de ello, su utilidad para derivar inferencias causales es algo limitada. No obstante, pueden proporcionar indicaciones importantes para futuras investigaciones utilizando diseños de cohortes, de casos y controles y de casos y controles anidados.

CAPÍTULO 10

Valoración de las medidas preventivas y terapéuticas: ensayos aleatorizados

Palabras clave

protocolos
selección de controles
aleatorización
aleatorización estratificada
tratamiento (asignado y recibido)
enmascaramiento
cruce (planificado y no planificado)
diseño factorial
incumplimiento (no adherencia)

Todos los que beben este tratamiento se recuperan en poco tiempo, excepto aquellos a los que no les ayuda, que mueren. Resulta evidente, por tanto, que fracasa únicamente en los casos incurables.

—Galeno[1] (129-c. 199).

Objetivos de aprendizaje

- Describir los elementos importantes de los ensayos aleatorizados.
- Definir el propósito de la aleatorización y el enmascaramiento.
- Introducir aspectos del diseño relacionados con los ensayos aleatorizados, como la aleatorización estratificada, el diseño cruzado planeado y no planeado, y el diseño factorial.
- Ilustrar los problemas planteados por la falta de cumplimiento en los ensayos aleatorizados.

En el capítulo 6 se expusieron algunas formas de cuantificar la historia natural de la enfermedad y de expresar el pronóstico de la enfermedad. Nuestro objetivo, tanto en el ámbito de la salud pública como en la práctica clínica, es modificar la historia natural de la enfermedad con el fin de prevenir o retrasar la muerte o la discapacidad y mejorar la salud del paciente o de la población. El reto es seleccionar las mejores medidas preventivas o terapéuticas para lograr este objetivo. Para ello, necesitamos llevar a cabo estudios que determinen el valor de estas medidas. El ensayo aleatorizado se considera el diseño ideal para evaluar tanto la eficacia como los efectos secundarios de nuevas formas de intervenciones.

La noción de usar una metodología rigurosa para valorar la eficacia de nuevos fármacos o de cualquier nueva modalidad terapéutica no es reciente. En 1883, Sir Francis Galton, antropólogo británico, explorador y eugenista, que tenía un gran interés en la inteligencia humana, escribió lo siguiente:

Hay quien afirma que el hombre posee la facultad de obtener resultados sobre los que tiene poco o ningún control personal por medio de la oración profunda y devota, mientras que otros dudan de la verdad de esta aseveración. La pregunta gira en torno a un hecho que debe ser determinado mediante la observación y no mediante la autoridad; parece ser un tema muy adecuado para un estudio estadístico… ¿Las plegarias reciben respuesta, o no…? ¿Las personas que rezan o por las que se reza se recuperan más rápidamente que el resto?[2]

Al igual que ocurre con muchas ideas pioneras en la ciencia y en la medicina, pasaron muchos años antes de que esta sugerencia fuese realmente puesta en marcha. En 1965, Joyce y Welldon comunicaron los resultados de un ensayo aleatorizado doble ciego sobre la eficacia de la oración[3]. Los hallazgos de este estudio no indicaron que los pacientes por los que se rezaba obtuviesen algún beneficio especial de dicha plegaria. Un estudio más reciente de Byrd[4], sin embargo, evaluó la eficacia de la oración intercesora en una población de una unidad de cuidados coronarios utilizando un protocolo aleatorizado doble ciego. Los hallazgos de este estudio sugirieron que la oración ejercía un efecto terapéutico beneficioso. *¿Cuál es correcto?*

En este capítulo y en el siguiente analizaremos los diseños de estudios que pueden emplearse para evaluar los abordajes terapéuticos y preventivos y nos centraremos en el uso de los ensayos aleatorizados. Aunque el término *ensayo clínico aleatorizado* a menudo se utiliza conjuntamente con su acrónimo, ECA, el diseño de ensayo aleatorizado también tiene aplicaciones importantes en estudios fuera del contexto clínico, como en los ensayos basados en la comunidad. Por este motivo, utilizamos el término *ensayo*

aleatorizado. Para facilitar nuestra exposición, nos referimos generalmente a tratamientos y fármacos; el lector debe tener en cuenta que los principios descritos se aplican igualmente a evaluaciones de medidas preventivas (como programas de cribado para la detección temprana de la enfermedad) y de otro tipo (por ejemplo, intervenciones conductuales). Los ensayos son esencialmente experimentos que están bajo el control del investigador. Esta característica es fundamental al compararlos con los estudios observacionales revisados en el capítulo 7, donde el investigador observa lo que ocurre, pero no interfiere ni controla.

En muchas descripciones anecdóticas de los primeros ensayos se proporcionan recomendaciones de muchos de los elementos que son importantes en los ensayos aleatorizados. En una revisión de la historia de los ensayos clínicos, Bull describió un ensayo accidental realizado por Ambroise Paré (1510-1590), una figura importante en la cirugía durante el Renacimiento[5]. Paré vivió en una época en la que el tratamiento estándar de las heridas de guerra era la aplicación de aceite hirviendo. En 1537 Paré fue el responsable del tratamiento de los heridos tras la conquista del castillo de Villaine. Los heridos eran tan numerosos que él dijo:

Me quedé sin aceite y me vi obligado a aplicar en su lugar un emoliente elaborado con yemas de huevo, aceite de rosas y trementina. Esa noche no pude dormir bien por miedo a que, por la falta de cauterización, los heridos en los que no había utilizado dicho aceite murieran por el veneno. Me levanté pronto para visitarlos y, en contra de lo que esperaba, observé que los que había tratado con el emoliente presentaban un dolor leve, sus heridas no estaban edematosas ni inflamadas y habían dormido toda la noche. Los otros, a los que había aplicado el aceite hirviendo, se encontraban febriles, tenían mucho dolor y sus heridas se encontraban tumefactas. En ese momento tomé la determinación de no quemar nunca de nuevo de ese modo tan cruel a los pobres heridos.

Aunque no se trató de un ensayo aleatorizado, fue una forma de ensayo no intencionado, que ha sido llevado a cabo en muchas ocasiones cuando no se disponía de una cantidad suficiente de un tratamiento que se pensaba que era el mejor disponible y no podía utilizarse en todos los pacientes que lo necesitaban.

Un ensayo planeado fue descrito por el cirujano escocés James Lind en 1747[6]. Lind se interesó en el escorbuto, del que morían miles de marineros británicos cada año. Estaba intrigado por la historia de un marinero que había presentado escorbuto y había sido abandonado en tierra en una isla aislada, donde subsistió gracias a una dieta de plantas y se recuperó del escorbuto. Lind realizó un experimento, que describió del siguiente modo:

Elegí a 12 pacientes con escorbuto que se encontraban navegando en el Salisbury. Elegí los casos tan parecidos como pude… Los mantuve juntos en una parte del barco y les di una dieta igual a todos. Dos de ellos tomaron un cuarto de sidra al día… Otros dos tomaron 25 gotas del elixir vitriolo… Otros dos tomaron dos cucharadas de vinagre… A dos se les dio a beber agua del mar… Otros dos recibieron dos naranjas y un limón al día… Otros dos tomaron una nuez moscada. Los efectos beneficiosos más rápidos y visibles se notaron en los que comieron naranjas y limones. Uno de ellos estaba apto para el trabajo al final del sexto día… El otro se convirtió en el enfermero del resto de los enfermos.

Resulta interesante que la idea de la causa dietética del escorbuto no fue aceptada en la época de Lind. Solo 47 años después, el Almirantazgo británico permitió repetir el experimento; esta vez, en una flota completa de navíos. Los resultados fueron tan espectaculares que, en 1795, el Ministerio de Marina ordenó que el zumo de limón fuese una parte necesaria en la dieta estándar de los marineros británicos; posteriormente se cambió por zumo de lima. El escorbuto prácticamente desapareció entre los marineros británicos, que, incluso hoy en día, son conocidos como los «limas».

Los ensayos aleatorizados pueden emplearse para muchos fines. Pueden usarse para evaluar nuevos fármacos y otros tratamientos de enfermedades, incluidas pruebas de nuevas técnicas de salud y asistencia sanitaria. Los ensayos también pueden utilizarse para valorar nuevos programas de cribado y detección precoz, comparar diferentes medidas preventivas o nuevos métodos para organizar y proporcionar servicios sanitarios.

El diseño básico de un ensayo aleatorizado se muestra en la figura 10.1.



FIG. 10.1 Diseño de un ensayo aleatorizado.

Comenzamos con una población definida en la que los participantes son aleatorizados a recibir un tratamiento nuevo o el tratamiento habitual y a continuación seguimos a los individuos de cada grupo para determinar cuántos mejoran en el grupo que recibe el tratamiento nuevo en comparación con los que mejoran en el grupo que recibe el tratamiento habitual (a menudo referido como «tratamiento habitual» o «estándar»). Si el tratamiento nuevo consigue un mejor resultado, esperaríamos encontrar un mejor resultado en una mayor proporción en el grupo que recibe el tratamiento nuevo que en el grupo que recibe el tratamiento habitual.

Podemos elegir comparar dos grupos que reciben tratamientos diferentes o podemos comparar más de dos grupos. Aunque, en ocasiones, puede compararse un tratamiento nuevo con individuos que no reciben tratamiento, a menudo se adopta la decisión de no utilizar un grupo no tratado. Por ejemplo, si quisiéramos evaluar un nuevo tratamiento del síndrome de inmunodeficiencia adquirida (SIDA), ¿estaríamos dispuestos a contar en nuestro estudio con un grupo de pacientes con SIDA no tratados? La respuesta claramente es no; lo que querríamos es comparar el nuevo tratamiento con un tratamiento recomendado en la actualidad, que claramente sería mejor que no administrar ningún tratamiento.

Fijémonos ahora en algunos de los aspectos que deben considerarse en el diseño de los ensayos aleatorizados. El principal de ellos es la especificación de los «brazos» o tratamientos del estudio. Estos deben estar claramente establecidos con criterios para su medición, así como la duración de los tratamientos y la duración del estudio. Primero, comencemos con quién reúne los requisitos para ser estudiado.

Selección de sujetos

Los criterios para determinar quién será incluido o no en el estudio deben ser determinados con gran precisión; se deben *exponer por escrito antes de que comience el estudio*. Una prueba excelente de la idoneidad de estos criterios escritos es preguntarnos: ¿si hemos expuesto nuestros criterios por escrito, y alguien que no participe en el estudio llega de nuevas y aplica nuestros criterios a la misma población, seleccionará esa persona a los mismos sujetos que nosotros habríamos seleccionado? No deben existir elementos subjetivos en la toma de decisiones por parte del investigador a la hora de elegir quién es incluido y quién no en el estudio. Todo estudio debe, en principio, ser replicable por otros, igual que ocurre con los experimentos de laboratorio. Claramente, esto es más fácil de decir que de hacer, porque en los ensayos aleatorizados a menudo tratamos con poblaciones relativamente numerosas. Sin embargo, este principio es importante, y, por tanto, los criterios de selección deben ser establecidos con precisión.

Asignación de sujetos a grupos de tratamiento sin aleatorización

Antes de exponer el proceso de aleatorización, preguntémonos si podríamos utilizar algunas alternativas a la aleatorización.

Estudios sin comparación

La primera alternativa posible es el *estudio de casos* o las *series de casos* (como presentamos en el cap. 7). En este tipo de estudio no se realizan comparaciones con un grupo no tratado o con un grupo que recibe algún otro tratamiento. La siguiente historia fue relatada por el Dr. Earl Peacock cuando era jefe del departamento de cirugía de la Universidad de Arizona:

Un día, cuando era estudiante de tercero de Medicina, un cirujano muy importante de Boston visitó la facultad y presentó un estupendo estudio sobre un gran número de pacientes que habían sido sometidos a intervenciones de reconstrucción vascular exitosas. Al final de la conferencia, un estudiante joven desde el fondo de la sala preguntó tímidamente: «¿Tiene un grupo control?». El ilustre cirujano se puso en pie, golpeó el atril y dijo: «¿Me está preguntando que si dejé sin operar a la mitad de los pacientes?». La sala enmudeció. La voz desde el fondo de la sala replicó con vacilación: «Sí, a eso me refería». Acto seguido el puño del conferenciante golpeó de nuevo el atril y este exclamó: «Por supuesto que no, eso habría condenado a muerte a la mitad de ellos». Ahora sí que se hizo el silencio y apenas se pudo escuchar la pregunta de la vocecilla: «¿A qué mitad?».

El asunto de las comparaciones es importante porque queremos ser capaces de extraer una inferencia causal respecto a la relación de un tratamiento y el consiguiente resultado. El problema de inferir una relación causal de una secuencia de eventos sin establecer una comparación queda demostrado en una historia citada por Ederer[8].

Durante la Segunda Guerra Mundial, el equipo de rescate, cavando en las ruinas de un piso destruido en el bombardeo de Londres, encontró a un anciano desnudo en la bañera, completamente consciente. El anciano dijo a sus rescatadores: «¿Saben?, esta ha sido la experiencia más asombrosa que he vivido. Cuando quité el tapón y el agua comenzó a drenar, estalló toda la casa.»

El problema ejemplificado en esta historia es este: si administramos un fármaco y el paciente mejora, ¿podemos atribuir la mejoría a la administración de ese fármaco? El catedrático Hugo Muensch, de la Universidad de Harvard, enunció su segunda ley: «Los resultados siempre pueden mejorarse omitiendo los controles.»

Estudios con comparación

Si reconocemos por tanto la necesidad de que nuestro estudio incluya alguna forma de comparación, ¿cuáles son los diseños posibles?

Controles históricos

Podríamos emplear un grupo de comparación del pasado, denominado *controles históricos*. En la actualidad contamos con un tratamiento que creemos que es bastante efectivo, nos gustaría probarlo en un grupo de pacientes; sabemos que necesitamos un grupo para realizar comparaciones. Por tanto, con el fin de realizar comparaciones, recurriremos a los historiales médicos de pacientes con la misma enfermedad que fueron tratados antes de que este nuevo tratamiento estuviese disponible. Este tipo de diseño parece intrínsecamente sencillo y atractivo.

¿Cuáles son los problemas del uso de controles históricos? En primer lugar, si decidimos realizar hoy el estudio que acabamos de describir, podemos establecer un sistema muy meticuloso de recogida de datos de los pacientes que están siendo tratados en la actualidad. Pero, por supuesto, no podemos hacer lo mismo en los pacientes tratados en el pasado, de los que debemos obtener datos imprecisos a partir de sus historiales médicos, que son probablemente útiles para administrar la atención individual, pero están llenos de errores y omisiones cuando se utilizan para fines de investigación. Por tanto, si al final del estudio encontramos una diferencia en el resultado entre los pacientes tratados en el periodo inicial (controles históricos) y los pacientes tratados en el periodo tardío (en el presente), no sabremos si existen diferencias reales en el resultado o si las diferencias observadas se deben únicamente a diferencias en la calidad de la recogida de datos. Los datos obtenidos de los grupos de estudio deben ser comparables en clase y calidad; en los estudios que utilizan controles históricos, a menudo no es el caso.

El segundo problema es que, si observamos una diferencia en el resultado entre el grupo inicial y el grupo tardío, no estaremos seguros de que la diferencia se deba al tratamiento, porque a lo largo del tiempo cambian muchas cosas aparte del propio tratamiento (p. ej., terapias de soporte, condiciones de vida y estilos de vida). Esto a menudo se conoce como «cambios seculares». Por tanto, si observamos una diferencia y hemos descartado que se deba a la calidad de los datos, no puede saberse si esa diferencia es el resultado del fármaco que estamos estudiando o de otros

cambios en muchos otros factores que pueden relacionarse con el resultado a lo largo del tiempo.

En ocasiones, sin embargo, este tipo de diseño puede resultar útil. Por ejemplo, cuando una enfermedad es uniformemente mortal y aparece un nuevo fármaco, la disminución de la tasa de letalidad paralela al uso del fármaco apoyaría con fuerza la conclusión de que el nuevo fármaco está produciendo un efecto. Los ejemplos incluyen el descubrimiento de la insulina para tratar la diabetes, de la penicilina para tratar infecciones graves y de los inhibidores de la tirosina cinasa (ITK), como el imatinib, para tratar la leucemia mieloide crónica. No obstante, aun así, deberíamos descartar la posibilidad de que la disminución pueda deberse a otros cambios en el entorno.

Controles simultáneos no aleatorizados

Debido a la importancia de los problemas encontrados con los controles históricos y las dificultades de tener en cuenta los cambios a lo largo del tiempo, un abordaje alternativo es el uso de controles simultáneos que no son seleccionados aleatoriamente. El problema de seleccionar controles simultáneos de modo no aleatorio se ilustra en la siguiente historia:

A un capitán de barco se le suministraron muestras de píldoras antináuseas para probarlas durante una travesía. Se le explicó detenidamente la necesidad de contar con controles. Tras el retorno del barco, el capitán comunicó los resultados con entusiasmo: «Prácticamente todos los controles sufrieron náuseas, y ninguno de los sujetos tuvo ningún problema. Es una píldora realmente maravillosa.» Un escéptico preguntó cómo había elegido a los controles y a los sujetos. «Pues les di la píldora a mis marineros y usé a los pasajeros como controles»[10].

Existen diferentes métodos posibles para elegir a los controles de modo no aleatorizado. Uno es asignar a los pacientes según el día del mes en el que el paciente es ingresado en el hospital: por ejemplo, si el paciente ingresa en un día impar, se le clasifica en el grupo A, y si ingresa en un día par, se le clasifica en el grupo B. En un ensayo sobre tratamientos anticoagulantes realizado tras la Segunda Guerra Mundial se empleó este método en función del día del mes y se observó que los días impares ingresaban más pacientes de lo esperado. Los investigadores comunicaron que «a medida que los médicos observaban los beneficios del tratamiento anticoagulante, aceleraban, dentro de lo posible, la hospitalización de dichos pacientes… con el fin de reunir a todos los que fuera posible dentro de la fecha límite de los días impares, ya que de otro modo hubiesen sido hospitalizados en un día par»[11].

El problema en este caso es que el sistema de asignación era predecible: era posible que el médico supiera cuál iba a ser la asignación del próximo paciente. El objetivo de la aleatorización es eliminar la posibilidad de que el investigador sepa cuál será la asignación del próximo paciente, porque dicho conocimiento introduce la posibilidad de sesgos por parte del investigador en lo que respecta al grupo de tratamiento al que el paciente será asignado.

Hace muchos años se llevó a cabo un estudio sobre los efectos de la vacunación con el bacilo de Calmette-Guérin (BCG) frente a la tuberculosis en niños de familias con tuberculosis en la ciudad de Nueva York[12]. Se dio instrucciones a los médicos para que dividiesen el grupo de niños elegibles en un grupo que iba a ser vacunado y en un grupo control no vacunado para comparar.

Como se observa en la tabla 10.1, la mortalidad por tuberculosis fue casi cinco veces superior en los controles que en los niños vacunados. Sin embargo, como escribieron los investigadores:

La experiencia posterior ha demostrado que, utilizando este método de selección, la tendencia era vacunar a los niños de los padres más inteligentes y colaboradores y usar como controles a los hijos de los padres no colaboradores. Este proceder probablemente era un error grave, ya que los padres colaboradores no solo adoptarían más precauciones, sino que por lo general traerían a los niños a consulta con más regularidad para recibir instrucciones sobre los cuidados y la alimentación de los niños[12].

Tabla 10.1

Resultados de un ensayo de vacunación con el bacilo de Calmette-Guérin (BCG): I

		MUERTES POR TUBERCULOSIS	
	Número de niños	Número	%
Vacunados	445	3	0,67
Controles	545	18	3,30

Datos de Levine MI, Sackett MF. Results of BCG immunization in New York City. *Am Rev Tuberculosis*. 1946;53:517–532.

Tras reconocer que las vacunaciones se estaban realizando selectivamente en niños de familias más preocupadas por la salud y los temas relacionados, los investigadores se dieron cuenta de que era posible que la tasa de mortalidad por tuberculosis fuese inferior en el grupo vacunado no por la propia vacunación sino porque estos niños fueron seleccionados de familias

más preocupadas por la salud, que presentaban un menor riesgo de mortalidad por tuberculosis, con o sin vacunación. Para abordar este problema, se realizó un cambio en el diseño del estudio: se vacunaron niños alternos y el resto servían de controles. Este método seguía sin ser una aleatorización, pero mejoraba mucho el diseño inicial. Como se muestra en la tabla 10.2, no se observaron diferencias entre los grupos.

Tabla 10.2

Resultados de un ensayo de vacunación con el bacilo de Calmette-Guérin (BCG): II

		MUERTES POR TUBERCULOSIS	
	Número de niños	Número	%
Vacunados	556	8	1,44
Controles	528	8	1,52

Datos de Levine MI, Sackett MF. Results of BCG immunization in New York City. *Am Rev Tuberculosis*. 1946;53:517–532.

Asignación de sujetos utilizando aleatorización

En vista de los problemas expuestos, la aleatorización es el mejor método de diseño de un ensayo. La aleatorización significa, de hecho, decidir la asignación de un paciente a un grupo del estudio u otro echándolo a suertes. El elemento crítico de la aleatorización es la impredictibilidad de la próxima asignación. En la figura 10.2 se muestra una tira cómica citada por Ederer para demostrar el problema de la predictibilidad de la siguiente asignación[13].

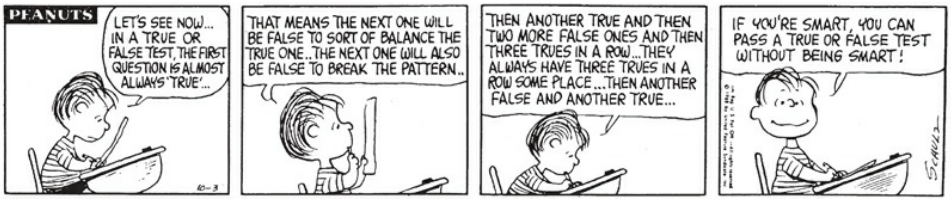


FIG. 10.2 Cómo predecir la asignación de tratamiento del siguiente paciente en un estudio aleatorizado. (PEANUTS © UFS. Reproducido con autorización.)

¿Cómo se lleva a cabo la aleatorización? Aunque en la actualidad la asignación aleatoria generalmente se realiza a través de programas informáticos, en ocasiones la asignación aleatoria manual se utiliza como respaldo de la asignación de generación informatizada o cuando el acceso a un ordenador es limitado. En este ejemplo hipotético de asignación manual utilizamos una selección de una tabla de números aleatorios (tabla 10.3). (Estas tablas de números aleatorios se encuentran disponibles en un apéndice en la mayoría de los libros de texto sobre estadística o pueden generarse en ordenadores.)

Tabla 10.3

Una tabla de números aleatorios

	00-04	05-09	10-14	15-19
00	56348	01458	36236	07253
01	09372	27651	30103	37004
02	44782	54023	61355	71692
03	04383	90952	57204	57810
04	98190	89997	98839	76129
05	16263	35632	88105	59090
06	62032	90741	13468	02647
07	48457	78538	22759	12188
08	36782	06157	73084	48094
09	63302	55103	19703	74741

En primer lugar, ¿cómo se lee la tabla 10.3? Obsérvese que la tabla se divide en 10 filas y 4 columnas numeradas (los números de fila aparecen en la primera columna de la izquierda). Las columnas se numeran en su parte superior: 00-04, 05-09, etc. Esto significa que el número en la columna 00 es el 5, el número en la columna 01 es el 6, el número en la columna 03 es el 3, etc. De modo similar, las filas se numeran en la izquierda: 00, 01, 02, etc. De este modo podemos referirnos a cualquier cifra de la tabla empleando los números de su fila y su columna. Esto es importante si alguien ajeno al estudio quiere comprobar la calidad del proceso de aleatorización. ¿Cómo se utiliza esta tabla? Supongamos que estamos realizando un estudio en el que habrá dos grupos: tratamiento A y tratamiento B. En este ejemplo, consideraremos cada número impar una asignación al grupo A y cada número par una asignación al grupo B. Cerramos los ojos y ponemos un dedo en cualquier parte de la tabla y escribimos el número de la fila y la columna que será nuestro punto de comienzo. También escribimos la dirección en la que nos moveremos en la tabla a partir del punto de comienzo (horizontalmente a la derecha, horizontalmente a la izquierda, hacia arriba o hacia abajo). Asumamos que apuntamos al «5» en la intersección de la columna 7 y la fila 7 y nos movemos horizontalmente a la derecha. El primer paciente, por tanto, es designado por un número impar, 5, y recibirá el tratamiento A. El segundo paciente también es designado por un número impar, 3, y recibirá el tratamiento A. El tercero es designado por un número par, 8, y recibirá el tratamiento B, y así sucesivamente. Obsérvese que la asignación del siguiente paciente no es predecible; *no* es una alternancia estricta, que sería predecible y, por lo tanto, sujeta a sesgo del investigador, a sabiendas o sin saberlo.

Una tabla de números aleatorios puede usarse de muchas formas para asignar a los pacientes a grupos de tratamiento en un ensayo aleatorizado (cuadro 10.1). Aunque existen muchos abordajes válidos, el punto importante es determinar por escrito el método elegido que se va a utilizar antes de comenzar en la práctica el proceso de aleatorización.

Cuadro 10.1 Ejemplos de usos de una tabla de números

aleatorios para distribuir a los pacientes en grupos de tratamiento en un ensayo aleatorizado

Si planeamos comparar dos grupos:

- Decidimos que los números pares designan el tratamiento A, los números impares designan el tratamiento B, *o*
- Decidimos que los números 0 a 4 designan el tratamiento A, los números 5 a 9 designan el tratamiento B.

Si planeamos comparar tres grupos:

- Decidimos que los números 1 a 3 designan el tratamiento A, los números 4 a 6 designan el tratamiento B, los números 7 a 9 designan el tratamiento C y el número 0 se ignora.

Tras decidir conceptualmente cómo usar los números aleatorios para asignar pacientes, ¿cómo tomamos una decisión práctica sobre qué pacientes reciben qué tratamiento? Asumamos, por ejemplo, que se ha decidido que los números impares designan la asignación al tratamiento A y los números pares designan el tratamiento B. La asignación al tratamiento, que es designado por el número aleatorio, se escribe en una tarjeta, y esta tarjeta se guarda dentro de un sobre opaco. Cada sobre se etiqueta en el exterior: paciente 1, paciente 2, paciente 3, y así sucesivamente, para indicar la secuencia en la que los pacientes serán incorporados al estudio. Por ejemplo, si el primer número aleatorio es el 2, en el primer sobre se introducirá una tarjeta de tratamiento B; si el siguiente número aleatorio es el 7, en el segundo sobre se introduce una tarjeta de tratamiento A, y así sucesivamente, según determinen los números aleatorios.

A continuación, se sellan los sobres. Cuando se incorpora el primer paciente, se abre el sobre 1 y se lee la asignación, y este proceso se repite para cada uno de los siguientes pacientes del estudio.

No obstante, este proceso no es infalible. La siguiente anécdota ilustra la necesidad de un control de calidad riguroso en todo estudio aleatorizado:

En un estudio aleatorizado que comparaba la mastectomía radical y la mastectomía simple en el cáncer de mama, uno de los cirujanos participantes estaba convencido de que la mastectomía radical era el tratamiento de elección y no podía permitirse realizar una mastectomía simple a ninguna de las pacientes incluidas en el estudio. Cuando se llevó a cabo la aleatorización de sus pacientes y abrió un sobre que indicaba una mastectomía simple para la siguiente asignación, apartó el sobre y siguió abriendo sobres hasta que encontró uno en el que la asignación era una mastectomía radical.

Esta historia refleja el conflicto que sufren muchos médicos clínicos que

incorporan a sus propios pacientes en ensayos aleatorizados. Por un lado, el médico tiene la obligación de hacer lo mejor para el paciente pero, por otro lado, cuando un médico clínico participa en un ensayo clínico, se le pide en realidad que se aparte de su papel habitual de toma de decisiones y que esencialmente «lance una moneda al aire» para decidir qué tratamiento recibirá el paciente. Por tanto, con frecuencia se produce un conflicto subyacente entre el papel clínico del médico y el papel del médico que participa en un ensayo clínico, y, como resultado, se pueden producir sesgos no intencionados.

Este es un problema común, especialmente en los ensayos multicéntricos de gran tamaño, en los que la aleatorización no se lleva a cabo en cada centro clínico participante, sino que se efectúa en un centro estadístico imparcial separado, encargado de la coordinación. Cuando un nuevo paciente es registrado en un centro clínico, se contacta con el centro coordinador o se descarga una asignación del centro coordinador. A continuación, el centro coordinador realiza la asignación del paciente, lo que queda registrado en ambos centros, el clínico y el centralizado.

¿Qué esperamos lograr con la aleatorización? Si realizamos la aleatorización correctamente, logramos que la siguiente asignación sea impredecible; no tenemos que preocuparnos acerca de qué sesgos subjetivos por parte de los investigadores, ya sean manifiestos o no, influyan en el proceso de selección de pacientes para un grupo de tratamiento u otro. Del mismo modo, si el estudio es lo suficientemente extenso y existen suficientes participantes, queremos que la aleatorización aumente la probabilidad de que los grupos sean comparables entre sí con respecto a características que nos interesan, como el sexo, la edad, la raza y la gravedad de la enfermedad, todos ellos factores que pueden afectar el pronóstico. La aleatorización no garantiza la comparabilidad, ya que el azar puede jugar un papel en el proceso de asignación aleatoria de tratamientos. Sin embargo, si los grupos de tratamiento que están siendo aleatorizados son lo suficientemente grandes y el proceso de aleatorización está libre de sesgos, la tendencia será que sean similares.

En la figura 10.3 se presenta un ejemplo hipotético del efecto de la falta de comparabilidad en una comparación de tasas de mortalidad en los grupos que están siendo estudiados. Asumamos una población de estudio de 2.000 pacientes con infarto de miocardio, de los cuales la mitad reciben una intervención y la otra mitad no. Sigamos suponiendo que, de los 2.000 pacientes, 700 presentan una arritmia y 1.300 no la presentan. La tasa de letalidad en los pacientes con arritmia es del 50% y en los pacientes sin arritmia, del 10%.

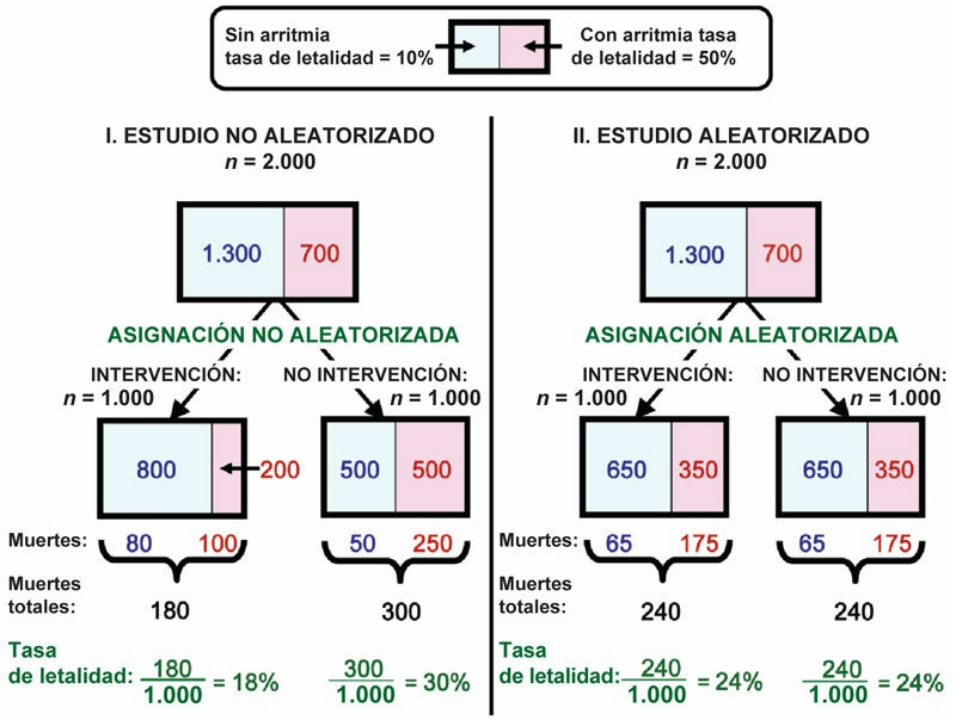


FIG. 10.3 Estudios aleatorizados frente a no aleatorizados. *I,* Si el estudio no es aleatorizado, las proporciones de pacientes con arritmia en los dos grupos de intervención pueden ser diferentes. En este ejemplo, es menos probable que los pacientes con arritmia reciban la intervención que los sujetos sin arritmias. *II,* Si el estudio es aleatorizado, es más probable que las proporciones de pacientes con arritmia en los dos grupos de intervención sean similares.

Fijémonos en el estudio no aleatorizado de la parte izquierda de la figura 10.3. Como no existe aleatorización, los grupos de intervención pueden no ser comparables en la proporción de pacientes que tienen arritmia. Quizá 200 pacientes del grupo de intervención pueden presentar arritmia (con una tasa de letalidad del 50%) y 500 en el grupo de no intervención pueden presentar arritmia (con su tasa de letalidad del 50%). La tasa de letalidad resultante será del 18% en el grupo de intervención y del 30% en el grupo sin intervención. Podríamos estar tentados por concluir que la intervención es más efectiva que la no intervención.

Sin embargo, fijémonos ahora en el estudio aleatorizado de la parte derecha de la figura 10.3. Como se observa aquí, los grupos son comparables, como es probable que ocurra cuando aleatorizamos, de modo que 350 de los 1.000 pacientes en el grupo de intervención y 350 de los 1.000 pacientes en el grupo sin intervención presentan la arritmia. Cuando se calcula la tasa de letalidad en este ejemplo, el resultado es un 24% en ambos grupos. Así, la diferencia observada entre la intervención o la no intervención cuando los grupos no eran comparables respecto a la arritmia se debía enteramente a la no comparabilidad y no a los efectos de la intervención propiamente dicha. (Téngase en cuenta que, aunque en la figura 10.3 se muestran 1.000 pacientes en el grupo de intervención y 1.000 en el de no intervención, la aleatorización *no* garantiza un número igual de participantes en cada grupo; sin embargo, con grandes números, en promedio los dos grupos serán generalmente

comparables.)

Nos podríamos preguntar: si estamos tan preocupados por la comparabilidad de los grupos, ¿por qué no emparejar simplemente los grupos en función de las variables específicas sobre las que estamos interesados en vez de mediante la aleatorización? La respuesta es que solo podemos emparejar variables que conocemos y que podemos medir. Por tanto, no podemos emparejar muchas variables que pueden afectar el pronóstico, como la carga genética del individuo, elementos del estado inmune del sujeto u otras variables de las cuales puede que no seamos conscientes. Además, si emparejamos respecto a una característica particular, no podemos analizar su asociación con el resultado porque los dos grupos ya serán idénticos. La aleatorización aumenta la probabilidad de que los grupos sean comparables, no solo en cuanto a las variables que reconocemos y podemos medir, sino también en cuanto a las variables que no podemos reconocer, que no podemos medir en la actualidad y que no somos capaces de medir con la tecnología presente. Sin embargo, la aleatorización no siempre garantiza la comparabilidad de los grupos que se estudian. Podemos analizar si hay diferencias importantes entre los dos grupos que pueden estar asociadas con el resultado del ensayo.

¿Cuál es el Principal Objetivo de la Aleatorización?

El principal objetivo de la aleatorización es evitar sesgos potenciales por parte de los investigadores que pudieran influir en la asignación de los participantes a los diferentes grupos de tratamiento. Cuando los participantes se asignan aleatoriamente a diferentes grupos de tratamiento, todas las decisiones sobre la asignación del tratamiento se dejan fuera del control de los investigadores. Así, el uso de la aleatorización es crucial para proteger el estudio de sesgos que pudieran ser introducidos consciente o inconscientemente por el investigador en el proceso de asignación.

Como se ha mencionado anteriormente, aunque la aleatorización a menudo aumenta la comparabilidad de los diferentes grupos de tratamiento, la aleatorización no garantiza la comparabilidad. Otro beneficio de la aleatorización es que, con independencia del grado con el que contribuya a la comparabilidad, esta contribución se aplica tanto a variables que pueden medirse como a variables que no podemos medir y puede que incluso no seamos conscientes de ellas, incluso aunque pueden ser importantes para interpretar los hallazgos del ensayo.

Aleatorización Estratificada

En ocasiones podemos estar especialmente interesados en la comparabilidad de los grupos en cuanto a una o algunas características importantes que pensamos firmemente que pueden influir en el pronóstico o la respuesta al

tratamiento de los grupos que se están estudiando. No obstante, como acabamos de exponer, la aleatorización no asegura la comparabilidad. Una opción que puede utilizarse es la *aleatorización estratificada*, un método de asignación que puede resultar muy útil para aumentar la probabilidad de comparabilidad de los grupos del estudio. En esta sección mostraremos cómo se utiliza este método para asignar participantes a diferentes grupos de estudio.

Por ejemplo, supongamos que estamos especialmente interesados en la edad como variable pronóstica: el pronóstico es mucho peor en los pacientes de edad avanzada que en los más jóvenes. Por tanto, estamos interesados en que los dos grupos de tratamiento sean comparables en cuanto a la edad. Aunque uno de los beneficios de la aleatorización es que puede aumentar la probabilidad de dicha comparabilidad, no la garantiza. Sigue siendo posible que tras la aleatorización podemos encontrar, por azar, que la mayoría de los pacientes de edad avanzada se encuentran en un grupo y la mayoría de los pacientes más jóvenes se encuentran en otro grupo. Nuestros resultados serían por tanto imposibles de interpretar, porque los pacientes de riesgo más elevado se encontrarían agrupados en un grupo y los pacientes de menor riesgo en el otro. Cualquier diferencia en el resultado entre los grupos de intervención puede ser atribuida entonces a esta diferencia en la distribución de edades de los dos grupos en vez de a los efectos de la intervención.

En la *aleatorización estratificada*, en primer lugar, estratificamos (estrato = capa) nuestra población de estudio por cada variable que consideremos importante y, a continuación, aleatorizamos a los participantes en los grupos de tratamiento dentro de cada estrato.

Consideremos el ejemplo que se muestra en la figura 10.4. Estamos estudiando 1.000 pacientes y creemos que el sexo y la edad son determinantes importantes del pronóstico. Si aleatorizamos, no sabemos cuál puede ser la composición de los grupos en cuanto a sexo y edad; por tanto, decidimos emplear la aleatorización estratificada.

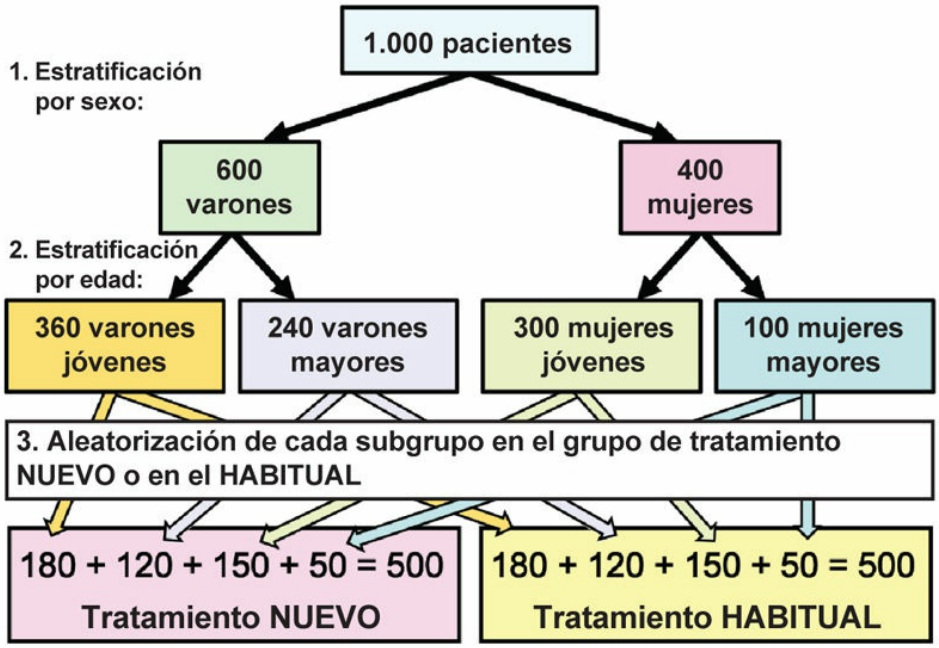


FIG. 10.4 Ejemplo de aleatorización estratificada. Véase la explicación en el texto.

En primer lugar, estratificamos los 1.000 pacientes por sexo en 600 varones y 400 mujeres. A continuación, estratificamos separadamente a los varones y a las mujeres por edad. Ahora tenemos cuatro grupos (estratos): varones jóvenes, varones de edad avanzada, mujeres jóvenes y mujeres de edad avanzada. Ahora aleatorizamos *dentro de cada grupo (estratos)* y el resultado es un grupo que recibe un tratamiento nuevo y un grupo que recibe el tratamiento habitual en cada uno de los cuatro grupos. Igual que en la aleatorización sin estratificación, acabamos con dos grupos de intervención, pero, al haber estratificado inicialmente los grupos, aumentamos la probabilidad de que los dos grupos sean comparables en cuanto a sexo y edad. (Igual que en la figura 10.3, en la figura 10.4 se muestra que la aleatorización da lugar a un mismo número de participantes en cada grupo de tratamiento, aunque este resultado no está garantizado por la aleatorización.)

Recogida de datos sobre los sujetos

Como se ha mencionado antes, resulta fundamental que los datos obtenidos para cada uno de los grupos del estudio sean de la misma calidad. No queremos que las diferencias en los resultados entre los grupos se deban a diferencias en la calidad o en la compleción de los datos recogidos en los grupos de estudio. Consideremos algunas de las variables sobre qué datos deben obtenerse de los sujetos.

Tratamiento (asignado y recibido)

¿Qué datos son necesarios? En primer lugar, debemos saber a qué grupo de tratamiento fue asignado el paciente. Además, debemos saber qué tratamiento recibió el paciente en realidad. Es importante conocer, por ejemplo, si el paciente fue asignado a recibir el tratamiento A, pero no lo cumplió. Un paciente puede aceptar el proceso de aleatorización, pero posteriormente puede cambiar de idea y negarse a cumplir el tratamiento. Por el contrario, también es muy importante saber si un paciente que no fue asignado a recibir el tratamiento A puede haber tomado dicho tratamiento por su cuenta, a menudo sin conocimiento de los investigadores.

Resultado

La necesidad de medidas comparables en todos los grupos de estudio es especialmente importante en las medidas de los resultados. Entre dichas medidas se incluyen tanto la mejoría (el efecto deseado) como cualquier efecto secundario que pudiera aparecer. Por tanto, es necesario establecer de modo explícito los criterios para todos los resultados que vayan a medirse en el estudio. Tras establecer los criterios explícitamente, debemos estar seguros de que son medidos de modo comparable en todos los grupos del estudio. En especial, debe evitarse el problema potencial de medir los resultados más cuidadosamente en los que reciben un tratamiento nuevo que en los que reciben el tratamiento disponible habitualmente. El enmascaramiento, que se expone más adelante, puede evitar en gran parte este problema; no obstante, como el enmascaramiento no siempre es posible, se debe prestar atención para asegurar la comparabilidad de las medidas y la calidad de los datos en todos los grupos del estudio.

Resultado de mortalidad por todas las causas («resultado de salud pública»)

En ocasiones, un medicamento o una estrategia preventiva para la mortalidad que sea eficaz con respecto al principal resultado de interés no aumenta la supervivencia libre de eventos. Por ejemplo, en los 13 años de seguimiento

del European Randomized Study of Screening for Prostate Cancer, se observó una reducción de aproximadamente el 27% en la mortalidad por cáncer de próstata[14]. Sin embargo, la mortalidad general (también conocida como «resultado de salud pública») fue similar en los dos grupos de estudio, lo que sugiere que la efectividad del cribado con respecto a la mortalidad por todas las causas fue nula.

Perfil pronóstico en la incorporación

Si conocemos los factores de riesgo de un mal resultado, queremos verificar que la aleatorización ha logrado un parecido razonable entre los dos grupos en relación con estos factores de riesgo. Por ejemplo, si la edad es un factor de riesgo importante, querríamos que la aleatorización lograse grupos que son comparables en edad. Se deben obtener datos sobre factores pronósticos en el momento de incorporación del sujeto al estudio y, a continuación, los dos (o más) grupos pueden compararse respecto a estos factores al inicio (es decir, antes de que se proporcione el tratamiento). Otra estrategia para evaluar la comparabilidad es examinar un resultado no relacionado con el tratamiento que se está evaluando. Por ejemplo, si el objetivo del ensayo aleatorizado es evaluar un nuevo medicamento para las migrañas, se espera que la mortalidad por cáncer sea similar en los dos grupos.

Enmascaramiento

El enmascaramiento implica la participación de varios componentes: en primer lugar, queremos que los sujetos no sepan a qué grupo son asignados. Este aspecto es de gran importancia cuando el resultado es una medida subjetiva, como la gravedad autoinformada de la cefalea o la lumbalgia. Si el paciente sabe que está recibiendo un tratamiento nuevo, el entusiasmo y ciertos factores psicológicos por parte del paciente pueden entrar en juego y suscitar una respuesta positiva incluso aunque el tratamiento por sí mismo no haya producido ningún efecto biológico o clínico positivo.

¿Cómo podemos enmascarar a los sujetos? Un método es empleando un *placebo*, una sustancia inerte que parece, sabe y huele como el agente activo. Sin embargo, el uso de un placebo no garantiza automáticamente que los pacientes están enmascarados. Algunos participantes pueden intentar determinar si están tomando el placebo o el fármaco activo. Por ejemplo, en un ensayo aleatorizado sobre la vitamina C para el resfriado común, los pacientes fueron enmascarados mediante el uso de placebo y posteriormente se les preguntó si sabían o sospechaban qué tratamiento estaban tomando.

Como se observa en la tabla 10.4, de los 52 pacientes que estaban tomando vitamina C y a los que se pidió que dijesen el agente que creían estar tomando, 40 dijeron que estaban tomando vitamina C. De los 50 que estaban recibiendo placebo, 39 dijeron que estaban recibiendo placebo. ¿Cómo lo sabían? Habían mordido la cápsula y lo averiguaron por el sabor amargo.

¿Qué supone el hecho de que lo supieran? Los datos sugieren que el número de resfriados fue más elevado en los sujetos que recibían vitamina C creyendo que era placebo que en los sujetos que tomaban el placebo creyendo que tomaban vitamina C. Por tanto, debemos prestar gran atención a la falta de enmascaramiento de los sujetos y los efectos potenciales sobre los resultados del estudio, en especial cuando tratamos con puntos finales subjetivos.

Tabla 10.4

Ensayo aleatorizado sobre vitamina C y placebo para el resfriado común: resultados de un estudio con cuestionario para determinar si los participantes sospechaban qué agente se les había dado

	FÁRMACO SOSPECHADO		
Fármaco real	**Vitamina C**	**Placebo**	**Total**
Vitamina C	40	12	52
Placebo	11	39	50
Total	51	51	102

P <0,001.

De Karlowski TR, Chalmers TC, Frenkel LD, et al. Ascorbic acid for the common cold. A prophylactic and therapeutic trial. JAMA 1975;231:1038. Copyright 1975, American Medical Association.

El uso de un placebo también es importante para estudiar la frecuencia de reacciones y efectos secundarios. El Physician's Health Study era un ensayo aleatorizado acerca del uso de aspirina para prevenir el infarto de miocardio. En la tabla 10.5 se muestran los efectos secundarios que fueron comunicados en los grupos que recibieron aspirina y en los que recibieron placebo en el estudio.

Tabla 10.5

Physicians' Health Study: efectos secundarios según el grupo de tratamiento

Efecto secundario	Grupo de aspirina (%)	Grupo de placebo (%)	*P*
Síntomas GI (excepto úlcera)	34,8	34,2	0,48
Úlceras en el tracto GI superior	1,5	1,3	0,08
Problemas hemorrágicos	27,0	20,4	<0,0001

GI, gastrointestinales.

Datos de Steering Committee of the Physicians' Health Study Research

Group: Final report on the aspirin component of the Ongoing Physicians' Health Study. *N Engl J Med.* 1989;321:129–135. Copyright 1989, Massachusetts Medical Society. Reservados todos los derechos.

Obsérvese la gran cantidad de reacciones comunicadas en el grupo que recibió placebo. Por tanto, no es suficiente decir que el 34% de las personas que recibieron aspirina sufrieron síntomas gastrointestinales; lo que realmente queremos saber es en qué grado el riesgo de efectos secundarios es mayor en las personas que toman aspirina en comparación con las que no la toman (es decir, los que toman placebo). Así pues, el placebo desempeña un papel importante en la identificación de los beneficios reales de un agente y de sus efectos secundarios. A veces es posible usar un medicamento tanto en el grupo asignado al nuevo tratamiento como en el grupo de placebo para evitar la aparición de los efectos secundarios más evidentes de la terapia. En el ejemplo de la aspirina, se podría administrar un inhibidor de la bomba de protones, que es una clase de medicamento que se usa para prevenir los síntomas gastrointestinales debidos al exceso de ácido, a ambos grupos aleatorizados, lo que enmascara a los participantes con respecto al grupo al que fueron asignados. Además de enmascarar a los sujetos, también queremos enmascarar a los observadores o a los que recogen los datos con respecto al grupo al que pertenece el paciente. El enmascaramiento de los participantes y del personal del estudio se denomina «doble ciego». Hace algunos años, se realizó un estudio en unidades de cuidados coronarios para evaluar el tratamiento del infarto de miocardio. Se planteó del siguiente modo:

Los pacientes que cumplían criterios estrictos de categorías de infarto de miocardio eran asignados aleatoriamente al grupo que era ingresado inmediatamente en la unidad de cuidados coronarios o al grupo que era enviado a sus casas para tratamiento domiciliario. Cuando se presentaron los datos preliminares, en las primeras fases del experimento era aparente que el grupo de pacientes clasificados como ingresados en la unidad de cuidados coronarios evolucionaban algo mejor que los pacientes enviados a sus domicilios. Un entusiasta de las unidades de cuidados coronarios insistió con firmeza en el hecho de que el experimento no era ético y que tenía que finalizar, ya que los datos demostraban que todos los pacientes deberían ser ingresados en unidades de cuidados coronarios. El estadístico reveló entonces que los encabezados de las columnas de datos habían sido intercambiados y que en realidad el grupo de tratamiento domiciliario había evolucionado ligeramente mejor. El entusiasta cambió entonces su punto de vista y no se le pudo convencer respecto a declarar como poco éticas a las unidades de cuidados coronarios[15].

El mensaje de este ejemplo es que cada uno de nosotros se enfrenta al estudio que se esté realizando con algunas ideas preconcebidas y sesgos conscientes o inconscientes. Los métodos expuestos en este capítulo y en el capítulo 11 están diseñados para proteger el estudio de los sesgos de los investigadores.

A continuación, nos ocuparemos de otros dos aspectos del diseño de los ensayos aleatorizados: el diseño cruzado y el diseño factorial.

Diseño cruzado

Otro aspecto importante de los ensayos clínicos es el *diseño cruzado*. El diseño cruzado puede ser de dos tipos: planeado o no planeado.

En la figura 10.5 se muestra un *diseño cruzado planeado*. En este ejemplo se compara un tratamiento nuevo con el tratamiento habitual. Los sujetos son aleatorizados al tratamiento nuevo o al tratamiento habitual (v. fig. 10.5A). Tras ser observados durante cierto periodo de tiempo mientras seguían un tratamiento y tras medir los cambios (v. fig. 10.5B), los pacientes cambian al otro tratamiento (v. fig. 10.5C). A continuación, ambos grupos son observados de nuevo durante cierto periodo de tiempo (v. fig. 10.5D). Los cambios en los pacientes del grupo 1 mientras siguen el tratamiento nuevo pueden compararse con los cambios en estos pacientes mientras siguen el tratamiento habitual (v. fig. 10.5E). Los cambios en los pacientes del grupo 2 mientras siguen el tratamiento nuevo también pueden compararse con los cambios en estos pacientes mientras siguen el tratamiento habitual (v. fig. 10.5F). De este modo, cada paciente sirve como su propio control, manteniendo constante la variación entre individuos de muchas características que en potencia podrían afectar a la comparación de la eficacia de los dos agentes.

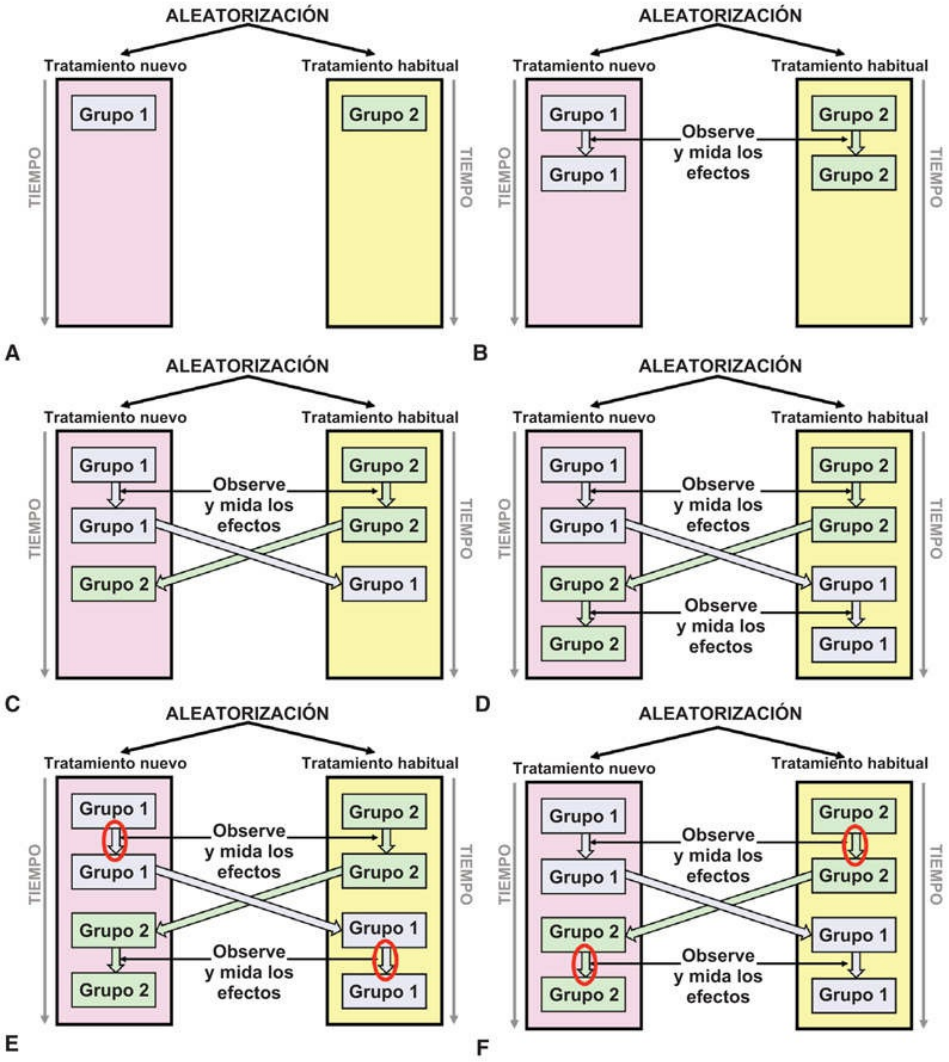


FIG. 10.5 (A-F) Diseño de un ensayo **cruzado planeado.** Véase la explicación en el texto.

Este tipo de diseño es muy atractivo y útil siempre que se tengan en cuenta ciertas precauciones. En primer lugar, el *efecto de arrastre:* por ejemplo, si un sujeto es cambiado del tratamiento A al tratamiento B y es observado con cada tratamiento, las observaciones con el tratamiento B serán válidas solo si no existe un arrastre residual desde el tratamiento A. Debe existir un «periodo de lavado» suficiente para asegurarse de que no permanece nada del tratamiento A o sus efectos antes de comenzar el tratamiento B. En segundo lugar, el orden en el que se administran los tratamientos puede suscitar respuestas psicológicas. Los pacientes pueden reaccionar de manera diferente al primer tratamiento administrado en el estudio como resultado del entusiasmo que a menudo se corresponde con un nuevo estudio; este entusiasmo puede disminuir con el paso del tiempo. Por tanto, queremos estar seguros de que las diferencias observadas se deben en realidad a los agentes que se están evaluando, y no a los efectos del orden en el que se administraron. Por último, el diseño cruzado planeado claramente no es posible si el tratamiento nuevo es quirúrgico o si el nuevo tratamiento cura la enfermedad.

Una consideración más importante es la de los *ensayos cruzados no planeados*. En la figura 10.6A se muestra el diseño de un ensayo aleatorizado sobre la cirugía de derivación coronaria, comparándola con el tratamiento médico de la cardiopatía coronaria. La aleatorización se llevó a cabo tras obtener el consentimiento informado. Aunque el diseño inicial es sencillo, en realidad pueden producirse cruces no planeados. Algunos sujetos aleatorizados al grupo de cirugía de derivación pueden decidir no someterse a la cirugía tras habérselo pensado (v. fig. 10.6B). Por tanto, se cruzan al grupo de tratamiento médico (v. fig. 10.6C). Además, el estado de algunos sujetos asignados al grupo de tratamiento médico puede comenzar a empeorar y puede ser necesaria una cirugía de derivación urgente (v. fig. 10.6B); estos pacientes se cruzan del grupo de tratamiento médico al grupo de tratamiento quirúrgico (v. fig. 10.6C). Los pacientes que se muestran en la parte izquierda de la figura 10.6D ahora son tratados quirúrgicamente y los de la derecha de la figura son tratados médicamente. En los tratados quirúrgicamente se incluyen los que fueron aleatorizados a cirugía (se muestran en rosa) y los que llegaron al tratamiento quirúrgico tras el cruce (se muestran en amarillo). En el grupo tratado médicamente se encuentran los que fueron aleatorizados al tratamiento médico (se muestran en amarillo) y los que llegaron al tratamiento médico tras el cruce (se muestran en rosa).

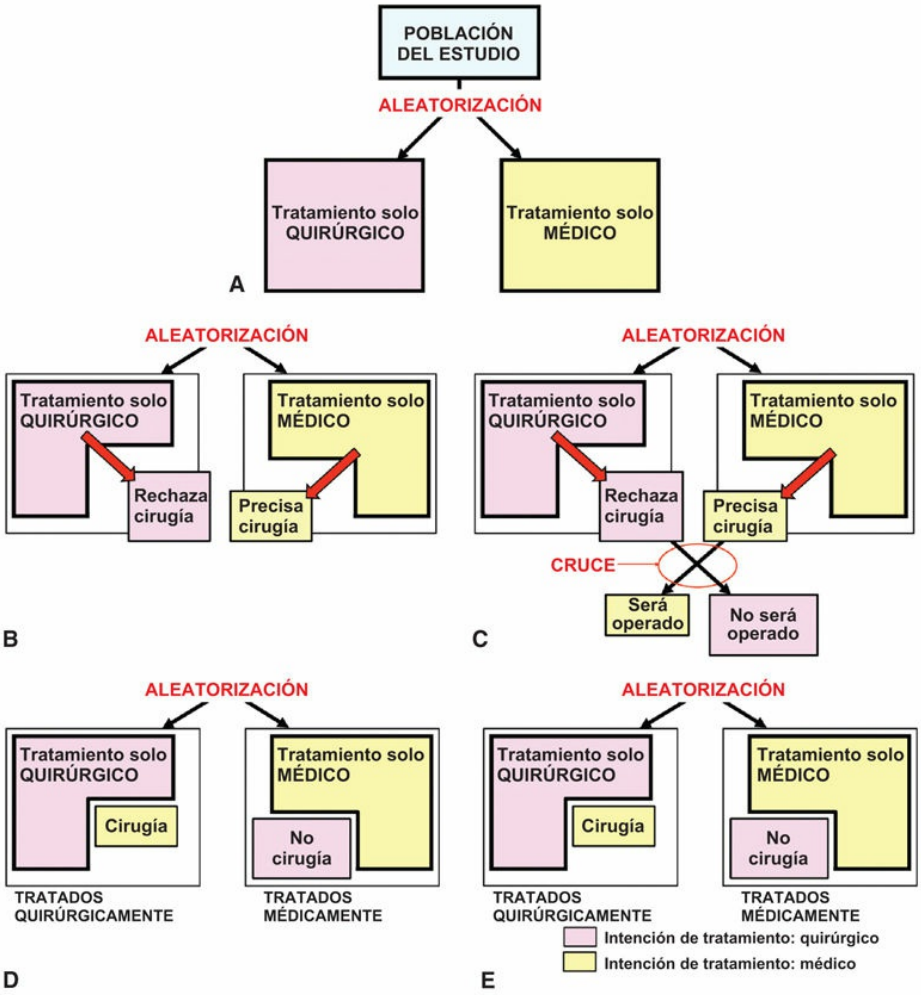


FIG. 10.6 (A-E) **Cruce no planeado** en un estudio de cirugía de derivación cardiaca y el uso del análisis de la intención terapéutica. (A) Diseño original del estudio. (B-D) Cruces no planeados. (E) Uso del análisis de la intención terapéutica.

Los cruces no planeados suponen un reto importante durante el análisis de los datos. Si realizamos el análisis teniendo en cuenta la asignación original (denominado *análisis de la intención de tratamiento*), incluiremos en el grupo quirúrgico a algunos pacientes que solo recibieron tratamiento médico e incluiremos en el grupo médico a algunos pacientes sometidos a la intervención quirúrgica. En otras palabras, compararíamos a los pacientes en función del tratamiento al que fueron aleatorizados inicialmente, sin tener en cuenta el tratamiento que recibieron en realidad. En la figura 10.6E se muestra el análisis de la intención de tratamiento en el que comparamos el grupo de color rosa (aleatorizado al tratamiento quirúrgico) con el grupo de color amarillo (aleatorizado al tratamiento médico). Sin embargo, si analizamos en función del tratamiento que los pacientes recibieron en realidad *(como análisis tratado)*, habremos puesto fin y, por tanto, habremos perdido los beneficios de la aleatorización.

No existe una solución perfecta para este dilema. La tendencia actual es realizar el análisis primario de la intención de tratamiento, en función de la asignación aleatorizada original. Esperaríamos que los resultados de otras

comparaciones fuesen compatibles con este primer abordaje. La conclusión es que, como no hay soluciones perfectas, el número de cruces no planeados debe ser el mínimo posible. Evidentemente, si analizamos teniendo en cuenta la aleatorización original y se han producido muchos cruces, la interpretación de los resultados del estudio será cuestionable. Si el número de cruces es grande, el problema de interpretar los resultados del estudio puede ser insalvable.

Diseño factorial

Una alternativa atractiva a los diseños de estudios expuestos en estos capítulos es el *diseño factorial*. Suponiendo que vamos a estudiar dos fármacos, que los resultados previstos de los dos fármacos son diferentes y que sus modos de acción son independientes, por economía podemos utilizar la misma población de estudio para probar ambos fármacos. Este tipo de diseño factorial se expone en la figura 10.7.

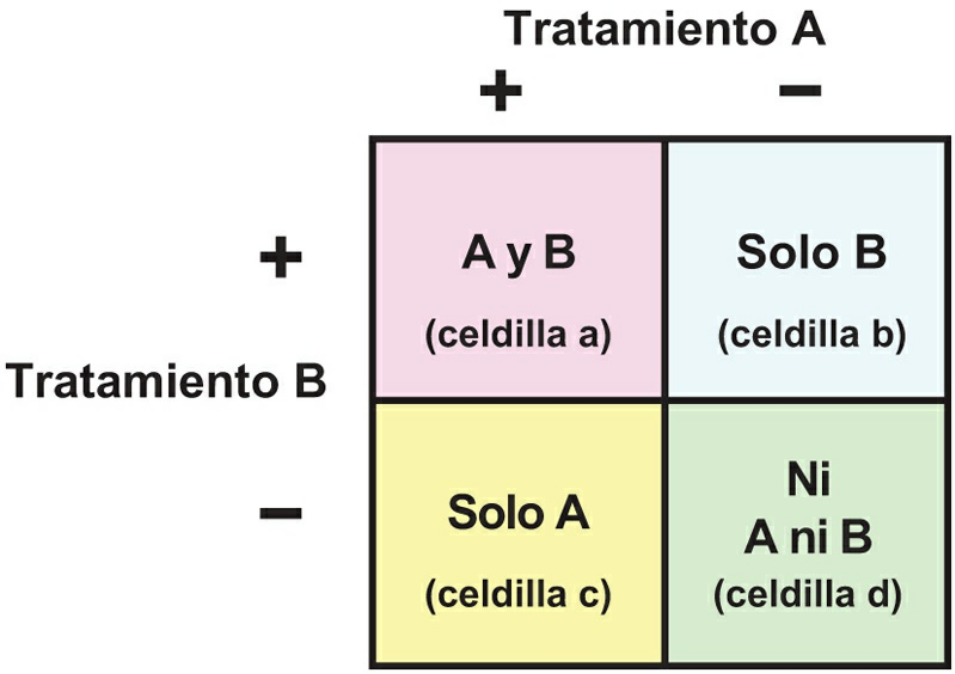


FIG. 10.7 Diseño factorial para el estudio de los efectos de dos tratamientos.

Si los efectos de los dos tratamientos son de hecho completamente independientes, podemos evaluar los efectos del tratamiento A comparando los resultados de las celdillas $a + c$ con los resultados de las celdillas $b + d$ (fig. 10.8A). De modo parecido, los resultados del tratamiento B pueden evaluarse comparando los efectos de las celdillas $a + b$ con los de las celdillas $c + d$ (v. fig. 10.8B). En caso de que se decida finalizar el estudio del tratamiento A, este diseño permite continuar el estudio para determinar los efectos del tratamiento B.

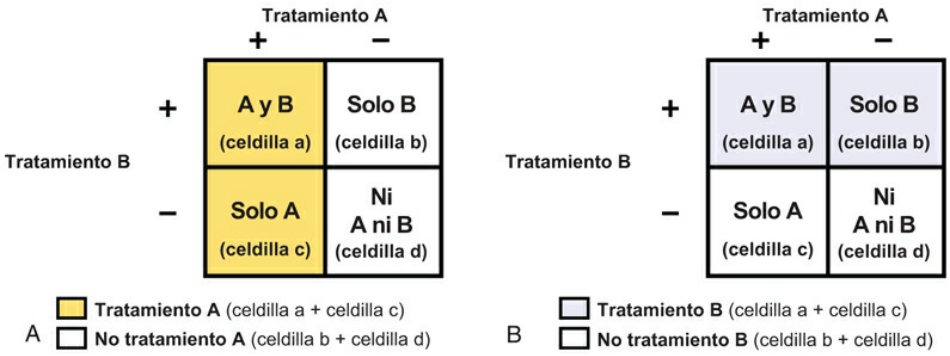


FIG. 10.8 (A-B) Diseño factorial. (A) Los efectos del tratamiento A *(celdillas naranjas)* frente a no emplear el tratamiento A. (B) Los efectos del tratamiento B *(celdillas púrpura)* frente a no emplear el tratamiento B.

El Physician's Health Study es un ejemplo de diseño factorial[16]. Más de 22.000 médicos fueron aleatorizados utilizando un diseño factorial 2 × 2 que estudiaba el uso de aspirina para la prevención primaria de enfermedades cardiovasculares y el uso de beta-caroteno para la prevención primaria del cáncer. Cada médico recibió una de las cuatro posibles intervenciones: aspirina y beta-caroteno, ni aspirina ni beta-caroteno, aspirina y placebo de beta-caroteno o beta-caroteno y placebo de aspirina. Los cuatro grupos resultantes se muestran en las figuras 10.9 y 10.10. La parte del estudio de la aspirina (fig. 10.11A) finalizó pronto, siguiendo el consejo del comité externo de control de datos, debido a que en el grupo que tomaba aspirina se observó una disminución estadísticamente significativa del 44% del riesgo de sufrir un primer infarto de miocardio. El grupo aleatorizado para el beta-caroteno (v. fig. 10.11B) continuó el estudio hasta la fecha programada de terminación. Tras 12 años de suplementos con beta-caroteno no se observaron beneficios ni efectos adversos en relación con la incidencia de cáncer, cardiopatía o muerte por todas las causas. Estudios posteriores han mostrado un mayor riesgo de cáncer con el uso de beta-caroteno en fumadores[17].



FIG. 10.9 Diseño factorial usado en un estudio sobre aspirina y beta-caroteno.

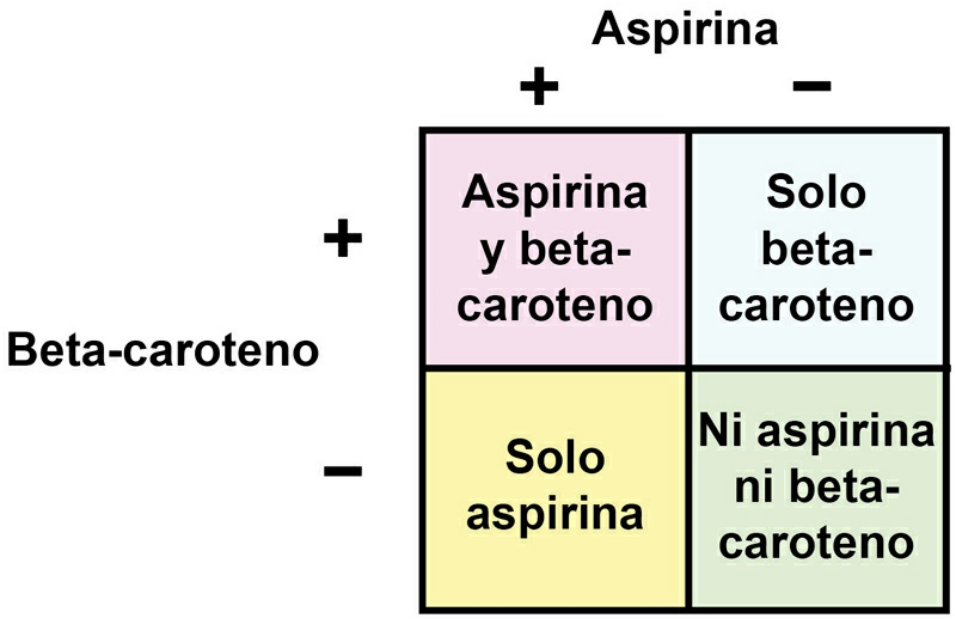


FIG. 10.10 Diseño factorial del estudio sobre aspirina y beta-caroteno en formato de tabla 2 × 2.

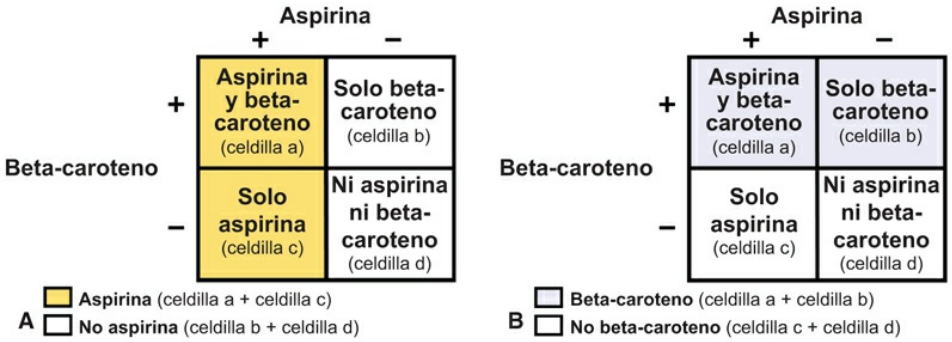


FIG. 10.11 (A-B) Diseño factorial. (A) Los efectos de la aspirina *(celdillas naranjas)* frente al no tratamiento con aspirina. (B) Los efectos del beta-caroteno *(celdillas púrpuras)* frente al no tratamiento con beta-caroteno.

Falta de cumplimiento

Los pacientes pueden aceptar el proceso de aleatorización, pero tras finalizarlo puede que no cumplan el tratamiento asignado. La falta de cumplimiento puede ser manifiesta o encubierta: por un lado, los pacientes pueden expresar abiertamente su negativa a cumplir el tratamiento o pueden interrumpir su participación en el estudio. Estos casos de falta de cumplimiento también se conocen como *abandonos* del estudio. Por otro lado, los pacientes pueden simplemente dejar de tomar el agente asignado sin informar al investigador o al personal del estudio. Siempre que sea posible, el estudio debe contar con comprobaciones para descubrir las posibles faltas de cumplimiento. Entre ellas se encuentran, por ejemplo, análisis de orina para el agente que se está estudiando o para uno de sus metabolitos.

Otro problema de los ensayos aleatorizados es el de los *cambios espontáneos* de tratamiento. Los pacientes de un grupo pueden tomar inadvertidamente el agente asignado al otro grupo. Por ejemplo, en un estudio sobre el efecto de la aspirina para la prevención del infarto de miocardio, los pacientes fueron aleatorizados a tomar aspirina o no tomarla. Sin embargo, surgió un problema porque, debido al gran número de fármacos de venta sin receta médica que contienen aspirina, muchos de los pacientes controles podrían estar tomando también aspirina sin saberlo. Se adoptaron dos medidas para abordar este problema: 1) se proporcionó a los controles listas de las preparaciones de venta sin receta médica que contenían aspirina, que debían evitar, y 2) se llevaron a cabo análisis de orina para identificar salicilatos tanto en el grupo que tomaba aspirina como en los controles.

El efecto neto de la falta de cumplimiento en los resultados del estudio será la reducción de las diferencias observadas (es decir, la anulación de las diferencias), porque el grupo de tratamiento incluirá pacientes que no tomaron el tratamiento y el grupo sin tratamiento puede incluir algunos que tomaron el tratamiento. Así, los grupos serán menos diferentes en cuanto al tratamiento de lo que habrían sido si no se hubiesen producido faltas de cumplimiento, de manera que, aunque haya una diferencia en los efectos del tratamiento, parecerá mucho menor.

Un abordaje utilizado en el Veterans Administration Study sobre el tratamiento de la hipertensión fue realizar un estudio piloto que identificaba a los pacientes que cumplían el tratamiento y a los que no lo cumplían. Cuando el estudio completo real fue realizado posteriormente, la población del estudio se limitó a aquellos que habían cumplido el tratamiento durante el estudio piloto (a veces referido como «periodo de prueba»). El problema con este abordaje es que, cuando queramos realizar generalizaciones a partir de los resultados de dicho estudio, solo podremos hacerlo en otras poblaciones de pacientes cumplidores, que pueden ser diferentes de la población en cualquier comunidad abierta, que estaría formada por cumplidores y no cumplidores.

La tabla 10.6 muestra los datos del Coronary Drug Project comunicados por Canner y cols.[18]. Este estudio comparaba el uso de clofibrato y de placebo en la reducción del colesterol. La tabla presenta la mortalidad en los dos grupos.

Tabla 10.6

Proyecto de fármacos coronarios: mortalidad a 5 años en pacientes a los que se administra clofibrato o placebo

	Número de pacientes	Mortalidad (%)
Clofibrato	1.065	18,2
Placebo	2.695	19,4

Modificada de Canner PL, Forman SA, Prud'homme GJ, para el Coronary Drug Project Research Group: Influence of adherence to treatment and response to cholesterol on mortality in the coronary drug project. *N Engl J Med*. 1980;303:1038–1041.

No se observaron grandes diferencias en la mortalidad a 5 años entre los dos grupos. Los investigadores especularon que quizá esto era el resultado de que los pacientes no hubiesen tomado la medicación. En la tabla 10.7 se muestran los resultados tras separar a los sujetos del grupo de clofibrato en buenos cumplidores y malos cumplidores. Observamos que la mortalidad a 5 años fue del 24,6% en el grupo de malos cumplidores y del 15% en el grupo de buenos cumplidores. Podríamos estar tentados a concluir que el cumplimiento fue de hecho el factor que produjo los resultados observados en la tabla 10.6: la ausencia de diferencias significativas entre los grupos que tomaban clofibrato y placebo.

Tabla 10.7

Proyecto de fármacos coronarios: mortalidad a 5 años en pacientes a los que se administra clofibrato o placebo, según el grado de cumplimiento

	Número de pacientes	Mortalidad (%)
Clofibrato		
Mal cumplidor (<80%)	357	24,6
Buen cumplidor (≥80%)	708	15,0
Placebo	2.695	19,4

Modificada de Canner PL, Forman SA, Prud'homme GJ, para el Coronary Drug Project Research Group: Influence of adherence to treatment and response to cholesterol on mortality in the coronary drug project. *N Engl J Med*. 1980;303:1038–1041.

En la tabla 10.8 se separan ambos grupos, el que tomaba clofibrato y el que tomaba placebo, en cumplidores y no cumplidores. Incluso en el grupo que tomaba placebo, la mortalidad a 5 años en el grupo de malos cumplidores era

más elevada que en el grupo de buenos cumplidores: 28% frente a 15%. Una forma de maximizar el cumplimiento es administrar una sola píldora que incluya una combinación de dos medicamentos necesarios para lograr el objetivo terapéutico. Esto se conoce comúnmente como combinaciones de dosis fijas (CDF). Una revisión sistemática y un metaanálisis encontraron que el uso de CDF de medicamentos antihipertensivos se asocia con una mejora significativa en el cumplimiento o la persistencia de la medicación, a pesar de las tendencias no significativas pero beneficiosas en la presión arterial y los efectos adversos[19].

Tabla 10.8

Proyecto de fármacos coronarios: mortalidad a 5 años en pacientes a los que se administra clofibrato o placebo, según el grado de cumplimiento

	CLOFIBRATO		PLACEBO	
Cumplimiento	Número de pacientes	Mortalidad (%)	Número de pacientes	Mortalidad (%)
Malo (<80%)	357	24,6	882	28,2
Bueno (≥80%)	708	15,0	1.813	15,1
Total	1.065	18,2	2.695	19,4

Adaptada de Canner PL, Forman SA, Prud'homme GJ, para el Coronary Drug Project Research Group: Influence of adherence to treatment and response to cholesterol on mortality in the coronary drug project. *N Engl J Med*. 1980;303:1038–1041.

¿Qué podemos sacar en claro de estas tablas? Las personas que no cumplen el tratamiento o las que no participan en estudios se diferencian de las que son cumplidoras o de las que sí participan. Por tanto, para realizar un estudio que valore un tratamiento u otra intervención, no podemos ofrecer el agente a una población y comparar los efectos en los que toman el agente y en los que se niegan o no lo toman, porque los dos grupos son básicamente diferentes en lo que respecta a diversas variables demográficas, sociales, psicológicas y culturales que pueden desempeñar un papel importante para determinar el resultado. Todos estos son *sesgos de selección*, que fueron abordados previamente cuando expusimos los diseños de los estudios observacionales. La aleatorización, u otro método que reduzca los sesgos de selección, es esencial en un ensayo clínico válido.

Conclusión

El ensayo aleatorizado se considera por lo general el método de referencia entre los diseños de los estudios. Cuando se crean jerarquías de diseños de estudio para valorar la solidez de las pruebas disponibles que respaldan las políticas de salud pública y clínica, los ensayos aleatorizados se encuentran prácticamente siempre al principio de la lista cuando los diseños de estudios se colocan en orden descendente en función de su calidad. Sin embargo, un enfoque de estudio observacional recientemente desarrollado —la aleatorización mendeliana— cuyo abordaje no está dentro del alcance de este libro de texto, imita la asignación aleatoria si se cumplen sus estrictos supuestos[20].

Este capítulo ha analizado muchos de los componentes de los ensayos aleatorizados que están diseñados para proteger el estudio de cualquier idea preconcebida y sesgos por parte del investigador y de otras personas que participan en la realización del estudio, así como de otros sesgos que podrían introducirse inadvertidamente. En el capítulo 11 abordaremos algunos otros aspectos relacionados con el diseño de los ensayos aleatorizados y consideraremos varios ejemplos interesantes y las aplicaciones del diseño de los ensayos aleatorizados. Más adelante en este libro analizaremos el uso de los ensayos aleatorizados y otros diseños de estudios para evaluar los servicios sanitarios y para estudiar la eficacia del cribado.

Bibliografía

1. Cited in Silverman WA. Where's the Evidence? Debates in Modern Medicine. New York: Oxford University Press; 1998.
2. Galton F. Inquiries Into Human Faculty and Its Development. London: Macmillan; 1883.
3. Joyce CRB, Welldon RMC. The efficacy of prayer: a double blind clinical trial. *J Chronic Dis*. 1965;18:367.
4. Byrd RC. Positive therapeutic effects of intercessory prayer in a coronary care unit population. *South Med J*. 1988;81:826.
5. Bull JP. The historical development of clinical therapeutic trials. *J Chronic Dis*. 1959;10:218.
6. Lind J. A Treatise of the Scurvy. Edinburgh: Sands, Murray & Cochran; 1753.
7. Peacock E. Cited in Tufte ER: Data Analysis for Politics and Policy. Englewood Cliffs, NJ: Prentice-Hall; 1974.
8. Ederer F. Why do we need controls? Why do we need to randomize? *Am J Ophthalmol*. 1975;79:758.
9. Bearman JE, Loewenson RB, Gullen WH. *Muensch's Postulates, Laws and Corollaries.* Biometrics Note No. 4. Bethesda, MD, Office of Biometry and Epidemiology, National Eye Institute, April 1974.
10. Wilson EB. Cited in Ederer F: Why do we need controls? Why do we need to randomize? *Am J Ophthalmol.* 1975;79:761.
11. Wright IS, Marple CD, Beck DF. Cited in Ederer F: Why do we need controls? Why do we need to randomize? *Am J Ophthalmol.* 1975;79:761.
12. Levine MI, Sackett MF. Results of BCG immunization in New York City. *Am Rev Tuber*. 1946;53:517–532.
13. Ederer F. Practical problems in collaborative clinical trials. *Am J Epidemiol*. 1975;102:111–118.
14. Schröder FH, Hugosson J, Roobol MJ, et al. Screening and prostate-cancer mortality in a randomized European study. *N Engl J Med*. 2009;360:1320–1328.
15. Cochrane AL. Cited in Ballintine EJ: Objective measurements and the double masked procedure. *Am J Ophthalmol.* 1975;79:764.
16. Hennekens CH, Buring JE, Manson JE. Lack of effect of long-term supplementation with beta carotene on the incidence of

malignant neoplasms and cardiovascular disease. *N Engl J Med*. 1996;334:1145–1149.
17. Goralczyk R. Beta-carotene and lung cancer in smokers: review of hypotheses and status of research. *Nutr Cancer*. 2009;61(6):767–774.
18. Canner PL, Forman SA, Prud'homme GJ. Influence of adherence to treatment and response of cholesterol on mortality in the coronary drug project. *N Engl J Med*. 1980;303:1038–1041.
19. Gupta AK, Arshad S, Poulter NR. Compliance, safety, and effectiveness of fixed-dose combinations of antihypertensive agents: a meta-analysis. *Hypertension*. 2010;55(2):399–407.
20. Smith GD, Ebrhaim S. 'Mendelian randomization': Can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol*. 2003;32(1):1–22.

Las preguntas de repaso de los capítulos 10 y 11 se encuentran al final del capítulo 11.

CAPÍTULO 11

Ensayos aleatorizados: algunos aspectos adicionales

Palabras clave

validez
error de tipo I y II
confiabilidad
generalización
eficacia
efectividad
número necesario para tratar
registro del ensayo

Objetivos de aprendizaje

- Definir conceptos clave del diseño de estudios epidemiológicos en el contexto de los ensayos aleatorizados: tamaño de la muestra, error de tipo I, error de tipo II, potencia, generalización (validez externa) y validez interna.
- Calcular e interpretar la eficacia en un ensayo aleatorizado.
- Describir el diseño y los resultados de cinco ensayos aleatorizados históricamente importantes.
- Definir las cuatro fases principales de los ensayos aleatorizados utilizadas por la agencia estadounidense del medicamento (FDA) para evaluar nuevos fármacos en Estados Unidos.
- Introducir algunas consideraciones éticas relacionadas

con los ensayos aleatorizados.

- Analizar el motivo del requerimiento del registro de los ensayos aleatorizados nuevos antes de su lanzamiento.

Tamaño de la muestra

En una reunión científica celebrada hace algunos años, un investigador presentó los resultados de un estudio que había realizado para valorar un nuevo fármaco para ovejas. «Tras administrar el fármaco», comentó, «un tercio de las ovejas mejoró considerablemente, otro tercio no experimentó ningún cambio y un tercio se escapó.»

Esta historia presenta una de las preguntas planteadas con mayor frecuencia por los médicos que realizan ensayos de nuevos fármacos o, de hecho, por cualquiera que realice estudios evaluadores: ¿cuántos sujetos se deben estudiar? El momento de responder esta pregunta es *antes* de realizar el estudio. Con demasiada frecuencia se realizan estudios, se invierten grandes sumas de dinero y otros recursos, y solo después de que el estudio se ha completado es cuando los investigadores descubren que desde el inicio contaban con muy pocos sujetos para obtener resultados significativos.

La cuestión de cuántos sujetos se necesitan para un estudio no se basa en la mística. Esta sección presenta la lógica para abordar la cuestión del tamaño de la muestra. Comencemos este análisis sobre el tamaño de la muestra con la figura 11.1.

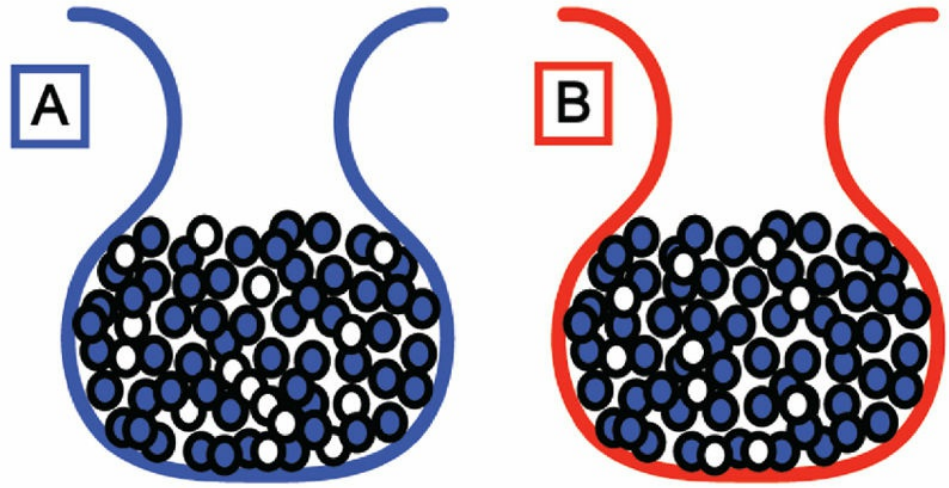


FIG. 11.1 Dos vasijas opacas (A) y (B); cada una de ellas contiene 100 cuentas, unas azules y otras blancas.

Tenemos dos vasijas con cuentas; cada una contiene 100 cuentas, unas azules y otras blancas. Las vasijas son opacas («enmascaradas»), de modo que (a pesar de su aspecto en la figura) no podemos ver los colores de las cuentas que hay en su interior. Queremos saber si la distribución de las cuentas por color es diferente en la vasija A y en la vasija B. Dicho de otro modo, ¿hay una proporción mayor (o menor) de cuentas azules en la vasija A que en la vasija B?

Para contestar a esta pregunta, tomemos una muestra de 10 cuentas de la

vasija A en una mano y una muestra de 10 cuentas de la vasija B en la otra. En función de la distribución de color de las 10 cuentas en cada mano, intentaremos alcanzar una conclusión acerca de la distribución de color de las 100 cuentas en cada una de las vasijas.

Supongamos que (como se muestra en la fig. 11.2) en una mano tenemos 9 cuentas azules y 1 cuenta blanca de la vasija A y en la otra mano tenemos 2 cuentas azules y 8 cuentas blancas de la vasija B. ¿Podemos concluir que el 90% de las cuentas de la vasija A son azules y que el 10% son blancas? Claramente, no. Es posible, por ejemplo, que de las 100 cuentas de la vasija A, 90 sean blancas y 10 azules, pero, por *azar*, nuestra muestra de 10 cuentas consta de 9 azules y 1 blanca. Esto es posible, pero muy poco probable. De modo similar, con respecto a la vasija B, no podemos concluir que el 20% de las cuentas sean azules y el 80% sean blancas. Es concebible que 90 de las 100 cuentas sean azules y 10 sean blancas, pero, por *azar*, la muestra de 10 cuentas contiene 2 azules y 8 blancas. Esto es posible, pero, de nuevo, muy improbable.

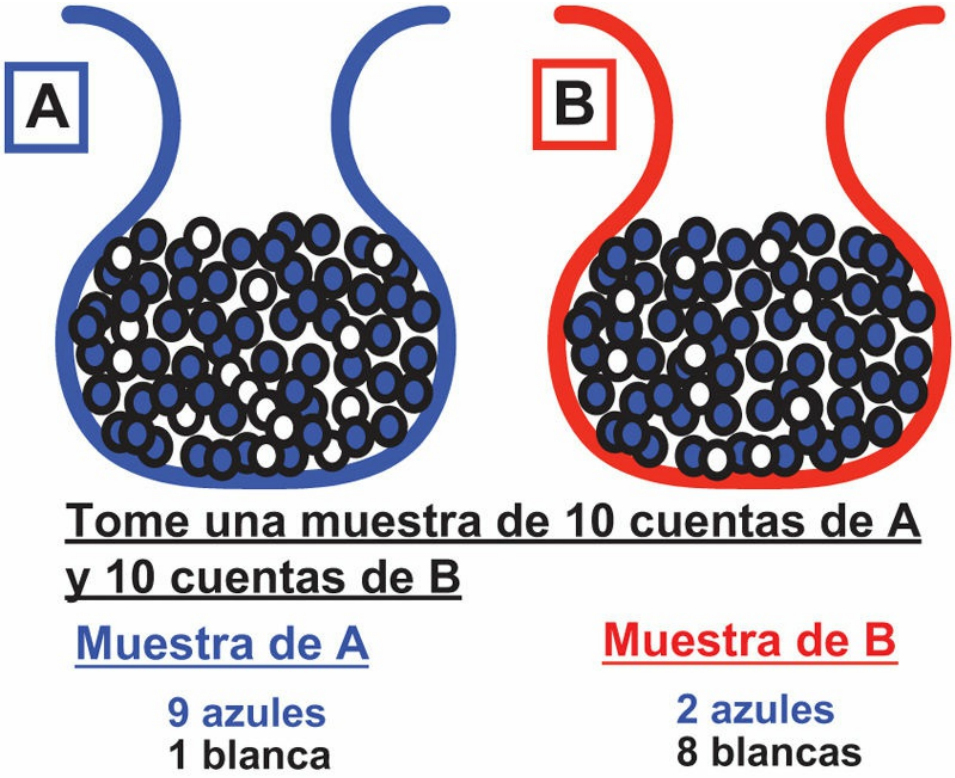


FIG. 11.2 Muestras de 10 cuentas de la vasija A y 10 cuentas de la vasija B.

Basándonos en las distribuciones de las muestras de 10 cuentas en cada mano, ¿podemos concluir que las distribuciones de las 100 cuentas en las dos vasijas son diferentes? Teniendo en cuenta las muestras en cada mano, ¿podría ocurrir, por ejemplo, que la distribución de cuentas en cada vasija fuese de 50 azules y 50 blancas? De nuevo, es posible, pero no es probable. No podemos excluir esta posibilidad basándonos en nuestras muestras. Miramos

las muestras y tratamos de llegar a una conclusión respecto a toda la población, las vasijas de las que hemos extraído las muestras.

Fijémonos ahora en el ejemplo que se muestra en la figura 11.3. De nuevo, obtenemos dos muestras. En esta ocasión, la muestra de 10 cuentas de la vasija A se compone de 7 cuentas azules y 3 blancas, y la muestra de 10 cuentas de la vasija B también se compone de 7 cuentas azules y 3 blancas. ¿Es posible que la distribución de color de las cuentas de las dos vasijas sea la misma? Claramente, podría ser. ¿Podríamos haber extraído estas dos muestras de 7 cuentas azules y 3 blancas de ambas vasijas si la distribución es en realidad de 90 cuentas blancas y 10 azules en la vasija A y 90 cuentas azules y 10 blancas en la vasija B? Sí, posiblemente, pero poco probable.

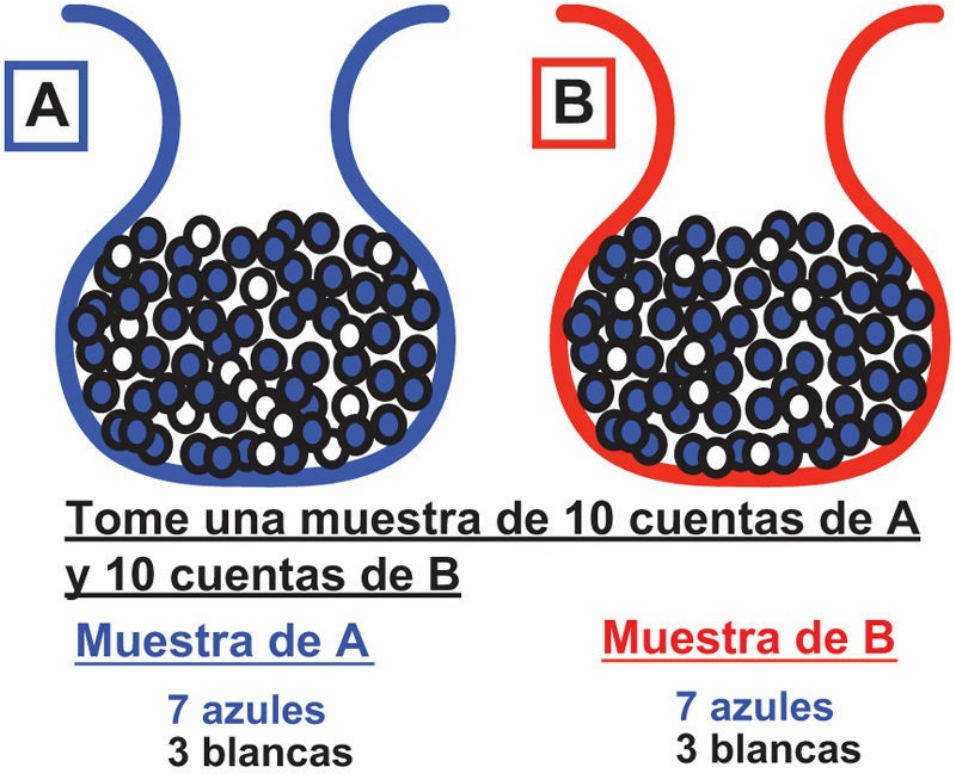


FIG. 11.3 Muestras de 10 cuentas de la vasija A y 10 cuentas de la vasija B.

Cuando realizamos un estudio, solo nos fijamos en la muestra de sujetos de nuestro estudio, como una muestra de pacientes con cierta enfermedad que están siendo tratados con el tratamiento A o con el tratamiento B. A partir de los resultados del estudio queremos extraer una conclusión de aplicación más allá de la población del estudio: ¿el tratamiento A es más eficaz que el tratamiento B en el universo total de todos los pacientes que tienen esta enfermedad que podrían tratarse con el tratamiento A o con el tratamiento B? El mismo problema que surgió con las muestras de 10 cuentas surge cuando queremos obtener una conclusión para todos los pacientes a partir de la muestra de los pacientes de nuestro estudio. Raramente, si es que ocurre alguna vez, un estudio se realiza en todos los pacientes que tienen una

enfermedad o en todos los pacientes que podrían ser tratados con el fármaco en cuestión.

Con estos antecedentes, consideremos ahora un ensayo en el que se comparan grupos que reciben un tratamiento dentro de dos posibles, A y B. (Recuérdese el muestreo de cuentas que acabamos de describir.) Antes de comenzar nuestro estudio, podemos enumerar los cuatro posibles resultados del estudio (cuadro 11.1):

1. Es posible que en realidad no existan diferencias en la eficacia entre el tratamiento A y el tratamiento B. En otras palabras, el tratamiento A no es ni mejor ni peor que el tratamiento B. Cuando realizamos nuestro estudio, concluimos correctamente en función de nuestras muestras que los dos grupos no se diferencian.
2. Es posible que en realidad no existan diferencias en la eficacia entre el tratamiento A y el tratamiento B, pero en nuestro estudio encontramos una diferencia entre los grupos y, por tanto, concluimos, basándonos en nuestras muestras, que existe una diferencia entre los tratamientos. Esta conclusión, basada en nuestras muestras, es errónea.
3. Es posible que en realidad existan diferencias en la eficacia entre el tratamiento A y el tratamiento B, pero, cuando examinamos los grupos en nuestro estudio, no encontramos diferencias entre ellos. Por tanto, basándonos en nuestras muestras, concluimos que no existen diferencias entre el tratamiento A y el tratamiento B. Esta conclusión es errónea.
4. Es posible que en realidad existan diferencias en la eficacia entre el tratamiento A y el tratamiento B, y, cuando examinamos los grupos de nuestro estudio, observamos que existen diferencias. Basándonos en estas muestras, concluimos correctamente que el tratamiento A difiere del tratamiento B.

Cuadro 11.1 Cuatro posibles conclusiones cuando se prueba si los tratamientos se diferencian o no

- **Cuando en realidad los tratamientos no difieren:**
    1. Podemos concluir correctamente que **no** se diferencian
    o
    2. Por error, podemos concluir que se diferencian
- **Cuando en realidad los tratamientos sí difieren:**
    1. Por error, podemos concluir que no se diferencian
    o
    2. Podemos concluir correctamente que **sí** se diferencian

Estas cuatro posibilidades representan todos los posibles resultados tras

completar nuestro estudio. Fijémonos en estas cuatro posibilidades, que se presentan en una tabla de 2 × 2 (fig. 11.4): dos columnas representan la realidad (o el tratamiento A se diferencia del tratamiento B o el tratamiento A no se diferencia del tratamiento B). Las dos filas representan nuestra decisión: concluimos que se diferencian o que no se diferencian. En esta figura, las cuatro posibilidades que acabamos de enumerar se representan en cuatro celdillas en la tabla de 2 × 2. Si no existen diferencias y, basándonos en las muestras incluidas en nuestro estudio, concluimos que no existen diferencias, se trata de una decisión correcta *(celdilla a)*. Si existen diferencias y, basándonos en nuestro estudio, concluimos que existen diferencias *(celdilla d)*, también se trata de una decisión correcta. En el mejor de los casos, todas las posibilidades caerían en una de estas dos celdillas. Desafortunadamente, raramente se produce este hecho, si es que se produce alguna vez. Existen ocasiones en las que no hay diferencias entre los tratamientos, pero, basándonos en las muestras de los sujetos incluidos en nuestro estudio, concluimos erróneamente que sí son diferentes *(celdilla c)*. Esta posibilidad se denomina *error de tipo I*. También es posible que realmente existan diferencias entre los tratamientos, pero, basándonos en las muestras de nuestro estudio, concluimos erróneamente que no existe tal diferencia *(celdilla b)*; es el denominado *error de tipo II*. (En este caso, los tratamientos son diferentes, pero no hemos sido capaces de detectar la diferencia en las muestras de nuestro estudio.)

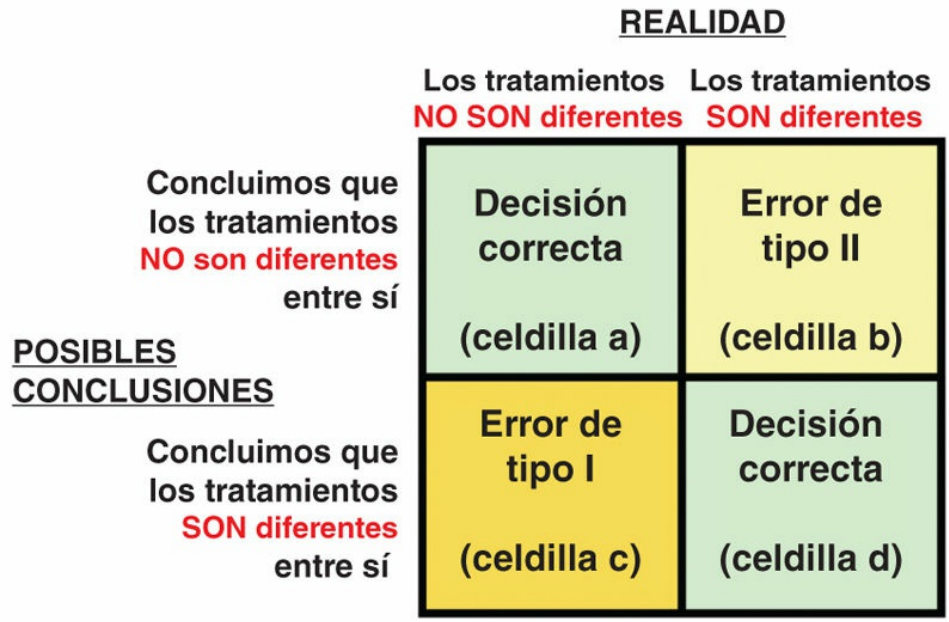


FIG. 11.4 Posibles resultados de un ensayo aleatorizado: errores de tipo I y de tipo II.

La *probabilidad* de cometer un error de tipo I se denomina α y la *probabilidad* de cometer un error de tipo II se denomina β (como se muestra en la fig. 11.5).

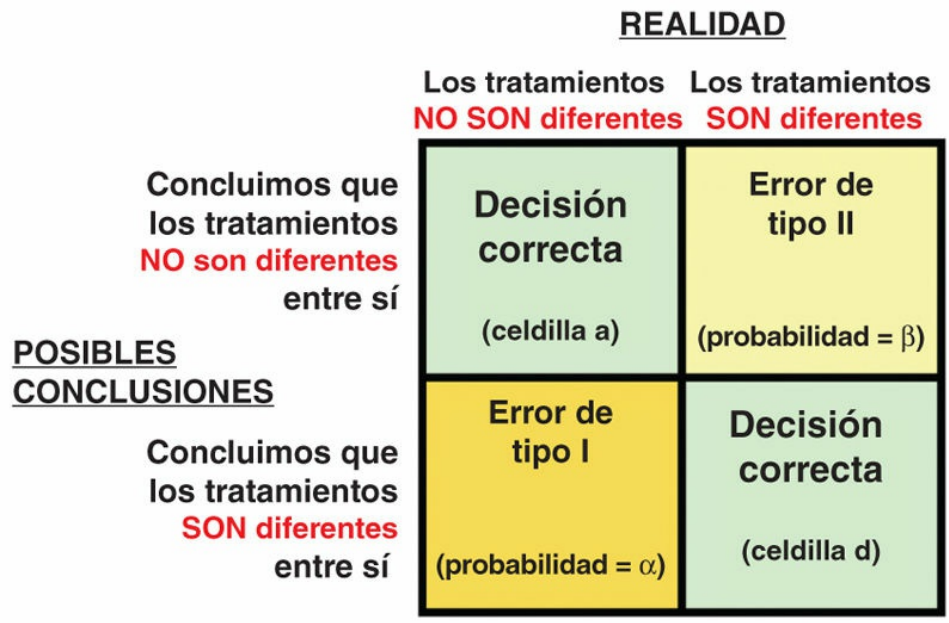


FIG. 11.5 Posibles resultados de un ensayo aleatorizado: α y β.

α es el denominado valor *P*, que vemos en muchas publicaciones y ha sido consolidado por muchos años de uso. Cuando leemos «*P* <0,05», se hace referencia a α. ¿Qué quiere decir que *P* <0,05? Nos indica que, basándonos en la muestra de sujetos incluidos en nuestro estudio, hemos concluido que el tratamiento A se diferencia del tratamiento B, porque hemos observado diferencias. La probabilidad de que dicha diferencia pudiera deberse al azar únicamente, y que dicha diferencia entre nuestros grupos no refleje una diferencia real entre el tratamiento A y el B, es de tan solo 0,05 (o 1 de 20).

Prestemos atención ahora a la mitad derecha de la tabla de 2 × 2, que muestra las dos posibilidades cuando existe una diferencia real entre el tratamiento A y el B, como se muestra en la figura 11.6. Si, como vemos aquí, la realidad es que existen diferencias entre los tratamientos, solo existen dos posibilidades: 1) Podríamos concluir, erróneamente, que los tratamientos no se diferencian (error de tipo II). La probabilidad de cometer un error de tipo II viene designada por β. O bien, 2) podríamos concluir, correctamente, que los tratamientos se diferencian. Como el total de todas las probabilidades debe ser igual a 1 y la probabilidad de un error de tipo II es = β, la probabilidad de decidir correctamente basándonos en nuestro estudio que los tratamientos son diferentes, cuando existen diferencias, será igual a 1 − β. Esta probabilidad, 1 − β, se denomina *potencia* del estudio. Nos dice cómo de bueno es nuestro estudio para identificar correctamente una diferencia entre los tratamientos cuando realmente son diferentes. ¿Cuál es la probabilidad de que nuestro estudio no pase por alto una diferencia si en realidad existe?

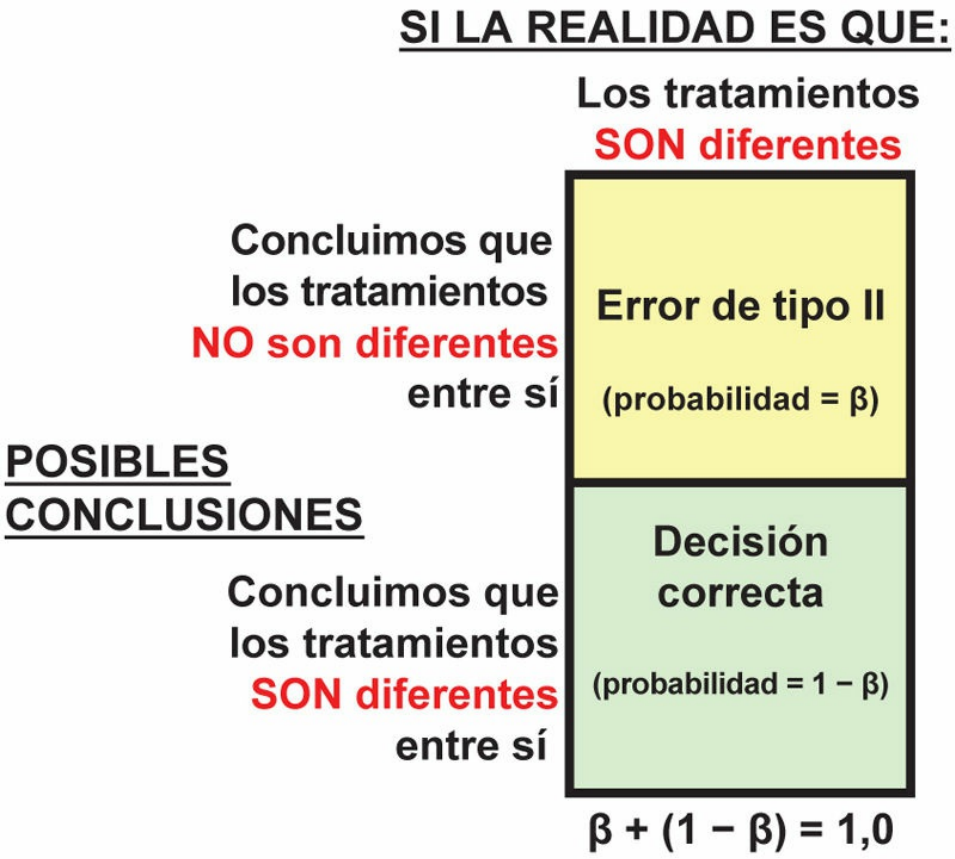


FIG. 11.6 Posibles resultados de un ensayo aleatorizado cuando los tratamientos difieren.

La tabla 2 × 2 completa de la figura 11.7 incluye todos los términos que hemos expuesto. En la tabla 11.1 se proporcionan múltiples definiciones para estos términos empleados habitualmente en la literatura epidemiológica.

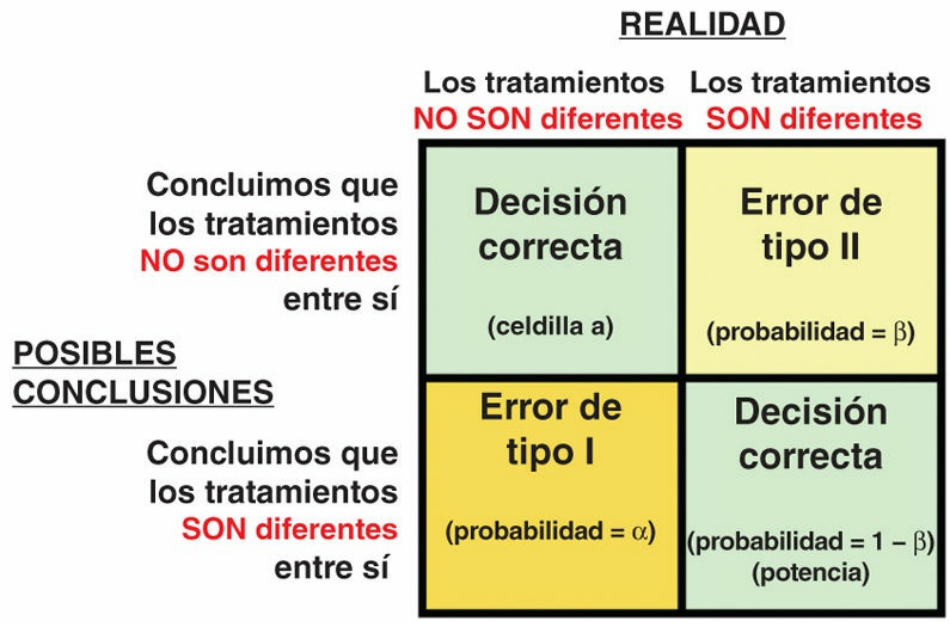


FIG. 11.7 Posibles resultados de un ensayo aleatorizado: resumen.

Tabla 11.1

Resumen de términos

Término	Definiciones
α	Probabilidad de cometer un error de tipo I
	Probabilidad de concluir que los tratamientos se diferencian cuando en realidad no difieren
β	Probabilidad de cometer un error de tipo II
	Probabilidad de concluir que los tratamientos no se diferencian cuando en realidad sí difieren
Potencia	1 − probabilidad de cometer un error de tipo II
	$1 - \beta$
	Probabilidad de concluir correctamente que los tratamientos son diferentes
	Probabilidad de detectar una diferencia entre los tratamientos si los tratamientos en realidad son diferentes

¿Cómo nos ayudan estos conceptos a estimar el tamaño de la muestra que necesitamos? Si nos planteamos la cuestión «¿cuántas personas tenemos que estudiar en un ensayo clínico?», debemos ser capaces de especificar los parámetros expuestos en el cuadro 11.2).

Cuadro 11.2 ¿Qué se debe especificar para estimar el tamaño de la muestra necesario en un ensayo aleatorizado?

1. La diferencia en las tasas de respuesta que se quiere detectar.
2. Una estimación de la tasa de respuesta en uno de los grupos.
3. El nivel de significación estadística (α).
4. El valor de la potencia deseada ($1 - \beta$).
5. Si la prueba es unilateral o bilateral.

En primer lugar, debemos especificar la diferencia esperada en la tasa de respuesta. Supongamos que el tratamiento existente cura al 40% de los pacientes y vamos a probar un tratamiento nuevo. Debemos ser capaces de decir si esperamos que el tratamiento nuevo cure al 50%, al 60% o a otro porcentaje de los pacientes tratados. Es decir, ¿el nuevo tratamiento será un 10% mejor que el tratamiento habitual y curará al 50% de los pacientes, o un 20% mejor que el tratamiento habitual y curará a un 60%, o un porcentaje diferente? ¿Qué tamaño de diferencia entre el tratamiento habitual y el tratamiento nuevo queremos ser capaces de detectar con nuestro estudio? ¿El tamaño de la diferencia tiene sentido clínico?

¿Cómo llegamos normalmente a la estimación de cuánto mejor será el nuevo tratamiento? ¿Qué pasa si no tenemos información sobre la que basar la estimación de la mejoría de la eficacia que podría anticiparse? Quizá estemos estudiando un nuevo tratamiento del que no existe experiencia previa. Un abordaje es buscar datos en poblaciones humanas sobre enfermedades y tratamientos similares. También podemos buscar datos relevantes en estudios en animales. En ocasiones, simplemente no podemos

establecer estimaciones. En estos casos, podemos hacer una conjetura (p. ej., una mejoría del 30%) pero *limitando* la estimación: es decir, calcular el tamaño de la muestra necesario basándonos en una mejoría del 40% en la tasa de respuesta y calcular también el tamaño de la muestra necesario basándonos en una mejoría del 20% en la tasa de respuesta. Esto se conoce a veces como «análisis de sensibilidad».

En segundo lugar, debemos contar con una estimación del resultado clínico (tasa de curación, tasa de mejoría) en uno de los grupos. En el ejemplo que acabamos de exponer, dijimos que la tasa de curación actual (o la tasa de respuesta) es del 40%. Esta es la estimación de la tasa de respuesta para el grupo que recibe el tratamiento habitual basándonos en la experiencia clínica actual.

En tercer lugar, debemos especificar el nivel de α con el que estaremos satisfechos. La elección depende del investigador; no existe nada sagrado en ningún valor específico, pero generalmente en las investigaciones se utilizan valores de 0,05 o 0,01.

En cuarto lugar, debemos especificar la potencia del estudio. De nuevo, no existe ningún valor sagrado, pero habitualmente se utilizan potencias del 80% o del 90%. ¡No queremos pasar por alto una diferencia si existe alguna!

Por último, debemos especificar si la prueba va a ser unilateral o bilateral. ¿Qué significa esto? Nuestra tasa de curación actual es del 40% y vamos a estudiar un nuevo tratamiento que creemos que tendrá una tasa de curación más elevada, quizás del 50% o del 60%. Con el tratamiento nuevo queremos detectar una diferencia que sea en la dirección de la mejoría, un aumento de la tasa de curación. Por tanto, podríamos decir que solo estudiaremos en busca de una diferencia en esa dirección (mejoría), porque ésa es la dirección en la que estamos interesados; es decir, es una prueba unilateral.

El problema es que en la historia de la medicina y de la salud pública a veces nos hemos sorprendido de descubrir que tratamientos nuevos que pensábamos que serían beneficiosos, realmente han sido dañinos o se han acompañado de efectos secundarios graves. Si esta posibilidad es real, en nuestro estudio querríamos encontrar una diferencia en la tasa de curación en *cualquier dirección* respecto de la tasa actual, es decir, utilizaríamos una prueba bilateral, que estudiaría no solo una diferencia que sea mejor que la tasa de curación actual, sino también una que sea peor que la tasa de curación actual. Los médicos clínicos y otros investigadores a menudo prefieren utilizar una prueba unilateral en sus estudios porque dichas pruebas requieren muestras de menor tamaño que las pruebas bilaterales. Como el número de pacientes disponibles para estudios a menudo es limitado, las pruebas unilaterales son atractivas. En ocasiones, los investigadores pueden tomar la decisión práctica de emplear una prueba unilateral incluso aunque no haya justificación conceptual para esta decisión.

Sobre este tema existen opiniones divergentes. Hay quien cree que si el investigador solo está interesado en una dirección (mejoría) está justificado emplear una prueba unilateral. Otros creen que, siempre que la diferencia

pueda ir en cualquiera de las direcciones, es necesario emplear una prueba bilateral. En una situación en la que una enfermedad concreta es mortal en el 100% de los casos, cualquier diferencia con un tratamiento nuevo solo podría dirigirse en la dirección de la mejoría, por lo que sería apropiado utilizar una prueba unilateral.

Prestemos atención ahora a la aplicación de estos cinco factores para estimar el tamaño de muestra necesario a partir de una tabla de tamaño muestral. Las tablas 11.2 y 11.3 son selecciones de tablas de tamaños muestrales publicadas por Gehan en 1979[1]. (En muchos libros de estadística estándar existen apéndices con tablas similares.) Ambas tablas proporcionan el número de pacientes necesarios *en cada grupo* para detectar diversas diferencias en las tasas de curación con un α de 0,05 y una potencia $(1 - \beta)$ de 0,80. La tabla 11.2 está concebida para ser utilizada en una prueba bilateral y la tabla 11.3 para una prueba unilateral.

Tabla 11.2

Número de pacientes necesarios en cada grupo para detectar varias diferencias en las tasas de curación; α = 0,05; potencia (1 − β) = 0,80 (prueba bilateral)

La menor de las dos tasas de curación	DIFERENCIAS EN LAS TASAS DE CURACIÓN ENTRE LOS DOS GRUPOS DE TRATAMIENTO													
	0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40	0,45	0,50	0,55	0,60	0,65	0,70
0,05	420	130	69	44	36	31	23	20	17	14	13	11	10	8
0,10	680	195	96	59	41	35	29	23	19	17	13	12	11	8
0,15	910	250	120	71	48	39	31	25	20	17	15	12	11	9
0,20	1.090	290	135	80	53	42	33	26	22	18	16	12	11	9
0,25	1.250	330	150	88	57	44	35	28	22	18	16	12	11	–
0,30	1.380	360	160	93	60	44	36	29	22	18	15	12	–	–
0,35	1.470	370	170	96	61	44	36	28	22	17	13	–	–	–
0,40	1.530	390	175	97	61	44	35	26	20	17	–	–	–	–
0,45	1.560	390	175	96	60	42	33	25	19	–	–	–	–	–
0,50	1.560	390	170	93	57	40	31	23	–	–	–	–	–	–

Modificada de Gehan E. Clinical trials in cancer research. *Environ Health Perspect*. 1979;32:31.

Tabla 11.3

Número de pacientes necesarios en cada grupo para detectar varias diferencias en las tasas de curación; α = 0,05; potencia (1 − β) = 0,80 (prueba unilateral)

La menor de las dos tasas de curación	DIFERENCIAS EN LAS TASAS DE CURACIÓN ENTRE LOS DOS GRUPOS DE TRATAMIENTO													
	0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40	0,45	0,50	0,55	0,60	0,65	0,70
0,05	330	105	55	40	33	24	20	17	13	12	10	9	9	8
0,10	540	155	76	47	37	30	23	19	16	13	11	11	9	8
0,15	710	200	94	56	43	32	26	22	17	15	11	10	9	8
0,20	860	230	110	63	42	36	27	23	17	15	12	10	9	8
0,25	980	260	120	69	45	37	31	23	17	15	12	10	9	–
0,30	1.080	280	130	73	47	37	31	23	17	15	11	10	–	–
0,35	1.160	300	135	75	48	37	31	23	17	15	11	–	–	–
0,40	1.210	310	135	76	48	37	30	23	17	13	–	–	–	–
0,45	1.230	310	135	75	47	36	26	22	16	–	–	–	–	–
0,50	1.230	310	135	73	45	36	26	19	–	–	–	–	–	–

Modificada de Gehan E. Clinical trials in cancer research. *Environ Health Perspect*. 1979;32:31.

Supongamos que estamos realizando un ensayo clínico sobre dos tratamientos: uno que se utiliza habitualmente (el «tratamiento de referencia») y uno nuevo. El tratamiento habitual tiene una tasa de curación del 40% y creemos que el tratamiento nuevo puede tener una tasa de curación del 60%, es decir, queremos detectar una mejoría en la tasa de curación del 20%. ¿Cuántos sujetos tenemos que estudiar? Supongamos que utilizaremos un α de 0,05, una potencia del 80% y una prueba bilateral. Por tanto, emplearemos la tabla 11.2. La primera columna de esta tabla indica la menor de las dos tasas de curación. Como la tasa de curación actual es del 40% y con nuestro nuevo tratamiento esperamos una tasa de curación del 60%, la menor de las dos tasas es el 40%, por lo que nos fijamos en esa fila de la tabla. Esperamos que el tratamiento nuevo tenga una tasa de curación del 60%, por lo que la diferencia entre las tasas de curación es del 20%. Nos desplazamos hacia abajo en la columna del 20% (la diferencia en las tasas de curación) hasta el punto de intersección con la fila del 40% (la menor de las tasas de curación) y encontramos el valor 97. Por tanto, *en cada uno de los grupos de nuestro estudio* necesitamos 97 sujetos.

Otro método es utilizar la tabla en una dirección inversa. Por ejemplo, consideremos una clínica para pacientes que sufren una cierta enfermedad rara. Cada año la clínica trata a 30 pacientes con la enfermedad y quiere probar un tratamiento nuevo. Dado que el número máximo de pacientes es 30, podríamos preguntarnos: «¿Qué diferencia de tamaño podríamos esperar detectar en las tasas de curación?». Podemos encontrar una diferencia de un cierto tamaño que puede ser aceptable o podemos encontrar que el número de sujetos disponibles para el estudio es simplemente demasiado pequeño. Si el número de pacientes es demasiado pequeño, tenemos varias opciones: podemos decidir no realizar el estudio, y dicha decisión debería adoptarse pronto, antes de invertir un gran esfuerzo; o podríamos decidir prolongar el estudio en el tiempo para acumular más sujetos. Por último, podríamos

decidir colaborar con investigadores de otros centros para aumentar el número total de sujetos disponibles para el estudio. En un estudio que se realiza en un solo sitio, puede ser difícil identificar sesgos en la selección de participantes, pero en un estudio multicéntrico, la presencia de algún sesgo en uno de los centros sería detectable más fácilmente.

Esta sección ha demostrado el uso de una tabla de tamaño muestral. También existen disponibles fórmulas y programas informáticos para calcular el tamaño de las muestras. Los tamaños muestrales pueden calcularse no solo para ensayos aleatorizados, sino también para estudios de cohortes o de casos-controles, como se expuso anteriormente.

Reclutamiento y retención de participantes del estudio

Un desafío importante en la realización de los ensayos aleatorizados es el reclutamiento de un número suficiente de voluntarios elegibles y dispuestos. El fallo en el reclutamiento de un número suficiente de voluntarios puede dejar a un ensayo bien diseñado sin un número suficiente de participantes para lograr resultados estadísticamente válidos. Los participantes potenciales también deben estar dispuestos a ser aleatorizados para el ensayo. Los ensayos pueden retrasarse mucho por este problema del reclutamiento limitado y los costes para completar dichos ensayos pueden aumentar, ya que puede llevar más tiempo realizar todo el estudio. Sin embargo, debido a las presiones para reclutar a un número suficiente de participantes, se necesita un alto nivel de vigilancia para asegurase de que los investigadores del estudio no han empleado la coacción, manifiesta o encubierta, consciente o inconscientemente, para convencer a posibles participantes para que se incorporen a un estudio. Dentro de los límites de un ensayo aleatorizado, los participantes deben ser completamente informados de los riesgos y los acuerdos adoptados con fines de compensación si se produjesen efectos adversos. También se deben adoptar los acuerdos adecuados para retribuir los gastos de los participantes, como transporte, alojamiento si es necesario, y el tiempo dedicado a estos, en especial si la participación se asocia con pérdida de salario. Sin embargo, el pago de incentivos en efectivo a participantes potenciales supone riesgo de coacción manifiesta o sutil y puede dar lugar a sesgos y distorsión de los resultados del estudio, en especial si los incentivos pagados son cuantiosos.

En ocasiones, la incorporación como participante en un estudio ha sido publicitada a voluntarios potenciales con el argumento de que solo a través de la participación el participante tendrá la oportunidad de ser tratado con los tratamientos disponibles más novedosos. Sin embargo, la justificación para llevar a cabo un ensayo aleatorizado es que no sabemos qué tratamiento es mejor. Por tanto, resulta fundamental que las personas que realizan el ensayo eviten ser muy entusiastas prometiendo a los participantes beneficios que aún no han sido demostrados de modo concluyente que estén asociados con el tratamiento que se está estudiando. Los International Review Boards supervisan todas las cuestiones éticas implicadas en los estudios de investigación, y no se puede reclutar a ningún participante en la investigación hasta que se apruebe el protocolo del estudio.

Un problema relacionado es el de retener a los voluntarios durante toda la duración del estudio. Las pérdidas de seguimiento y otras formas de falta de cumplimiento pueden convertir este aspecto en un problema importante. Los participantes pueden perder interés en el estudio con el paso del tiempo o considerar la participación demasiado inadecuada, especialmente a largo

plazo (como cuando el seguimiento es de años). Los investigadores deben valorar por qué los participantes a menudo abandonan los estudios y adoptar las medidas adecuadas para evitar pérdidas de seguimiento.

Formas de expresar los resultados de los ensayos aleatorizados

Los resultados de los ensayos aleatorizados pueden expresarse de diversas formas. Pueden calcularse los riesgos de morir o de desarrollar una enfermedad o una complicación en cada grupo, y posteriormente puede calcularse la *reducción del riesgo* (eficacia). La *eficacia* del agente que se está estudiando, como una vacuna, puede expresarse mediante las tasas de desarrollar la enfermedad en el grupo vacunado y en el grupo al que se administra placebo:

$$\text{Eficacia} = \frac{\left(\begin{array}{c}\text{Tasa en los que}\\ \text{recibieron el placebo}\end{array}\right) - \left(\begin{array}{c}\text{Tasa en los que}\\ \text{recibieron la vacuna}\end{array}\right)}{\text{Tasa en los que recibieron el placebo}}$$

Esta fórmula nos informa de la cuantía de la disminución de la enfermedad gracias al uso de la vacuna. Los riesgos a menudo se calculan por *personas-años* de observación.

La eficacia, o cómo de bien funciona un tratamiento bajo condiciones «ideales» (como las de un ensayo clínico), puede diferenciarse de la efectividad, o cómo de bien funciona el tratamiento en situaciones «reales». Aunque los ensayos aleatorizados evalúan con mayor frecuencia la eficacia del tratamiento, los dos términos (eficacia y efectividad) a menudo se emplean indistintamente (de modo erróneo). La eficacia y la efectividad se analizan con mayor detalle más adelante.

Otra forma de comunicar los resultados de los ensayos aleatorizados es calculando la *razón de los riesgos* entre los dos grupos de tratamiento (el riesgo relativo), que se analizará más adelante. Además, con frecuencia comparamos las *curvas de supervivencia* en cada grupo, como hemos ilustrado previamente, para determinar si existen diferencias significativas.

Un objetivo importante de los ensayos aleatorizados es producir un efecto en la forma de ejercer la medicina clínica y en el ámbito de la salud pública. No obstante, en ocasiones los médicos pueden tener dificultades para situar los hallazgos de dichos ensayos en una perspectiva que sea relevante para su práctica. Por tanto, otro método para expresar los resultados de los ensayos aleatorizados es estimar el *número de pacientes que sería necesario tratar* (NNT)

para prevenir un resultado adverso, como una muerte. Esto puede calcularse del siguiente modo:

$$\text{NNT} = \frac{1}{\left(\begin{array}{c}\text{Tasa en el}\\ \text{grupo no tratado}\end{array}\right) - \left(\begin{array}{c}\text{Tasa en el}\\ \text{grupo tratado}\end{array}\right)}$$

Así, por ejemplo, si la tasa de mortalidad en el grupo no tratado es del 17% y la tasa de mortalidad en el grupo tratado es del 12%, necesitaríamos tratar:

$$\frac{1}{17\% - 12\%} = \frac{1}{0,05} = 20$$

personas para evitar una muerte. Las estimaciones del NNT suelen redondearse hacia arriba hasta el siguiente número entero. Este método puede emplearse en estudios de varias intervenciones, tanto para tratamientos como para medidas preventivas.

Por ejemplo, como se mencionó en el capítulo 10, en el European Randomized Study of Screening for Prostate Cancer, hubo una reducción de alrededor del 27% en la mortalidad por cáncer de próstata[2]. Por lo tanto, si en 100 individuos examinados se evitaran 27 muertes por cáncer de próstata, el cribado se habría realizado en (100 ÷ 27 = 1 ÷ 0.27) ≈ 4 individuos con el fin de evitar una muerte por cáncer de próstata. El mismo abordaje también puede utilizarse para valorar el riesgo de efectos secundarios calculando el *número necesario para dañar* (NND) para causar daño en una persona más. Estas estimaciones están sujetas a un error considerable y suelen presentarse con un intervalo de confianza del 95% para poder interpretarlas correctamente. Además, tienen otras limitaciones: no tienen en cuenta la calidad de vida y tienen un valor limitado para los pacientes. Estas estimaciones, no obstante, pueden ayudar a los médicos a estimar la cuantía del efecto que podrían esperar observar utilizando el nuevo tratamiento o la nueva medida preventiva en sus prácticas.

Interpretación de los resultados de los ensayos aleatorizados

Generalización de los resultados más allá de la población del estudio

Cuando realizamos un ensayo, el objetivo último es generalizar los resultados más allá de la población del estudio. Consideremos un ejemplo. Supongamos que queremos evaluar un nuevo fármaco para el lupus eritematoso sistémico (una enfermedad del tejido conjuntivo) utilizando un ensayo aleatorizado. Los diagramas de la figura 11.8 representan un ensayo aleatorizado en el que una población definida es identificada en el total de la población, y un subgrupo de esa población definida será la población del estudio. Por ejemplo, la *población total* podrían ser todos los pacientes con lupus eritematoso, la *población definida* podrían ser todos los pacientes con lupus eritematoso en nuestra comunidad y la *población del estudio* podrían ser los pacientes con la enfermedad que reciben asistencia médica en alguna de las distintas clínicas de nuestra comunidad.

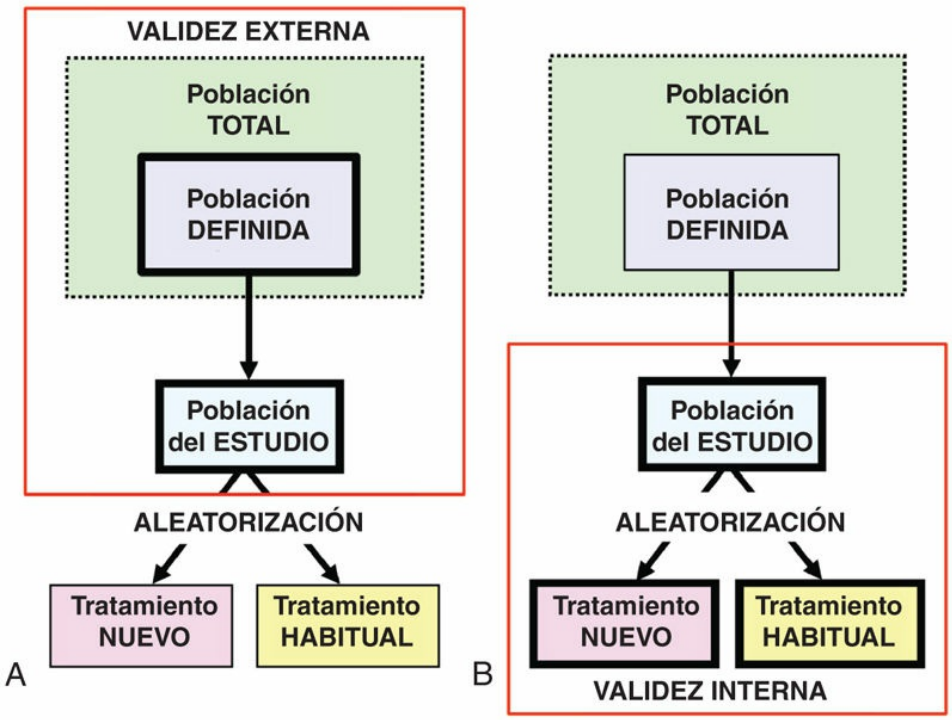


FIG. 11.8 (A) Validez externa (generalización) en un ensayo aleatorizado. Los hallazgos del estudio son generalizables de la población del estudio a la población definida y, presumiblemente, al total de la población. (B) Validez interna en un ensayo aleatorizado. El estudio se realizó correctamente y los hallazgos del estudio son, por tanto, válidos en la población del estudio.

Si realizamos un estudio en los pacientes reclutados de varias clínicas de nuestra comunidad y observamos que un tratamiento nuevo es mejor que el tratamiento empleado habitualmente (el tratamiento estándar), querríamos poder afirmar que el tratamiento nuevo es mejor para la enfermedad con independencia de dónde recibe el tratamiento el paciente, y no únicamente para los pacientes de esas clínicas. Nuestra capacidad para aplicar los resultados obtenidos en nuestra población de estudio a una población más general se denomina *generalización* o *validez externa* del estudio. Queremos ser capaces de *generalizar* a partir de los hallazgos del estudio a todos los pacientes con la enfermedad en nuestra comunidad. Para ello, debemos conocer hasta qué grado los pacientes que hemos estudiado son representativos de la población definida, es decir, de todos los pacientes con la enfermedad en cuestión en nuestra comunidad (v. fig. 11.8A). Debemos caracterizar a los que no participaron en el estudio e identificar características de los pacientes del estudio que pudieran ser diferentes de las de los pacientes que no participaron en el estudio. Dichas diferencias pueden descartar que podamos generalizar los resultados del estudio a otros pacientes que no fueron incluidos en el estudio. También podemos querer generalizar nuestros resultados, no solo a todos los pacientes con la enfermedad en nuestra comunidad, sino a todos los pacientes con la enfermedad, con independencia de dónde vivan, es decir, a la totalidad de pacientes con la enfermedad. Sin embargo, la población total de un ensayo aleatorizado raramente es conocida (menos aún enumerada). Aunque se espera que la población definida sea representativa de la población total, esta suposición raramente se verifica, si es que alguna vez se hace.

La validez externa debe diferenciarse de la *validez interna* (v. fig. 11.8B). Un ensayo aleatorizado tiene validez interna si la aleatorización se ha realizado correctamente y el estudio no sufre otros sesgos ni ninguno de los principales problemas metodológicos que hemos analizado. Esencialmente, se debería concluir idealmente que las diferencias observadas en los resultados en los dos grupos que se comparan se atribuyen a la exposición hipotética en estudio, además de los errores de muestreo. Los ensayos aleatorizados pueden considerarse el método de diseño de estudios de referencia, porque la aleatorización, si se realiza correctamente, evita sesgos por parte de los investigadores del estudio que puedan influir en la asignación del tratamiento para cada paciente. Si el estudio es lo suficientemente amplio, la aleatorización probablemente logrará la comparabilidad entre los grupos de tratamiento en cuanto a factores que pueden ser importantes para el resultado, como la edad, el sexo, la raza, etc., así como para factores que no hemos medido o de cuya importancia no somos conscientes. Los aspectos de la validez interna y de la validez externa (generalización) son puntos básicos a la hora de realizar cualquier ensayo aleatorizado y en los diseños de estudios observacionales, expuestos en capítulos anteriores. Debido a que los ensayos aleatorizados generalmente estudian voluntarios, su validez externa tiende a ser más baja que la de los estudios observacionales. Lo contrario es

cierto para la validez interna, dado que la comparabilidad entre grupos es mucho más probable en los ensayos clínicos.

¿Qué información pueden proporcionar a un médico en ejercicio los resultados de un ensayo aleatorizado acerca de un paciente concreto?

Consideremos un escenario hipotético sencillo. Un médico está a punto de recetar un tratamiento a uno de sus pacientes. Conoce un ensayo aleatorizado de gran calidad publicado recientemente que comparaba el tratamiento A y el tratamiento B para la patología que presenta su paciente (fig. 11.9A). Como se observa en el diagrama, en el ensayo la proporción de pacientes que lograron un buen resultado (partes azules de las barras) tras recibir el tratamiento A fue mucho mayor que la proporción de pacientes que lograron un buen resultado tras recibir el tratamiento B. Los resultados del ensayo se comunicaron, por tanto, mostrando que el tratamiento A es superior al tratamiento B para esta enfermedad.

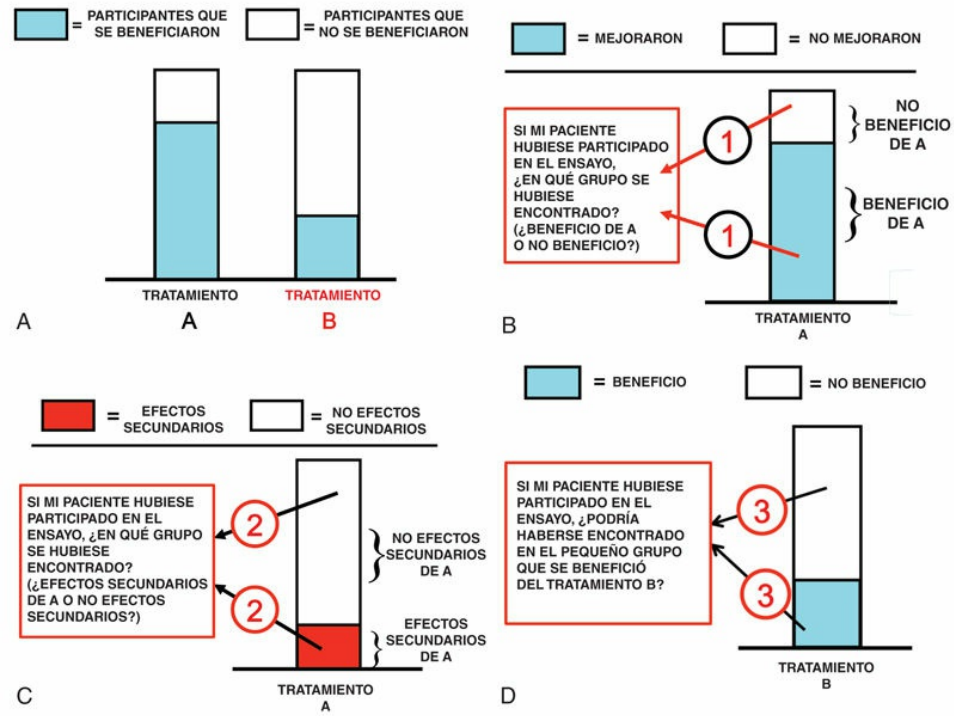


FIG. 11.9 (A) Resultados de un ensayo aleatorizado hipotético que compara un tratamiento A y un tratamiento B. Las áreas azules indican el número de pacientes que se beneficiaron de cada tratamiento y las áreas blancas indican los que no respondieron a cada tratamiento. (B) Primera pregunta del médico. (C) Segunda pregunta del médico. (D) Tercera pregunta del médico. (Véase la explicación en el texto.)

El médico conoce bien los resultados publicados del ensayo aleatorizado. Sin embargo, antes de recetar el tratamiento para su paciente basándose en

los resultados del ensayo, tiene algunas preguntas cuya respuesta le podría proporcionar una guía valiosa para elegir el mejor tratamiento *para este paciente*. A continuación, se exponen tres de sus preguntas a modo de ejemplo:

1. «Si mi paciente hubiese participado en el ensayo aleatorizado y hubiese sido asignado al grupo aleatorizado a recibir el tratamiento A (v. fig. 11.9B), ¿habría sido uno de los que mejoraron (se muestran en azul) o habría sido uno de los que no respondieron al tratamiento A (la parte blanca superior de la barra)?».
2. «Si mi paciente hubiese participado en el ensayo aleatorizado y hubiese sido asignado al grupo que recibió el tratamiento A (v. fig. 11.9C), ¿habría sido uno de los que sufrieron efectos secundarios (se muestran en rojo) o habría sido uno de los que no presentaron efectos secundarios con el tratamiento A (la parte blanca superior de la barra)?».
3. «Si mi paciente hubiese participado en el ensayo aleatorizado y hubiese sido asignado al grupo que recibió el tratamiento B (v. fig. 11.9D), ¿se habría encontrado en el grupo que mejoró tras recibir el tratamiento B (se muestra en azul) o se habría encontrado entre los que no respondieron al tratamiento B (la parte blanca superior de la barra)?»

Desafortunadamente, la mayoría de los ensayos aleatorizados no proporcionan la información que el médico necesitaría para caracterizar a un paciente concreto lo suficiente como para predecir qué respuesta podría tener ese paciente a los tratamientos disponibles. El médico por lo general no posee la suficiente información que le ayude a decidir si sería razonable generalizar a partir de los resultados del ensayo aleatorizado a un paciente específico antes de elegir e iniciar el tratamiento. Si generaliza a su paciente, ¿a partir de qué subgrupo de participantes en el ensayo debería generalizar?

Otro factor limitante en muchos ensayos aleatorizados es que, aunque asumamos que los abandonos del ensayo fueron mínimos y que todos los participantes aceptaron ser aleatorizados, quedan preguntas por contestar: ¿podemos asumir que en el mundo «real» no aleatorizado un paciente determinado respondería del mismo modo que un paciente aleatorizado podría responder en un ensayo? ¿Qué sabemos de la personalidad y las preferencias de los participantes en los ensayos aleatorizados que nos indicarían si un paciente específico que debe ser tratado posee características similares, como los mismos valores, personalidad y preocupaciones? ¿Una persona que acepta ser aleatorizada es parecida a la población general de la que un paciente específico puede proceder para recibir tratamiento? Como destacó David Mant, los participantes de los ensayos aleatorizados generalmente no son representativos de la población general[3]. Los participantes de los ensayos son por lo general más sanos, más jóvenes y

están mejor informados que los pacientes que acuden a ser tratados. Una última cuestión que se debe abordar es si hemos perdido nuestra preocupación sobre los individuos cuando reducimos a todo el mundo en el estudio a ser parte de un grupo de estudio y a menudo solo examinamos los resultados para el grupo como un todo, perdiendo de vista las diferencias y preferencias individuales.

Investigación comparativa de eficacia

Algunos ensayos aleatorizados están diseñados para comparar un tratamiento nuevo con un placebo. Otros ensayos aleatorizados se ocupan de la comparación de un tratamiento nuevo con un tratamiento más antiguo aceptado con el fin de determinar si el nuevo tratamiento es superior al tratamiento establecido. Más adelante en este capítulo estudiaremos dos ejemplos de ensayos utilizados para evaluar intervenciones ampliamente aceptadas. En los últimos años también ha surgido interés en lo que se ha denominado investigación comparativa de eficacia (ICE), en la que dos o más intervenciones existentes son comparadas con el fin de «determinar qué intervención sería más útil en una población dada o en un paciente determinado»[4]. En este tipo de abordaje, los resultados de otros tipos de diseños de estudios, que se analizaron en capítulos previos, pueden utilizarse conjuntamente con los hallazgos de ensayos aleatorizados para intentar responder a estas preguntas.

Otro aspecto es el coste de las intervenciones. Por ejemplo, muchos tratamientos de las infecciones por el virus de la inmunodeficiencia humana (VIH) son muy caros y dichos tratamientos pueden ser asequibles en países de altos ingresos, pero puede que no lo sean en muchos países de bajos y medianos ingresos. A medida que aparecen medicaciones más novedosas y menos caras (genéricos), a menudo se realizan estudios para determinar si las alternativas más nuevas y menos caras son igual de efectivas que las intervenciones más caras, cuya eficacia ya ha sido documentada. Estos estudios a menudo se denominan *estudios de equivalencia* y están diseñados para determinar si las intervenciones más baratas son igual de eficaces que los tratamientos más caros. Para estos estudios también se utiliza el término de *estudios de no inferioridad*. Estos estudios deben distinguirse de los *estudios de superioridad,* en los que fármacos de nueva aparición son evaluados para determinar si son más eficaces (superiores) que intervenciones disponibles en la actualidad.

Las cuatro fases para probar nuevos fármacos en Estados Unidos

A medida que aparecen nuevos fármacos, la Food and Drug Administration estadounidense sigue una secuencia estándar de cuatro fases para probar y evaluar estos nuevos agentes:

Ensayos de fase I. Estos ensayos son estudios farmacológicos clínicos: estudios pequeños de 20-80 pacientes que se ocupan de aspectos de seguridad del nuevo fármaco o de otros tratamientos. Examinan efectos tóxicos farmacológicos, como la seguridad, los márgenes de seguridad de las dosis en el ser humano y los efectos secundarios observados con el nuevo tratamiento. Si el fármaco pasa estas pruebas, a continuación se realizan estudios de *fase II*.

Ensayos de fase II. Los estudios de fase II consisten en investigaciones clínicas de 100-300 pacientes con el fin de evaluar la eficacia del nuevo fármaco o tratamiento y estudiar aún más su seguridad relativa. Si el fármaco pasa los estudios de fase II, a continuación pasa a ensayos de *fase III*.

Ensayos de fase III. Estos estudios son ensayos controlados aleatorizados a gran escala diseñados para valorar la eficacia y la seguridad relativa. Estos estudios a menudo se realizan con 1.000-3.000 o más participantes. El reclutamiento de esta gran cantidad de participantes puede ser muy difícil y a menudo necesita la participación de más de un centro de estudio. Cuando desde el comienzo se anticipan dificultades en el reclutamiento, el estudio puede diseñarse en la fase de planificación como un ensayo multicéntrico. Si el fármaco pasa la fase III puede ser aprobado y recibir la licencia para su comercialización.

Estudios de fase IV. Cada vez es un hecho más reconocido que ciertos efectos adversos de los medicamentos, como la carcinogénesis (cáncer) y la teratogénesis (malformaciones congénitas), pueden no manifestarse durante muchos años. También es posible que dichos efectos adversos de los nuevos fármacos puedan ser tan infrecuentes que puede que no se detecten incluso en ensayos clínicos aleatorizados relativamente extensos, o puedan volverse evidentes únicamente cuando el fármaco es usado por una gran cantidad de pacientes, una vez comercializado. Por este motivo, los estudios de *fase IV*, que también se conocen como de *vigilancia tras la comercialización*, son importantes para controlar nuevos fármacos que ya son utilizados por la población. Los estudios de fase IV no son estudios aleatorizados y en realidad no son ensayos, a diferencia de los ensayos de fase I, II y III. Como los estudios de fase IV estudian los

efectos secundarios de tratamientos nuevos una vez que el tratamiento ya se comercializa, los participantes no son aleatorizados. Con el fin de que los hallazgos de los estudios de vigilancia tras la comercialización sean válidos, resulta fundamental contar con un sistema de comunicación de efectos adversos de gran calidad. Aunque el objetivo de los estudios de fase IV a menudo es el número de efectos secundarios comunicados y el número de pacientes que recibieron el nuevo tratamiento y sufrieron efectos secundarios, los estudios de fase IV a menudo son muy valiosos para aportar pruebas adicionales sobre los beneficios y ayudan a optimizar el uso del nuevo agente.

La secuencia rigurosa que acabamos de describir ha protegido a la población estadounidense de muchos tratamientos peligrosos. En los últimos años, sin embargo, la presión para acelerar el procesamiento de nuevos fármacos para tratar la infección por VIH y el síndrome de inmunodeficiencia adquirida (SIDA) ha dado lugar a un replanteamiento de este proceso de aprobación. Parece probable que las modificaciones que terminen haciéndose en el proceso de aprobación no se limitarán a los fármacos utilizados para tratar el SIDA, sino que de hecho tendrán ramificaciones extensas en el proceso general de aprobación. Los cambios que se realicen en el futuro tendrán, por tanto, implicaciones importantes para la salud de los pacientes de Estados Unidos y de todo el mundo.

Cinco ensayos aleatorizados importantes en Estados Unidos

Hypertension Detection and Follow-up Program

La decisión de si controlar agresivamente o no la presión arterial en individuos hipertensos es un desafío clínico continuo. En la década de 1960, un estudio de la Veterans Administration demostró que el tratamiento de personas que sufren elevaciones importantes de la presión arterial puede reducir significativamente su mortalidad, lo que sentó las bases para el control de la presión arterial en Estados Unidos[5]. La cuestión de si el tratamiento antihipertensivo beneficia únicamente a personas con elevaciones leves de la presión arterial (presión arterial diastólica de 90-104 mmHg) no fue resuelta. Aunque podríamos ser capaces de reducir la presión arterial en dichas personas, debemos tener en cuenta el problema de los efectos secundarios de los fármacos antihipertensivos. A menos que pueda demostrarse algún beneficio para la salud de los pacientes, el uso de estos fármacos antihipertensivos no estaría justificado en las personas cuya presión arterial se encuentra mínimamente elevada.

El estudio multicéntrico Hypertension Detection and Follow-up Program (HDFP) fue diseñado, por tanto, para investigar los beneficios del tratamiento de la hipertensión leve a moderada. En este estudio, de 22.994 sujetos que fueron elegibles porque sufrían elevación de la presión arterial diastólica, 10.940 fueron aleatorizados a tratamiento escalonado o al grupo de tratamiento referido (fig. 11.10).

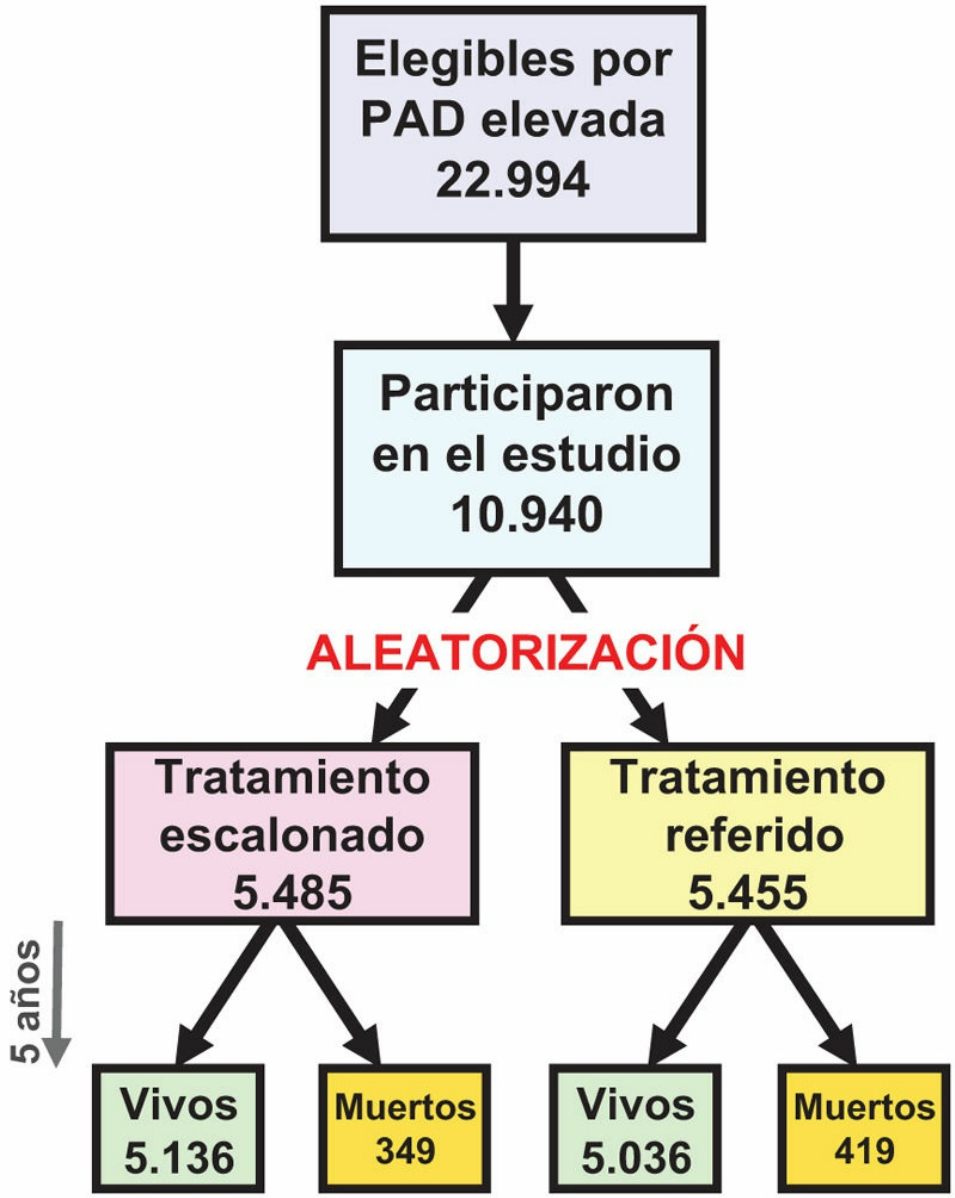


FIG. 11.10 Diseño del Hypertension Detection and Follow-up Program. *PAD*, presión arterial diastólica.

El *tratamiento escalonado* hacía referencia al tratamiento siguiendo un protocolo definido con precisión en el que el tratamiento se cambiaba cuando no se lograba la disminución especificada de la presión arterial durante un cierto periodo. El grupo de comparación suponía un problema: desde el punto de vista del diseño del estudio, hubiese sido deseable un grupo que no recibiese tratamiento para la hipertensión. Sin embargo, los investigadores creyeron que no sería éticamente justificable dejar sin tratamiento antihipertensivo a pacientes hipertensos. De modo que los pacientes del grupo de comparación fueron remitidos de vuelta a sus médicos (tratamiento habitual [TH]), y este grupo se denominó *grupo de tratamiento referido*. A continuación, se estudió la mortalidad en ambos grupos a lo largo de un periodo de 5 años[6].

En la figura 11.11 observamos que, en cada intervalo tras la incorporación al estudio, los pacientes del grupo de tratamiento escalonado presentaban

una mortalidad inferior que los del grupo de tratamiento referido. En dicha figura vemos que se mantuvo el mismo patrón en aquellos que únicamente presentaban elevaciones leves de la presión arterial.

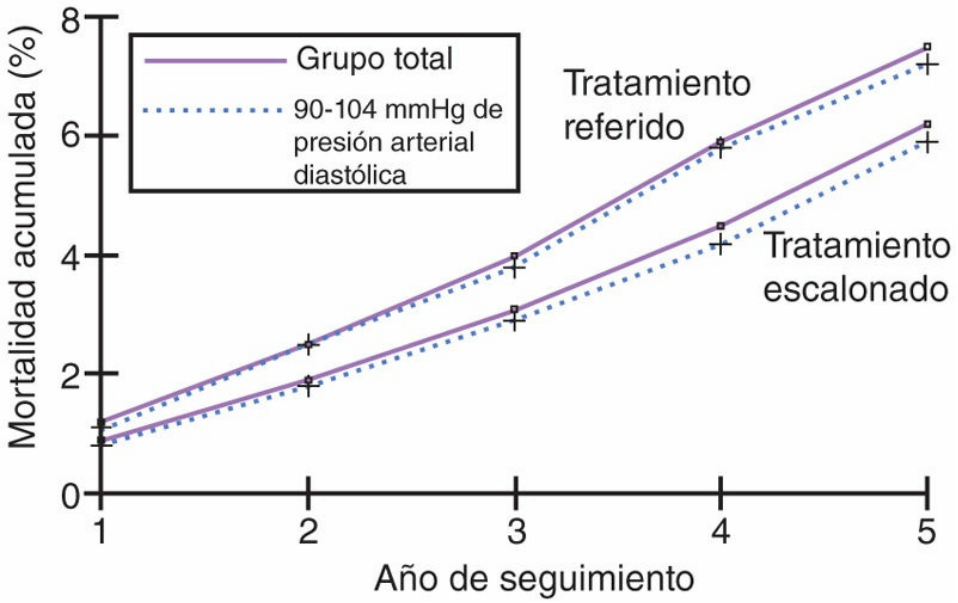


FIG. 11.11 Mortalidad acumulada por todas las causas en función del nivel de presión arterial y el tipo de tratamiento recibido en el Hypertension Detection and Follow-up Program. (Modificado de Hypertension Detection and Follow-up Program Cooperative Group: Five-year findings of the Hypertension Detection and Follow-up Program: I. Reduction in mortality of persons with high blood pressure, including mild hypertension. *JAMA.* 1979;242:2562–2571.)

Los resultados se exponen con mayor detalle en la tabla 11.4, en la que se presentan los datos en función de la presión arterial diastólica al incorporarse al estudio, antes del tratamiento antihipertensivo. La columna de la derecha muestra el porcentaje de reducción de la mortalidad en el grupo de tratamiento escalonado: la mayor reducción se produjo en los pacientes con una elevación leve de la presión diastólica.

Tabla 11.4

Mortalidad por todas las causas durante el Hypertension Detection and Follow-up Program

			TASA DE MORTALIDAD A 5 AÑOS		
Presión arterial diastólica al comienzo (mmHg)	Tratamiento escalonado	Tratamiento referido	TE	TR	Reducción de la mortalidad en el grupo de TE (%)
90-104	3.903	3.922	5,9	7,4	20,3
105-114	1.048	1.004	6,7	7,7	13,0
≥115	534	529	9,0	9,7	7,2
Total	5.485	5.455	6,4	7,7	16,9

TE, tratamiento escalonado; *TR*, tratamiento referido.

De Hypertension Detection and Follow-up Program Cooperative Group: Five-year findings of the Hypertension Detection and Follow-up Program: I. Reduction in mortality of persons with high blood pressure, including mild hypertension. *JAMA*. 1979;242:2562–2571.

Este estudio ha tenido un gran impacto y ha logrado que los médicos traten elevaciones incluso leves de la presión arterial. Sin embargo, ha sido criticado porque carecía de un grupo no tratado para comparación. No solo fueron remitidos estos pacientes de vuelta a sus médicos, sino que no hubo control del tratamiento que les fue proporcionado por sus médicos. Por tanto, la interpretación de estos datos es algo problemática. Incluso hoy, existe controversia acerca de si de hecho existió una objeción ética legítima a incluir un grupo no tratado en este estudio o si existió un problema ético a la hora de diseñar un estudio caro que fue difícil de organizar y que dejó tanta incertidumbre y dificultad en su interpretación.

Multiple Risk Factor Intervention Trial

Un problema grave de los ensayos a gran escala que requieren la inversión de gran cantidad de recursos, económicos y de otro tipo, y que se tardan años en completar es que su interpretación a menudo se ve empañada por un problema en el diseño o en la metodología que puede no haber sido apreciado en una fase inicial del estudio. El Multiple Risk Factor Intervention Trial (MRFIT) fue un estudio aleatorizado diseñado para determinar si la mortalidad por infarto de miocardio podría disminuir por cambios del estilo de vida y otras medidas. En este estudio, un grupo recibió una intervención especial (IE) que consistía en el tratamiento escalonado de la hipertensión y educación y formación intensiva sobre cambios del estilo de vida. El grupo de comparación recibió su TH en la comunidad. A lo largo de un periodo de seguimiento medio de 7 años, los factores de riesgo de cardiopatía coronaria (CC) disminuyeron más en los varones que recibieron la IE que en los que recibieron el TH (fig. 11.12).

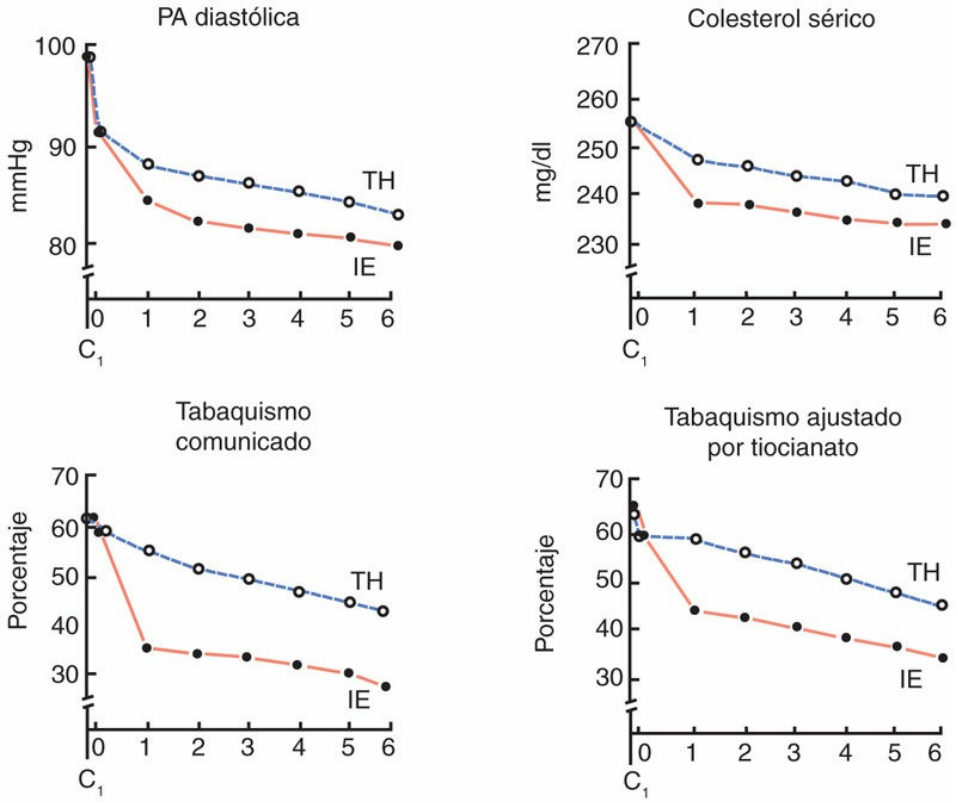


FIG. 11.12 Niveles medios de factores de riesgo por año de seguimiento en los participantes del Multiple Risk Factor Intervention Trial Research Group. *C_1,* primera visita de cribado; *IE,* intervención especial; *PA,* presión arterial; *TH,* tratamiento habitual. (De Multiple Risk Factor Intervention Trial Research Group: Multiple Risk Factor Intervention Trial: Risk factor changes and mortality results. *JAMA*. 1982;248:1465–1477.)

Sin embargo, al finalizar el estudio, no se observaron diferencias estadísticamente significativas entre los grupos, ni en la mortalidad por CC ni en la mortalidad por todas las causas (fig. 11.13). ¿Cuál pudo haber sido la causa?

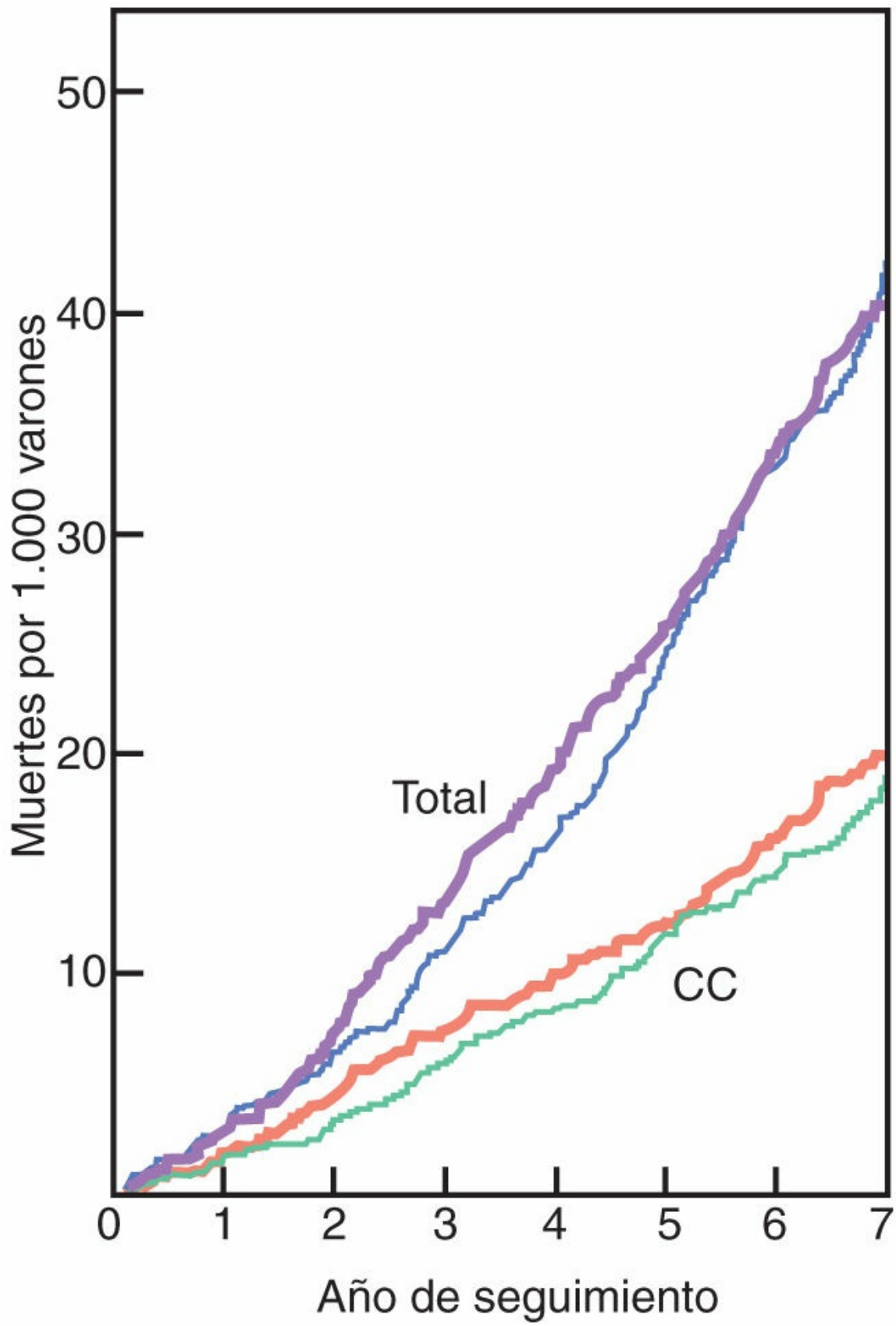


FIG. 11.13 Tasas de mortalidad acumuladas por cardiopatía coronaria *(CC)* y totales de los participantes del Multiple Risk Factor Intervention Trial Research Group. La línea gruesa indica los varones que reciben el tratamiento habitual; la línea fina indica los varones que reciben una intervención especial. (De Multiple Risk Factor Intervention Trial Research Group: Multiple Risk Factor Intervention Trial: Risk factor changes and mortality results. *JAMA*. 1982;248:1465–1477.)

La interpretación de estos resultados se vio complicada por problemas serios. En primer lugar, el estudio fue realizado en una época en la que la mortalidad por enfermedad coronaria estaba disminuyendo en Estados Unidos. Además, no quedó claro si la falta de diferencias encontrada en este estudio se debió a que los cambios del estilo de vida no eran un factor importante o porque el grupo control, por su cuenta, adoptó los mismos cambios de estilo de vida que adoptaron muchas otras personas en Estados Unidos en ese periodo. Gran parte de la población adoptó cambios como modificaciones generalizadas de la dieta, aumento del ejercicio e interrupción del tabaquismo, por lo que el grupo control pudo haberse «contaminado» con algunos de los cambios de conducta que habían sido recomendados en el grupo de estudio de modo formal y estructurado.

Este estudio también muestra el problema de usar medidas intermedias como puntos finales de la eficacia en los ensayos aleatorizados. Como todo efecto sobre la mortalidad puede tardar años en manifestarse, resulta tentador utilizar medidas que podrían verse afectadas antes por la intervención («indicadores indirectos»). Sin embargo, como se observa aquí, aunque la intervención fue exitosa para reducir el tabaquismo, los niveles de colesterol y la presión arterial diastólica, no podemos concluir basándonos en estos cambios que la intervención fue efectiva, porque el objetivo del estudio era determinar si la intervención podría reducir la mortalidad por CC, lo que no ocurrió.

Debido a estos problemas, que a menudo ocurren en los estudios muy amplios y caros, hay quien piensa que la inversión de los mismos fondos en diversos estudios más pequeños realizados por diferentes investigadores en diferentes poblaciones podría ser una elección más inteligente: si los resultados fueran consistentes, serían más creíbles debido a la generalización, a pesar de que las series individuales se acompañarían de los problemas de las muestras de menor tamaño (que afectan a la potencia estadística).

Estudio sobre prevención del cáncer de mama utilizando tamoxifeno

La observación de que las mujeres con cáncer de mama tratadas con tamoxifeno presentaban una menor incidencia de cáncer en la otra mama (contralateral) sugirió que el tamoxifeno podría ser útil para la prevención del cáncer de mama. Para estudiar esta hipótesis, se inició un ensayo aleatorizado en 1992. En septiembre de 1997, 13.388 mujeres de 35 o más años de edad se habían reclutado para el ensayo y habían sido asignadas aleatoriamente a recibir placebo o 20 mg de tamoxifeno al día durante 5 años. En marzo de 1998, un comité independiente que controlaba los datos decidió que las pruebas sobre la reducción del riesgo de cáncer de mama eran lo suficientemente sólidas como para justificar la interrupción del estudio antes de la fecha de finalización planeada. Como se observa en la figura 11.14, las tasas acumuladas del cáncer de mama invasivo y no invasivo se redujeron de

modo importante en las mujeres tratadas con tamoxifeno. Al mismo tiempo, como se observa en la figura 11.15, las tasas de cáncer endometrial invasivo aumentaron en el grupo tratado con tamoxifeno. Cuando se toma la decisión de utilizar tamoxifeno para la prevención del cáncer de mama, los beneficios potenciales del tamoxifeno deben sopesarse frente a la mayor incidencia de efectos adversos inesperados, en este caso el cáncer endometrial. El cuadro se ve aún más complicado por el hecho de que en la época en la que se publicaron los resultados de este estudio, dos estudios europeos más pequeños no observaron una reducción similar a la comunicada en el estudio americano. Así pues, nos encontramos ante el dilema del beneficio frente al daño; además, surge la duda de por qué otros estudios no han demostrado el mismo efecto destacado sobre la incidencia del cáncer de mama y cómo se deben tener en cuenta los resultados de dichos estudios a la hora de elaborar políticas públicas sobre esta materia.

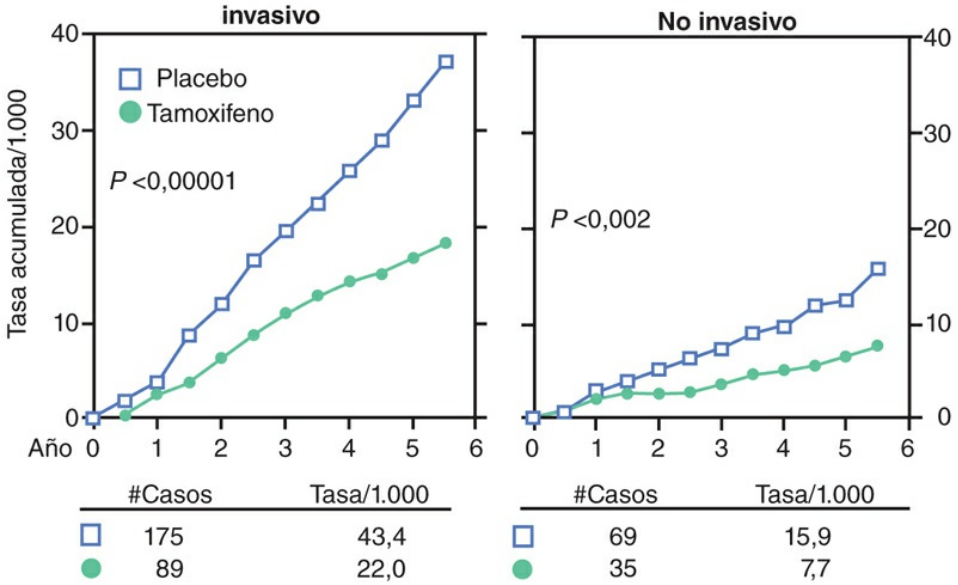


FIG. 11.14 Tasas acumuladas de cáncer de mama invasivo y no invasivo en participantes que reciben placebo o tamoxifeno. (De Fisher B, Costantino JP, Wickerham DL, et al. Tamoxifen for prevention of breast cancer: Report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. *J Natl Cancer Inst*. 1998;90:1371–1388.)

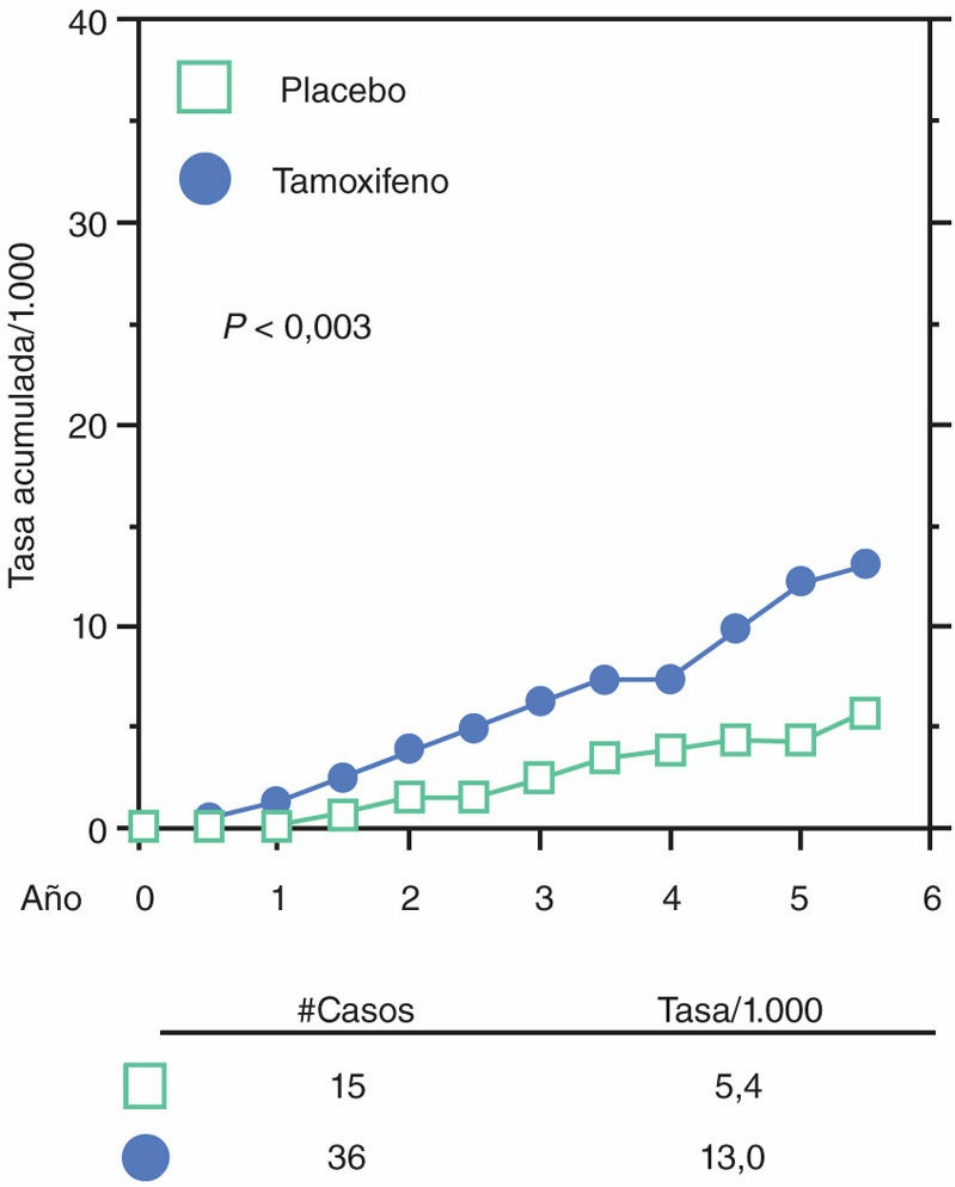


FIG. 11.15 Tasas acumuladas de cáncer de endometrio invasivo en participantes que reciben placebo o tamoxifeno. (De Fisher B, Costantino JP, Wickerham DL, et al. Tamoxifen for prevention of breast cancer: Report of the National Surgical Adjuvant Breast and Bowel Project P-1 Study. *J Natl Cancer Inst*. 1998;90:1371–1388.)

Ensayos aleatorizados para evaluar intervenciones ampliamente aceptadas

Los ensayos controlados aleatorizados pueden usarse con dos propósitos fundamentales: 1) para evaluar nuevas formas de intervención antes de que sean aprobadas y recomendadas para su uso general, y 2) para evaluar intervenciones que son muy controvertidas o que han sido ampliamente usadas o recomendadas sin haber sido evaluadas adecuadamente. Para evaluar el impacto que ejercen los ensayos controlados aleatorizados sobre la práctica médica, el segundo uso demuestra el desafío de cambiar los abordajes empleados en la práctica médica habitual que pueden no haber sido evaluados adecuadamente. En esta sección presentamos dos ejemplos sobre esta materia.

Un ensayo sobre la cirugía artroscópica de la rodilla por artrosis

Alrededor del 6% de los adultos de más de 30 años de edad y el 12% de los adultos mayores de 65 años sufren dolor de rodilla intenso como resultado de la artrosis. En Estados Unidos, una intervención realizada con frecuencia en los pacientes con dolor de rodilla y signos de artrosis ha sido la cirugía artroscópica con lavado o desbridamiento de la articulación de la rodilla utilizando un artroscopio. Se ha estimado que la intervención se realiza cada año en más de 225.000 adultos de mediana edad y de edad avanzada, con un coste anual de más de 1.000 millones de dólares.

En diversos ensayos controlados aleatorizados se compararon pacientes que fueron sometidos a un desbridamiento o a un lavado de la rodilla con controles que no recibieron tratamiento. Los pacientes tratados comunicaron más mejoría de su dolor de rodilla que los que no fueron tratados. Otros estudios, sin embargo, en los que solo se inyectó solución salina en la rodilla, también comunicaron mejoría de los síntomas. Así pues, resultó claro que los beneficios percibidos podrían relacionarse más con las expectativas del paciente (efecto placebo) que, con la eficacia real, porque la mejoría subjetiva comunicada por los pacientes era más probable cuando a los pacientes no se les ocultaba si recibían o no tratamiento quirúrgico. Para resolver el asunto de si el lavado o el desbridamiento artroscópico reduce los síntomas de dolor de rodilla en los pacientes con artrosis, se necesitaba un ensayo controlado aleatorizado en el que los controles fuesen sometidos a un tratamiento placebo. En julio de 2002, Moseley y cols.[7] publicaron los resultados de un ensayo aleatorizado muy bien realizado sobre esta intervención, utilizando una artroscopia placebo en los controles.

El diseño de este estudio se muestra en la figura 11.16. 180 veteranos fueron aleatorizados a grupos sometidos a desbridamiento artroscópico (59), lavado

artroscópico (61) o artroscopia placebo (60). La intervención placebo consistió en una incisión cutánea y en un desbridamiento simulado sin introducción de un artroscopio. Los parámetros medidos fueron el grado de dolor de la rodilla, determinado por encuestas, y el grado de función física, determinado por encuestas y observación directa. Estos resultados fueron valorados durante un periodo de 2 años. Los encargados de valorar el dolor y el estado funcional en los participantes, así como los mismos participantes, desconocían a qué grupo de tratamiento habían sido asignados.

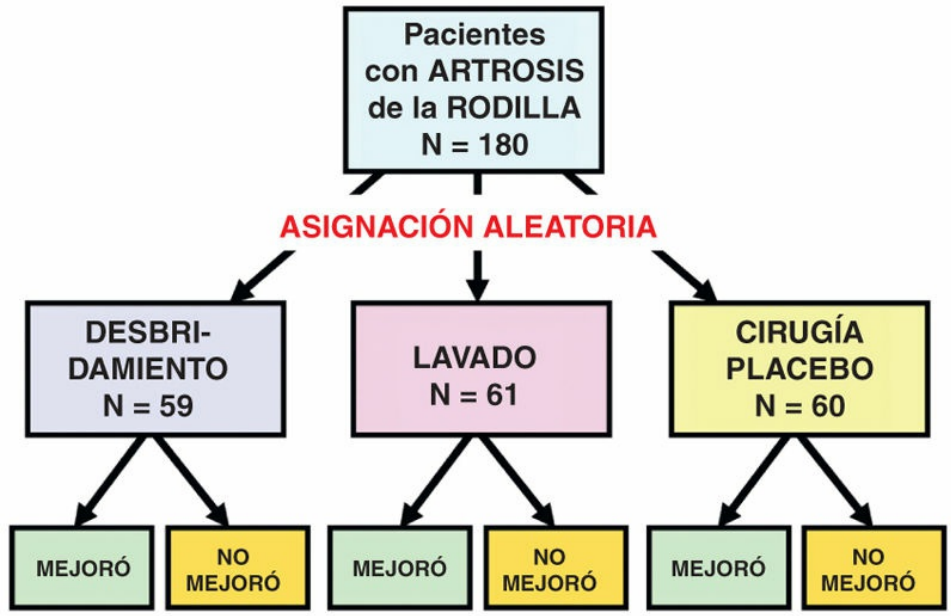
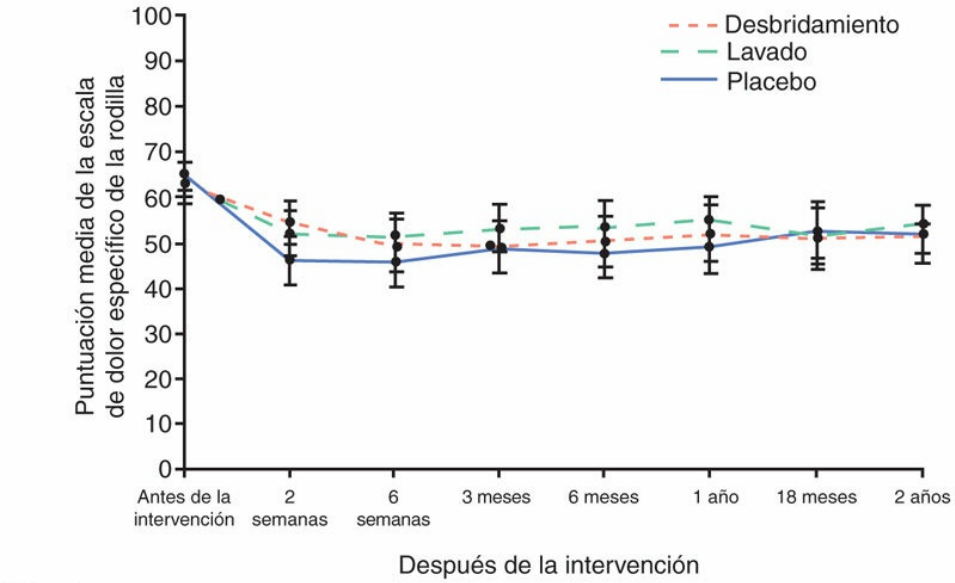


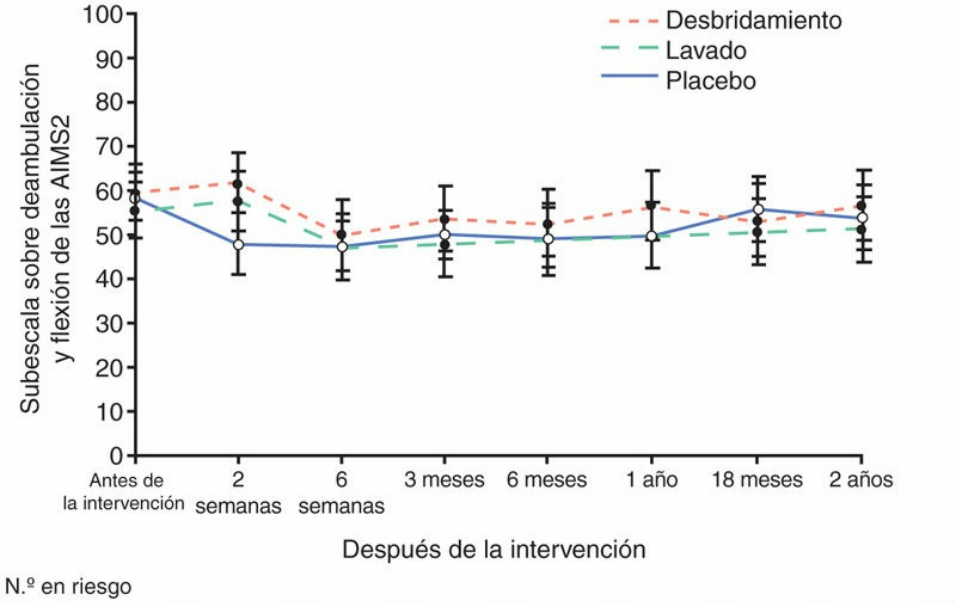
FIG. 11.16 Diseño de un ensayo controlado sobre cirugía artroscópica para la artrosis de la rodilla. (Basado en Moseley JB, O'Malley K, Petersen NJ, et al. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med*. 2002;347:81–88.)

Los resultados se muestran en las figuras 11.17 y 11.18. En ninguno de los grupos sometidos a intervención artroscópica se logró un mayor alivio del dolor que en el grupo sometido a la intervención placebo (v. fig. 11.17). Además, en ninguno de los grupos sometidos a alguna de las dos intervenciones artroscópicas se logró una mayor mejoría de la función física que en el grupo sometido a la intervención placebo (v. fig. 11.18).



N.º en riesgo								
Placebo	60	59	57	56	57	53	52	55
Lavado	61	59	57	59	59	57	56	55
Desbridamiento	58	59	59	58	56	50	51	53

FIG. 11.17 Valores medios (e intervalos de confianza del 95%) de la escala de dolor específico de la rodilla. Las determinaciones se realizaron antes de la intervención y 2 semanas, 6 semanas, 3 meses, 6 meses, 12 meses, 18 meses y 24 meses después de la intervención. Las puntuaciones más altas indican un dolor más intenso. (De Moseley JB, O'Malley K, Petersen NJ, et al. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med.* 2002;347[2]:81–88.)



N.º en riesgo								
Placebo	60	59	57	56	57	54	52	55
Lavado	61	59	57	59	59	57	57	56
Desbridamiento	58	58	59	58	55	51	51	53

FIG. 11.18 Valores medios (e intervalos de confianza del 95%) de la subescala sobre deambulación y flexión de las escalas de medición del impacto de la artrosis (AIMS2, por sus siglas en inglés). Las determinaciones se realizaron antes de la intervención y 2 semanas, 6 semanas, 3 meses, 6 meses, 12 meses, 18 meses y

24 meses después de la intervención. Las puntuaciones más altas indican una función peor. (De Moseley JB, O'Malley K, Petersen NJ, et al. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med.* 2002;347[2]:81–88.)

El investigador principal del estudio, el Dr. Nelda Wray, del Houston Veterans Affairs Medical Center, donde se realizó el estudio, resumió los resultados diciendo: «Nuestro estudio demuestra que la cirugía no es mejor que el placebo; la intervención por sí misma no es útil.» Un mes después de la publicación de este estudio, el Department of Veterans Affairs emitió una nota de recomendación a sus médicos exponiendo que no debería realizarse la intervención quirúrgica hasta que se publicaran estudios adicionales. Según la nota de recomendación, el dolor de rodilla no era un dato suficiente para indicar la cirugía a menos que también existieran signos de «alteraciones anatómicas o mecánicas» que presumiblemente pudieran mejorar con dicha intervención.

Efecto de los grupos de apoyo psicosocial en la supervivencia de las pacientes con cáncer de mama metastásico

En 1989 se realizó un estudio en el que las mujeres con cáncer de mama metastásico eran asignadas aleatoriamente a terapia de grupo de apoyo-expresiva (intervención conductual) o a un grupo control. La terapia de apoyo-expresiva es un tratamiento estandarizado para pacientes con enfermedades potencialmente mortales que fomenta que un grupo de participantes, dirigidos por un terapeuta, expresen sus sentimientos y preocupaciones acerca de la enfermedad y su impacto. Este estudio demostró un beneficio en la supervivencia, aunque en él no se había planeado inicialmente un análisis de supervivencia. Otros ensayos de otras intervenciones psicosociales no han demostrado un beneficio en la supervivencia.

Para aclarar este aspecto, Goodwin y cols.[8] realizaron un ensayo aleatorizado multicéntrico en el que 235 mujeres con cáncer de mama metastásico fueron aleatorizadas a un grupo que recibió terapia de apoyo-expresiva o a un grupo control que no recibió esta intervención (fig. 11.19). De las 235 mujeres, 158 fueron asignadas al grupo de intervención y 77 al grupo control.



FIG. 11.19 Diseño de un ensayo controlado aleatorizado sobre apoyo psicosocial en grupo en la supervivencia de pacientes con cáncer de mama metastásico. (Datos de Goodwin PJ, Leszcz M, Ennis M, et al. The effect of group psychosocial support on survival in metastatic breast cancer. *N Engl J Med*. 2001;345[24]:1719–1726.)

A lo largo del periodo del estudio, la supervivencia no se vio prolongada en las pacientes que recibieron terapia de apoyo-expresiva (fig. 11.20). Sin embargo, el estado de ánimo y la percepción del dolor sí mejoraron, especialmente en las mujeres que estaban más angustiadas al comienzo del estudio. Aunque los hallazgos en las publicaciones médicas son contradictorios acerca de la supervivencia y se siguen realizando estudios adicionales, los resultados de este estudio sugieren que esta intervención no se acompaña de beneficios en la supervivencia. Por tanto, los deseos de las mujeres que deciden hacer frente a su enfermedad de diferente forma, como no compartir sus sentimientos en un grupo, deben ser respetados. Además, no se debe sugerir a las mujeres que prefieren no participar en dicha terapia de grupo en ese momento difícil de sus vidas que su negativa puede estar acelerando su propia muerte.

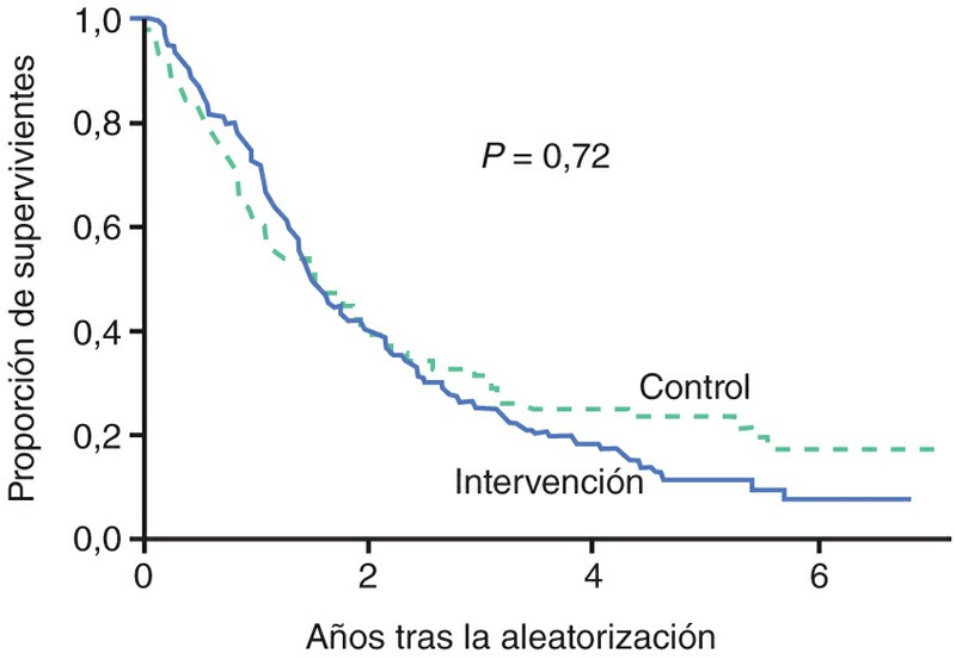


FIG. 11.20 Curvas de supervivencia de Kaplan-Meier de mujeres asignadas al grupo de intervención y al grupo control. No se observaron diferencias significativas en la supervivencia entre los dos grupos. (De Goodwin PJ, Leszcz M, Ennis M, et al. The effect of group psychosocial support on survival in metastatic breast cancer. *N Engl J Med.* 2001;345[24]:1719–1726.)

Registro de ensayos clínicos

Es un hecho conocido desde hace mucho tiempo que no se publican todos los resultados de los ensayos clínicos. Esto puede suponer un problema serio cuando se revisan los resultados de todos los ensayos clínicos publicados. Por ejemplo, si se revisan los ensayos clínicos de un nuevo fármaco pero solo se han publicado los que muestran resultados beneficiosos y no los que muestran resultados negativos (por algún motivo), de los estudios publicados podría llegarse a la conclusión errónea de que *todos* los estudios sobre el nuevo fármaco han mostrado un beneficio claro. Este tipo de problema frecuente se denomina *sesgo de publicación* o sesgo de no publicación. Por ejemplo, Liebeskind y cols.[9] identificaron 178 ensayos clínicos controlados sobre ictus isquémico agudo publicados en inglés a lo largo de un periodo de 45 años desde 1955 a 1999 por medio de una búsqueda sistemática de varias bases de datos de gran envergadura. Estos ensayos reclutaron a un total de 73.949 sujetos y evaluaron 75 fármacos u otros tipos de intervención. Encontraron que el problema del sesgo de publicación era un factor importante a la hora de revisar las publicaciones sobre los ictus isquémicos agudos. Era mucho más probable que *no* fuesen publicados los ensayos que demostraban que el fármaco estudiado era dañino que los ensayos en los que los resultados indicaban que el fármaco estudiado fue neutro o beneficioso.

Varios factores explican el problema de los sesgos de publicación. Las revistas médicas tienen más interés en publicar resultados de estudios que muestran efectos espectaculares que resultados de estudios que no encontraron beneficios con un fármaco nuevo, un dispositivo u otra intervención médica. Tanto los investigadores como las publicaciones están menos interesados en los estudios que concluyen que un tratamiento nuevo es inferior al tratamiento habitual o en los que los hallazgos no apuntan claramente hacia una u otra dirección (hallazgos nulos). Existe un factor más importante que está contribuyendo a este problema: las empresas que lanzan nuevos fármacos y financian los estudios sobre dichos fármacos con frecuencia prefieren no publicar los resultados cuando el fármaco estudiado es menos eficaz que los tratamientos ya disponibles. Las empresas están claramente preocupadas por si los resultados de dichos estudios pudieran influir negativamente en las ventas del producto y afectar a los grandes ingresos potenciales que habían calculado tener con el nuevo fármaco. El resultado neto, sin embargo, es la ocultación de los datos, lo que proporciona una visión del fármaco (incluyendo su eficacia y seguridad) que no es completa, por lo que los reguladores, los médicos y la población no pueden adoptar decisiones basadas en evidencias (es decir, decisiones basadas en la información total generada por los ensayos clínicos).

La importancia del riesgo para la salud pública por el hecho de comunicar selectivamente los resultados de los ensayos clínicos y la frecuencia con la que esta comunicación selectiva tiene lugar condujo al Committee of Medical

Journal Editors a adoptar una política, que entró en vigor en 2005, que obligaba a registrar *todos* los ensayos clínicos sobre intervenciones médicas en un registro público de ensayos antes de reclutar a participantes en el estudio[10]. Se consideran intervenciones médicas a los fármacos, las intervenciones quirúrgicas, los dispositivos, los tratamientos conductuales y los procesos de asistencia sanitaria. Antes de considerar la publicación de un ensayo clínico en alguna de las principales revistas médicas que adoptaron esta política, es obligatorio inscribirlo en un registro accesible al público sin coste alguno. El depósito federal de ensayos clínicos se encuentra en Internet en ClinicalTrials.gov. A partir de 2017 se espera que todos los ensayos clínicos financiados por los NIH se registren y presenten los resultados del estudio a ClinicalTrials.gov, de acuerdo con la *«NIH Policy on Dissemination of NIH-Funded Clinical Trial Information»* (NOT-OD-16-149)[11].

Consideraciones éticas

En el contexto de los ensayos clínicos surgen muchos aspectos éticos. Una pregunta que se plantea con frecuencia es si la aleatorización es en sí misma ética. ¿Cómo podemos dejar a pacientes sin un tratamiento farmacológico, especialmente cuando tienen enfermedades graves y potencialmente mortales? La aleatorización es ética solo cuando desconocemos si el fármaco A es mejor que el fármaco B. Podemos tener alguna indicación de que un tratamiento es mejor que el otro (a menudo este es el motivo para realizar un ensayo en primer lugar), pero no estamos seguros. A menudo, sin embargo, no está claro en qué momento «descubrimos» que el fármaco A es mejor que el fármaco B. Un mejor planteamiento de la pregunta es: ¿cuándo tenemos pruebas adecuadas que apoyen la conclusión de que el fármaco A es mejor que el fármaco B? Una cuestión que ha recibido mucha atención en los últimos años es si es ético utilizar un placebo[12]. Este aspecto conlleva implícito el tema de si es ético no administrar un tratamiento cuya eficacia ha sido demostrada[13].

La pregunta también puede plantearse de modo inverso: «¿Es ético no aleatorizar?». Cuando consideramos fármacos, medidas preventivas o sistemas de asistencia sanitaria de aplicación en gran cantidad de personas, tanto en Estados Unidos como en otros países, la norma puede ser llevar a cabo un ensayo aleatorizado que aclare los aspectos del beneficio y el daño y no seguir sometiendo a las personas a efectos tóxicos innecesarios y crear falsas esperanzas, a menudo con un gran coste. Por tanto, las dudas sobre la ética de la aleatorización deben plantearse en ambas direcciones: aleatorización o no aleatorización.

Otro punto importante es si puede obtenerse realmente un consentimiento informado en los estudios. Muchos protocolos de ensayos clínicos multicéntricos requieren la incorporación de los pacientes en el estudio inmediatamente tras el diagnóstico. El paciente puede que no sea capaz de otorgar su consentimiento en ese momento y los familiares pueden estar tan angustiados por el diagnóstico que acaban de recibir y por sus implicaciones que tienen gran dificultad en comprender la noción de la aleatorización y dar el visto bueno para que el paciente sea aleatorizado. Por ejemplo, gran parte del progreso alcanzado en las últimas décadas en el tratamiento de la leucemia infantil ha sido resultado de protocolos multicéntricos rigurosos que han precisado la incorporación del niño al estudio inmediatamente después de establecer el diagnóstico de la leucemia. Claramente, en unos momentos en los que los padres están tan angustiados, nos podemos cuestionar si son capaces de otorgar realmente su consentimiento informado. Sin embargo, el progreso ha tenido lugar gracias únicamente a dichos ensayos rigurosos, que han salvado las vidas de tantos niños con leucemia aguda.

Por último, ¿bajo qué circunstancias debería interrumpirse un ensayo antes de lo que se había planeado inicialmente? Este tema también es complicado y

puede surgir porque desde el principio se observan o efectos beneficiosos o efectos nocivos del agente estudiado, antes de haber reclutado a toda la muestra de participantes o antes de que los sujetos hayan sido estudiados durante el periodo de seguimiento completo. Muchos estudios cuentan con un comité externo examinador de datos y seguridad que revisa los datos a medida que son recibidos; el comité toma la decisión, como se ve, por ejemplo, en el Physicians' Health Study expuesto en el capítulo 10, en el que se estudiaban simultáneamente dos medicaciones en un diseño factorial: la aspirina se estudiaba para la prevención primaria de las enfermedades cardiovasculares y el beta-caroteno se estudiaba para la prevención primaria del cáncer. El comité externo examinador de datos decidió que los hallazgos sobre la aspirina eran lo suficientemente claros como para finalizar el estudio sobre la aspirina, pero que el estudio sobre el beta-caroteno debía continuar.

Conclusión

Los ensayos aleatorizados son el método de referencia para evaluar la eficacia de las medidas terapéuticas, preventivas y de otro tipo tanto en la medicina clínica como en el ámbito de la salud pública. En este capítulo y en el capítulo 10 se ha proporcionado una visión global de los métodos de diseño de estudio en los ensayos aleatorizados y las medidas empleadas para minimizar o evitar sesgos de selección y de otro tipo. Desde un punto de vista social, la generalización y los aspectos éticos son consideraciones importantes, y estos aspectos también se han analizado.

Epílogo

Concluiremos esta exposición sobre los ensayos aleatorizados citando un artículo de Caroline y Schwartz que fue publicado en la revista *Chest* en 1975. El artículo se titulaba «Chicken soup rebound and relapse of pneumonia: Report of a case»[14].

En la introducción los autores expusieron lo siguiente:

Desde hace mucho tiempo se sabe que el caldo de pollo posee una potencia terapéutica inusual frente a una gran variedad de agentes víricos y bacterianos. De hecho, ya en el siglo XII, el teólogo, filósofo y médico Moses Maimónides escribió: «El caldo de pollo [...] se recomienda como alimento excelente, así como medicación.» Estudios anecdóticos previos sobre la eficacia terapéutica de este agente, sin embargo, no han logrado proporcionar detalles sobre la duración adecuada del tratamiento. A continuación, exponemos un caso clínico en el que la interrupción abrupta del caldo de pollo dio lugar a una recidiva grave de una neumonía[14].

A continuación, los autores presentan el caso clínico de un médico de 47 años que fue tratado de neumonía con caldo de pollo. La administración de caldo de pollo se interrumpió prematuramente y el paciente sufrió una recidiva. Al no poder contar con más caldo de pollo, la recidiva fue tratada con penicilina intravenosa.

La exposición de los autores es de especial interés. A continuación, se muestra una parte:

La eficacia terapéutica del caldo de pollo fue descubierta por primera vez hace varios miles de años cuando una epidemia de gran mortalidad para los varones jóvenes de Egipto parecía no afectar a una minoría étnica que residía en la misma zona. La investigación epidemiológica contemporánea reveló que la dieta del grupo no afectado por la epidemia contenía grandes cantidades de una preparación cocinada hirviendo pollo con varias verduras y hierbas. Se debe destacar a este respecto que las órdenes relativas a la dieta dadas a Moisés en el monte Sinaí, aunque restringían el consumo de no menos de 19 tipos de aves, no incluían el pollo en la prohibición. Algunos eruditos creen que la receta del caldo de pollo fue transmitida a Moisés en la misma ocasión, pero fue relegada a la tradición oral cuando las escrituras fueron canonizadas. […] Aunque el caldo de pollo se utiliza ampliamente en la actualidad frente a una variedad de trastornos orgánicos y funcionales, su

elaboración está en gran medida en manos de particulares y la estandarización es casi imposible. Las investigaciones preliminares de la farmacología del caldo de pollo han demostrado que se absorbe inmediatamente tras su administración oral. […] No se recomienda la administración parenteral[14].

Este trabajo suscitó el envío de varias cartas al editor. En una, el Dr. Laurence F. Greene, catedrático de Urología de la Clínica Mayo, escribió:

Puede que esté interesado en saber que hemos tratado exitosamente la impotencia masculina con otro compuesto derivado del pollo, la hexametilacetil lututria tetrazolamina citarabina sódica (Schmaltz [Upjohn]). Este compuesto, cuando se aplica en pomada en el pene, no solo cura la impotencia, sino que también aumenta la libido y evita la eyaculación precoz. […] Los estudios preliminares indican que sus efectos dependen de la dosis en vista de que la relación sexual dura 5 minutos cuando se aplica la pomada al 5%, 15 minutos cuando se aplica la pomada al 15%, y así sucesivamente. Hemos recibido una beca de 650.000 dólares de la National Scientific Foundation para llevar a cabo un estudio prospectivo controlado aleatorizado doble ciego. Desafortunadamente, somos incapaces de obtener un número adecuado de sujetos debido a que todos los voluntarios se niegan a participar a menos que se les asegure que serán sujetos y no controles[15].

Bibliografía

1. Gehan E. Clinical trials in cancer research. *Environ Health Perspect*. 1979;32:31.
2. Schröder FH, Hugosson J, Roobol MJ, et al. Screening and prostate-cancer mortality in a randomized European study. *N Engl J Med*. 2009;360:1320–1328.
3. Mant D. Can randomized trials inform clinical decisions about individual patients? *Lancet*. 1999;353:743–746.
4. Institute of Medicine Initial National Priorities for Comparative Effectiveness Research. Washington, DC: National Academy Press; 2009: [fecha de última consulta: 5 de octubre de 2017]. http://www.nap.edu/catalog.php?record_id=12648.
5. Veterans Administration Cooperative Study Group on Hypertensive Agents Effects of treatment on morbidity in hypertension: results in patients with diastolic blood pressure averaging 115 through 129 mm Hg. *JAMA*. 1967;213:1028–1034.
6. Hypertension Detection and Follow-up Program Cooperative Group Five year findings of the Hypertension Detection and Follow-up Program: I. Reduction of mortality of persons with high blood pressure, including mild hypertension. *JAMA*. 1979;242:2562.
7. Moseley JB, O'Malley K, Petersen NJ, et al. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med*. 2002;347:81–88.
8. Goodwin PJ, Leszcz M, Ennis M, et al. The effect of group psychosocial support on survival in metastatic breast cancer. *N Engl J Med*. 2001;345:1719–1726.
9. Liebeskind DS, Kidwell CS, Sayre JW, et al. Evidence of publication bias in reporting acute stroke clinical trials. *Neurology*. 2006;67:973–979.
10. DeAngelis CD, Drazen JM, Frizelle FA, eds. Clinical trial registration: a statement from the International Committee of Medical Journal. *JAMA*. 2004;292:1363–1364.
11. National Institutes of Health. *NIH policy on the dissemination of NIH funded clinical trial information*. https://grants.nih.gov/grants/guide/notice-files/NOT-OD-16-

149.html. Consultado el 5 de octubre de 2017.
12. Emanuel EJ, Miller FG. The ethics of placebo-controlled trials: a middle ground. *N Engl J Med*. 2001;345:915–919.
13. Huston P, Peterson R. Withholding proven treatment in clinical research. *N Engl J Med*. 2001;345:912–914.
14. Caroline NL, Schwartz H. Chicken soup rebound and relapse of pneumonia: report of a case. *Chest*. 1975;67:215–216.
15. Greene LF. The chicken soup controversy [letter]. *Chest*. 1975;68:605.

Preguntas de repaso de los capítulos 10 y 11

1. *El principal objetivo de la asignación aleatoria en un ensayo clínico es:*

a. Ayudar a asegurar que los sujetos del estudio son representativos de la población general.

b. Facilitar el doble ciego (enmascaramiento).

c. Facilitar la medición de las variables del resultado.

d. Asegurar que los grupos del estudio poseen características basales comparables.

e. Reducir el sesgo de selección en la asignación del tratamiento.

2. *Un anuncio en una revista médica decía: «Dos mil pacientes con faringitis fueron tratados con nuestra nueva medicina. En 4 días, el 94% se encontraban asintomáticos.» El anuncio asegura que la medicina es eficaz. Basado en lo expuesto anteriormente, la afirmación:*

a. Es correcta.

b. Puede ser incorrecta porque la conclusión no está basada en una tasa.

c. Puede ser incorrecta porque no reconoce el fenómeno de cohorte a largo plazo.

d. Puede ser incorrecta porque no se utilizó ninguna prueba estadísticamente significativa.

e. Puede ser incorrecta porque no se utilizó un grupo control o de comparación.

3. *El objetivo de un estudio* doble ciego *o con* doble enmascaramiento *es:*

a. Lograr comparabilidad entre los sujetos tratados y no tratados.

b. Reducir los efectos de la variación del muestreo.

c. Evitar sesgos del observador y del sujeto.

d. Evitar sesgos del observador y variación del muestreo.

e. Evitar sesgos del sujeto y variación del muestreo.

4. *En muchos estudios que examinaban la asociación entre los estrógenos y el cáncer uterino endometrial se utilizó una prueba significativa unilateral. La suposición subyacente que justificó una prueba unilateral en vez de bilateral fue que:*

a. La distribución de la proporción expuesta seguía un patrón «normal».

b. Antes de realizar el estudio se pensaba que los estrógenos causaban cáncer uterino endometrial.

c. El patrón de asociación podría expresarse mediante una función en línea recta.

d. El error de tipo II era el error potencial más importante que se debía evitar.

e. Solo se utilizó un grupo control.

5. *En un ensayo aleatorizado, un diseño cruzado planeado:*

a. Elimina el problema de un posible efecto de orden.

b. Debe tener en cuenta el problema de posibles efectos residuales del primer tratamiento.

c. Necesita aleatorización estratificada.

d. Elimina la necesidad de controlar el cumplimiento o la falta de cumplimiento.

e. Mejora la generalización de los resultados del estudio.

6. *Un ensayo aleatorizado que comparaba la eficacia de dos fármacos demostró una diferencia entre los dos (con un valor de* $P < 0{,}05$*). Suponga que, sin embargo, los dos fármacos en realidad no se diferencian. Este es, por tanto, un ejemplo de lo siguiente:*

a. Error de tipo I (error α).

b. Error de tipo II (error β).

c. $1 - \alpha$.

d. $1 - \beta$.

e. Ninguno de los anteriores.

7. *Todos los siguientes son beneficios potenciales de un ensayo clínico*

aleatorizado excepto:

a. La probabilidad de que los grupos del estudio sean comparables es mayor.

b. Se elimina la autoselección para un tratamiento particular.

c. La validez externa del estudio es mayor.

d. La asignación del siguiente sujeto no puede predecirse.

e. El tratamiento que recibe un sujeto no está influido por sesgos conscientes o subconscientes del investigador.

Número de pacientes necesarios en un grupo control y en un grupo experimental para una probabilidad dada de obtener un resultado significativo (prueba bilateral)

La menor de las dos	DIFERENCIAS EN LAS TASAS DE CURACIÓN ENTRE LOS DOS GRUPOS DE TRATAMIENTO					
tasas de curación	0,05	0,10	0,15	0,20	0,25	0,30
0,05	420	130	69	44	36	31
0,10	680	195	96	59	41	35
0,15	910	250	120	71	48	39
0,20	1.090	290	135	80	53	42
0,25	1.250	330	150	88	57	44
0,30	1.380	360	160	93	60	44
0,35	1.470	370	170	96	61	44
0,40	1.530	390	175	97	61	44

α = 0,05; potencia (1 − β) = 0,80.

Datos de Gehan E: Clinical trials in cancer research. Environ Health Perspect 1979;32:31.

La pregunta 8 se basa en la tabla precedente:

8. *Una compañía farmacéutica sostiene que un nuevo fármaco G para una cierta enfermedad posee una tasa de curación del 50% en comparación con el fármaco H, cuya tasa de curación es solo del 25%. Se le encarga a usted que diseñe un ensayo clínico para comparar los fármacos G y H. Utilizando la tabla precedente, calcule el número de pacientes necesarios en cada grupo de tratamiento para detectar dicha diferencia con unos valores de α = 0,05, bilateral, y β = 0,20.*

El número de pacientes necesarios en cada grupo de tratamiento es

_____.

Utilice sus conocimientos sobre diseño de estudios para responder la pregunta 9.

9. *Elija el mejor diseño de estudio de la siguiente lista para cada una de las siguientes preguntas de investigación. Cada diseño de estudio solo puede ser usado una vez.*

a. Estudio ecológico.

b. Estudio transversal.

c. Estudio de casos y controles.

d. Cohorte prospectiva.

e. Ensayo aleatorizado.

9a. _____ Un investigador desea determinar si la prevalencia de sífilis es mayor entre los hombres que entre las mujeres.

9b. _____ Un investigador cree que una enfermedad rara puede estar asociada con el uso de un fertilizante común para césped.

9c. _____ Las autoridades sanitarias de la ciudad de Baltimore creen que un programa para dejar de fumar en combinación con el uso de parches de nicotina será más efectivo que un programa para dejar de fumar solo.

9d. _____ Los investigadores quieren determinar el riesgo de posibles efectos adversos sobre la salud asociados al consumo excesivo de alcohol entre los adultos jóvenes.

9e. _____ Un investigador cree que la tasa de hospitalización a nivel provincial por enfermedades cardiovasculares aumentará con los niveles crecientes de contaminación del aire exterior.

Las preguntas 10 y 11 se basan en el siguiente ensayo controlado aleatorizado, que estudia los efectos de un fármaco para reducir los accidentes cerebrovasculares recurrentes:

Dado el aumento de la carga de la enfermedad cardiovascular, un investigador diseña un ensayo controlado aleatorizado dirigido a pacientes que han sufrido un accidente cerebrovascular en los últimos 30 días. El ensayo quiere evaluar si el fármaco A reduce la probabilidad de recurrencia del accidente cerebrovascular en comparación con el tratamiento habitual (ensayo de superioridad). En el ensayo, 300 pacientes con accidente cerebrovascular se

distribuyen aleatoriamente en dos grupos, en los que el grupo 1 recibe el fármaco A y el grupo 2 recibe el tratamiento habitual. El investigador comparó la incidencia acumulada de accidente cerebrovascular recurrente entre ambos grupos. Algunos pacientes aleatorizados a seguir el tratamiento habitual recibieron el fármaco A por otros medios. Por otro lado, algunos participantes asignados al azar a recibir el fármaco A terminaron no tomándolo. Los resultados del ensayo controlado aleatorizado son los siguientes:

	GRUPO 1		GRUPO 2	
	ALEATORIZADO AL FÁRMACO A		ALEATORIZADO AL TRATAMIENTO HABITUAL	
	Tomó el fármaco A	Tratamiento habitual	Tomó el fármaco A	Tratamiento habitual
Sufrió un accidente cerebrovascular recurrente	10	20	5	40
No sufrió un accidente cerebrovascular recurrente	80	40	15	90

10. *Según el análisis de la intención de tratar, calcule la tasa de incidencia acumulada del accidente cerebrovascular recurrente utilizando el tratamiento habitual como referencia. Las respuestas deben redondearse a dos decimales.*

Tasa de incidencia acumulada = _____

11. *Según el análisis por protocolo, calcule la tasa de incidencia acumulada para el accidente cerebrovascular recurrente utilizando el tratamiento habitual como referencia. Las respuestas deben redondearse a dos decimales.*

Tasa de incidencia acumulada = _____

CAPÍTULO 12

Estimación del riesgo: ¿existe una asociación?

Palabras clave

riesgo absoluto
medidas de asociación
cociente de posibilidades (posibilidades relativas)
riesgo relativo

Objetivos de aprendizaje

- Revisar el concepto de riesgo absoluto.
- Introducir y comparar el riesgo relativo y el cociente de posibilidades (*odds ratio,* OR) como medidas de asociación entre una exposición y una enfermedad.
- Calcular e interpretar un riesgo relativo en un estudio de cohortes.
- Calcular e interpretar el cociente de posibilidades en un estudio de cohortes y en un estudio de casos y controles, y describir cuándo este cociente es una buena estimación del riesgo relativo.
- Calcular e interpretar el cociente de posibilidades en un estudio de casos y controles de datos emparejados.

En los capítulos anteriores se han descrito los diseños de estudio básicos que se utilizan comúnmente en las investigaciones epidemiológicas. Estos se muestran esquemáticamente en las figuras 12.1 a 12.3.



FIG. 12.1 Diseño de un ensayo clínico aleatorizado.



FIG. 12.2 Diseño de un estudio de cohortes.

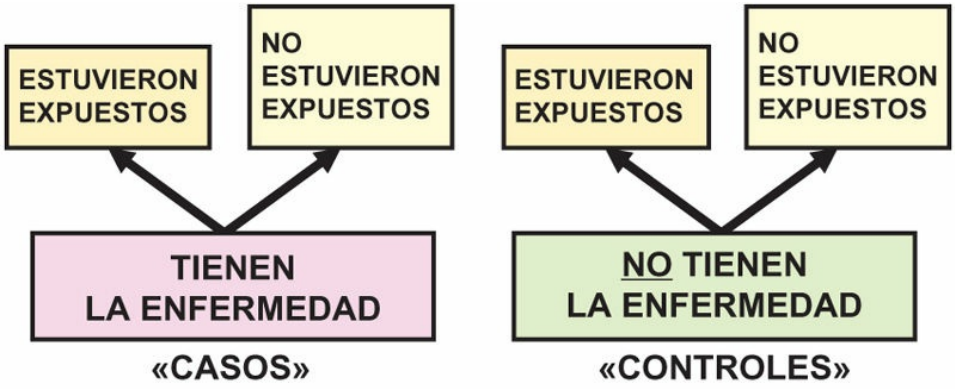


FIG. 12.3 Diseño de un estudio de casos y controles.

Debe recordarse que la diferencia fundamental entre un ensayo clínico aleatorizado y un estudio de cohortes es que, en este último, los sujetos no son asignados al azar para ser expuestos o no expuestos, porque es evidente que la asignación aleatoria a la exposición a agentes tóxicos o posiblemente cancerígenos no sería aceptable. Aparte de eso, los estudios de cohorte y los ensayos aleatorios son esencialmente equivalentes. Por tanto, los estudios de cohortes se utilizan en muchas investigaciones sobre la etiología, ya que este diseño de estudio permite aprovechar las poblaciones que han tenido una exposición específica documentada y compararlas con las que no han tenido esa exposición. Los estudios de casos y controles también se utilizan para abordar cuestiones sobre la etiología, aunque a menudo en una fase más exploratoria. Con independencia de qué diseño se use, el objetivo es determinar si hay un exceso de riesgo (incidencia), o tal vez un riesgo menor (protección), de una determinada enfermedad en asociación con una cierta exposición específica o característica. Como ha sido mencionado anteriormente, la incidencia es una medida del riesgo de enfermedad. El riesgo puede definirse como la probabilidad de que se produzca un evento (como el desarrollo de una enfermedad).

Antes de describir estas estrategias comparativas, se expondrá el concepto de riesgo absoluto.

Riesgo absoluto

La incidencia de una enfermedad en una población se denomina *riesgo absoluto.* El riesgo absoluto puede indicar la magnitud del riesgo en un grupo de personas con una exposición determinada, pero, debido a que no tiene en cuenta el riesgo de la enfermedad en los individuos no expuestos, no indica si la exposición se asocia con un mayor riesgo de la enfermedad. La comparación es fundamental para la epidemiología. Sin embargo, el riesgo absoluto puede tener implicaciones significativas tanto en la medicina clínica como en la política de salud pública: por ejemplo, una mujer que contrae la rubéola durante el primer trimestre del embarazo y pregunta a su médico cuál es el riesgo de malformación de su hijo recibirá un cierto número como respuesta. Basándose en esta información, puede decidir abortar o continuar su embarazo. Ella no recibe explícitamente datos comparativos, pero por lo general se está haciendo una comparación implícita: la mujer no solo desea saber cuál es su riesgo, sino también cómo es en comparación con la situación si no hubiese contraído la rubéola (esto es lo que se denomina el *contrafactual* en términos de inferencia causal, que se estudiará más adelante en este libro). Por tanto, aunque el riesgo absoluto no establece ninguna comparación explícita, suele realizarse una comparación implícita siempre que se observa la incidencia de una enfermedad. Sin embargo, para abordar la cuestión de la asociación, hay que utilizar estrategias que implican comparaciones explícitas.

¿Cómo se puede determinar si una enfermedad concreta se asocia a una exposición específica?

Para determinar si existe una asociación entre una exposición específica y una enfermedad concreta, se debe determinar, a menudo a partir de los datos obtenidos en estudios transversales, de casos y controles o de cohortes, si existe un exceso de riesgo de la enfermedad en las personas que han estado expuestas a un determinado agente. A modo de ejemplo, se considerarán los resultados de una investigación sobre un brote de enfermedad transmitida por los alimentos. Los alimentos sospechosos se identificaron y, para cada uno de ellos, se calculó la tasa de ataque (o tasa de incidencia) de la enfermedad para quienes comieron dicho alimento (expuestos) y para quienes no lo comieron (no expuestos), como se muestra en la tabla 12.1.

Tabla 12.1

Brote de enfermedad transmitida por alimentos: I. Porcentaje de enfermos entre los que comieron y los que no comieron alimentos específicos

Alimento	Comieron (% enfermos)	No comieron (% enfermos)
Ensalada de huevo	83	30
Macarrones	76	67
Queso cottage	71	69
Ensalada de atún	78	50
Helado	78	64
Otros	72	50

¿Cómo se puede determinar si un exceso de riesgo se asocia a cada uno de los alimentos? Una estrategia, mostrada en la columna C de la tabla 12.2, consiste en calcular el *cociente* de la tasa de ataque (los que sospechamos que estaban enfermos) en quienes comieron cada alimento entre la tasa de ataque en quienes no comieron el alimento. Una estrategia alternativa para identificar cualquier exceso de riesgo en los individuos expuestos se muestra en la columna D. Se puede restar el riesgo de quienes no comieron el alimento del riesgo de quienes sí lo comieron. La *diferencia* representa el exceso de riesgo absoluto en quienes estuvieron expuestos.

Tabla 12.2

Brote de enfermedad transmitida por alimentos: II. Formas de calcular el exceso de riesgo

	(A)	(B)	(C)	(D)
Alimento	Comieron (% enfermos)	No comieron (% enfermos)	(A)/(B)	(A) – (B) (%)
Ensalada de huevo	83	30	2,77	53
Macarrones	76	67	1,13	9
Queso cottage	71	69	1,03	2
Ensalada de atún	78	50	1,56	28
Helado	78	64	1,21	14
Otros	72	50	1,44	22

Por tanto, como se ve en este brote de origen alimentario, para determinar si una cierta exposición se asocia a una cierta enfermedad, se debe determinar si hay un exceso de riesgo de la enfermedad en las poblaciones expuestas comparando el riesgo de la enfermedad (en este caso, la intoxicación alimentaria) en poblaciones expuestas con el riesgo de la enfermedad en poblaciones no expuestas. Como acabamos de ver, un exceso de riesgo de este tipo puede calcularse de las dos formas siguientes:

1. El *cociente* de los riesgos (o de las tasas de incidencia):

$$\frac{\text{Riesgo de enfermedad en expuestos}}{\text{Riesgo de enfermedad en no expuestos}}$$

2. La *diferencia* entre los riesgos (o entre las tasas de incidencia):

$$\left(\begin{array}{c}\text{Riesgo de enfermedad}\\ \text{en expuestos}\end{array}\right) - \left(\begin{array}{c}\text{Riesgo de enfermedad}\\ \text{en no expuestos}\end{array}\right)$$

Aquí surge la pregunta de si el método que se elige para calcular el exceso de riesgo supone alguna diferencia. A continuación, se expone un ejemplo hipotético de dos comunidades, A y B (se presenta en la tabla 12.3).

Tabla 12.3

Ejemplo que compara dos formas de calcular el exceso de riesgo

	POBLACIÓN	
	A	B
Incidencia (%)		
En expuestos	40	90
En no expuestos	10	60
Diferencia entre las tasas de incidencia (%)	30	30

Cociente de las tasas de incidencia	4,0	1,5

En la comunidad A, la incidencia de una enfermedad en las personas expuestas es del 40% y la incidencia en las personas no expuestas, del 10%. La duda es si hay un exceso de riesgo asociado con la exposición. Al igual que en el ejemplo de la intoxicación alimentaria, se puede calcular el cociente de las tasas o la diferencia entre dichas tasas. El *cociente* de las tasas de incidencia es 4,0. Si se calcula la *diferencia* entre las tasas de incidencia, es del 30%. En la comunidad B, la incidencia en las personas expuestas es del 90% y la incidencia en las personas no expuestas, del 60%. Si se calcula el *cociente* de la incidencia entre las personas expuestas y las personas no expuestas en la población B, es de 90/60, o 1,5. Si se calcula la *diferencia* entre la incidencia en las personas expuestas y no expuestas en la comunidad B, es, de nuevo, del 30%.

¿Qué indican estas dos medidas? ¿Hay una diferencia sobre la información que aporta el cociente de las tasas de incidencia en comparación con la diferencia entre las tasas de incidencia? Esta pregunta es el tema de este capítulo y del capítulo 13.

Riesgo relativo

Concepto de riesgo relativo

Tanto los estudios de casos y controles como los estudios de cohortes están diseñados para determinar si existe una asociación entre la exposición a un factor y el desarrollo de una enfermedad. Si existe una asociación, hay que averiguar su grado de solidez. Si se lleva a cabo un estudio de cohortes, se puede formular la pregunta de otra manera: «¿Cuál es el cociente del riesgo de enfermedad en los individuos expuestos y el riesgo de enfermedad en los individuos no expuestos?» Este cociente es el *riesgo relativo* (RR):

$$\text{Riesgo relativo} = \frac{\text{Riesgo en expuestos}}{\text{Riesgo en no expuestos}}$$

El riesgo relativo también se puede definir como la probabilidad de que un evento (el desarrollo de una enfermedad) se produzca en personas expuestas en comparación con la probabilidad del evento en personas no expuestas, o como el cociente de estas dos probabilidades.

Interpretación del riesgo relativo

¿Cómo se interpreta el valor de un riesgo relativo?

1. Si el riesgo relativo es igual a 1, el numerador es igual al denominador: el riesgo en las personas expuestas es igual al riesgo en las no expuestas. Por tanto, no existe evidencia de ningún mayor riesgo en los individuos expuestos ni de ninguna asociación de la enfermedad con la exposición en cuestión.
2. Si el riesgo relativo es superior a 1, el numerador es mayor que el denominador: el riesgo en las personas expuestas es mayor que el riesgo en las no expuestas. Esto indica una asociación positiva, que puede ser causal (como se explica en un capítulo posterior).
3. Si el riesgo relativo es inferior a 1, el numerador es menor que el denominador: el riesgo en las personas expuestas es menor que el riesgo en las no expuestas. Esto indica una asociación negativa, y puede apuntar hacia un efecto protector. Este hallazgo se puede observar en las personas que reciben una vacuna eficaz («expuestas» a la vacuna).

Estas tres posibilidades se resumen en la tabla 12.4.

Tabla 12.4

Interpretación del riesgo relativo (RR) de una enfermedad

Si RR = 1	El riesgo de los expuestos es igual al riesgo de los no expuestos (ausencia de asociación)
Si RR >1	El riesgo de los expuestos es mayor que el riesgo de los no expuestos (asociación positiva; posiblemente causal)
Si RR <1	El riesgo de los expuestos es menor que el riesgo de los no expuestos (asociación negativa; posiblemente de protección)

Cálculo del riesgo relativo en los estudios de cohortes

En un estudio de *cohortes*, el riesgo relativo puede calcularse *directamente*. Recuerde el diseño de un estudio de cohortes, que se expone en la tabla 12.5.

Tabla 12.5

Cálculos del riesgo en un estudio de cohortes

		DESPUÉS, SEGUIR PARA VER SI			
		La enfermedad se desarrolla	La enfermedad no se desarrolla	Totales	Incidencia de la enfermedad
Primero, seleccionar	Expuestos	a	b	$a+b$	$\frac{a}{a+b}$
	No expuestos	c	d	$c+d$	$\frac{c}{c+d}$
	$\frac{a}{a+b}$ = incidencia en expuestos		$\frac{c}{c+d}$ = incidencia en no expuestos		

En esta tabla se observa que la incidencia en los individuos expuestos es

$$\frac{a}{a+b}$$

y la incidencia en los individuos no expuestos es

$$\frac{c}{c+d}$$

El riesgo relativo se calcula de la siguiente manera:

$$\text{Riesgo relativo} = \frac{\textit{Incidencia en expuestos}}{\textit{Incidencia en no expuestos}} = \frac{\left(\dfrac{a}{a+b}\right)}{\left(\dfrac{c}{c+d}\right)}$$

En la tabla 12.6 se muestra un estudio de cohortes hipotético con 3.000 fumadores y 5.000 no fumadores para investigar la relación del tabaquismo con el desarrollo de arteriopatía coronaria (AC) en un periodo de 1 año.

Tabla 12.6

Tabaquismo y arteriopatía coronaria (AC): estudio de cohortes hipotético con 3.000 fumadores de cigarrillos y 5.000 no fumadores

	Se desarrolla AC	No se desarrolla AC	Totales	Incidencia por 1.000 por año
Fuman cigarrillos	84	2.916	3.000	28,0
No fuman cigarrillos	87	4.913	5.000	17,4

En este ejemplo:

$$\text{Incidencia en los expuestos} = \frac{84}{3.000} = 28,0 \text{ por } 1.000$$

e

$$\text{Incidencia en los no expuestos} = \frac{87}{5.000} = 17,4 \text{ por } 1.000$$

Por tanto:

$$\text{Riesgo relativo} = \frac{\text{Incidencia en expuestos}}{\text{Incidencia en no expuestos}} = \frac{28,0}{17,4} = 1,61$$

Una expresión similar de riesgos se observa en un ejemplo histórico en la tabla 12.7, donde se muestran los datos de los primeros 12 años del estudio Framingham que relacionan el riesgo de enfermedad coronaria con la edad, el sexo y la concentración de colesterol.

Tabla 12.7

Relación entre la concentración sérica de colesterol y el riesgo de arteriopatía coronaria por edad y sexo: estudio Framingham durante los primeros 12 años

	VARONES		MUJERES	
Colesterol sérico (mg/dl)	30-49 años	50-62 años	30-49 años	50-62 años
Incidencia (por 1.000)				
<190	38,2	105,7	11,1	155,2
190-219	44,1	187,5	9,1	88,9
220-249	95,0	201,1	24,3	96,3
250+	157,5	267,8	50,4	121,5
Riesgos relativos[a]				
<190	1,0	2,8	0,3	4,1
190-219	1,2	4,9	0,2	2,3
220-249	2,5	5,3	0,6	2,5
250+	4,1	7,0	1,3	3,2

[a] La incidencia para cada subgrupo se compara con la de los varones de 30-49 años de edad, con una concentración sérica de colesterol inferior a 190 mg/dl (riesgo = 1,0).

De Truett J, Cornfield J, Kannel W. A multivariate analysis of the risk of coronary heart disease in Framingham. *J Chronic Dis*. 1967;20:511–524.

En primer lugar, hay que prestar atención a la parte superior de la tabla, que muestra las tasas de incidencia por 1.000 habitantes de la población de Framingham, por edad, sexo y concentración sérica de colesterol. En los varones, la asociación del riesgo con la concentración de colesterol parece estar relacionada con la dosis; el riesgo aumenta para ambos grupos de edad con los incrementos de la concentración de colesterol. La relación no es tan constante en las mujeres.

En la mitad inferior de la tabla, los valores se han convertido en riesgos relativos. Los autores han tomado la tasa de incidencia de 38,2 en los varones más jóvenes con una baja concentración de colesterol y le han asignado un riesgo de 1,0; estos sujetos se consideran «no expuestos». Todos los demás riesgos de la tabla se expresan en relación con este riesgo de 1,0. Por ejemplo, la incidencia de 157,5 en varones más jóvenes con una concentración de colesterol superior a 250 mg/dl se compara con la tasa de incidencia de 38,2; al dividir 157,5 entre 38,2, se obtiene un riesgo relativo de 4,1. Utilizando estos riesgos relativos, es más fácil comparar los riesgos e identificar cualquier tendencia. Aunque el riesgo más bajo en los varones se ha escogido como el estándar y se ha establecido en 1,0, los autores podrían haber optado por fijar cualquiera de los valores de la tabla en 1,0 y establecer todos los demás respecto a este. Por lo tanto, al describir el riesgo relativo, se deben especificar las categorías de numerador y denominador (p.ej., en la tabla 12.7, el riesgo de los hombres más jóvenes con niveles de colesterol ≥250 mg/dl es aproximadamente cuatro veces mayor que el de los hombres más jóvenes con niveles de colesterol <190 mg/dl). Una de las razones para la elección de un valor bajo como el estándar es que la mayoría de los otros valores serán mayores de 1,0; para la mayoría de las personas, la tabla es más fácil de leer cuando hay menos valores completamente a la derecha del decimal.

En la figura 12.4 se muestran los datos basados en la fusión de 2.282 varones de mediana edad seguidos durante 10 años en el estudio Framingham y de 1.838 varones de mediana edad seguidos durante 8 años en Albany, Nueva York. Los datos relacionan el tabaquismo, la concentración de colesterol y la presión arterial con el riesgo de infarto de miocardio y fallecimiento por AC. Recuerde que, en ese momento de la historia, estas asociaciones aún no se conocían, aunque en la actualidad todos estamos de acuerdo en que son factores de riesgo establecidos para la enfermedad cardiaca. Los autores han asignado un valor de 1 al riesgo más bajo en cada una de las dos partes de la figura, y los otros riesgos se calculan respecto a este valor. A la izquierda se muestra el riesgo en los no fumadores con concentraciones bajas de colesterol (que se ha fijado en 1) y el riesgo en los no fumadores con concentraciones altas de colesterol; los riesgos para los fumadores con concentraciones bajas y altas de colesterol se calculan respecto a los riesgos para los no fumadores con concentraciones bajas de colesterol. Debe tenerse en cuenta que el riesgo es mayor con concentraciones altas de colesterol, y que esto es válido tanto en fumadores como en no fumadores (aunque el riesgo es mayor en los fumadores, incluso con concentraciones de colesterol bajas). Por consiguiente, tanto el tabaquismo como las concentraciones elevadas de colesterol contribuyen al riesgo de infarto de miocardio y de fallecimiento por AC. Un análisis comparable con la presión arterial y el tabaquismo se muestra en el lado derecho de la figura.

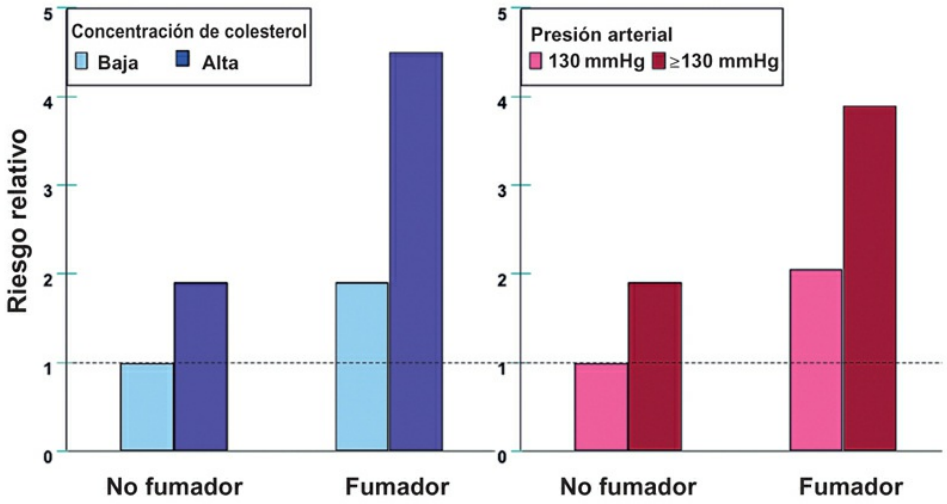


FIG. 12.4 Riesgo relativo de infarto de miocardio y fallecimiento por arteriopatía coronaria en varones de 30-62 años por concentración sérica de colesterol *(izquierda)* y niveles de presión arterial *(derecha)* en relación con el consumo de cigarrillos. Las concentraciones altas de colesterol se definen como 220 mg/dl o mayores. (Datos de Doyle JT, Dawber TR, Kannel WB, et al. The relationship of cigarette smoking to coronary heart disease. *JAMA*. 1964;190:886.)

Cociente de posibilidades (posibilidades relativas)

Ya hemos visto que para calcular un riesgo relativo hay que tener los valores de la incidencia de la enfermedad en las personas expuestas y en las no expuestas, que pueden obtenerse a partir de un estudio de cohortes. En un estudio *de casos y controles,* sin embargo, no se sabe la incidencia en la población expuesta o la incidencia en la población no expuesta, porque se comienza con personas enfermas (casos) y personas no enfermas (controles). Por tanto, en un estudio de casos y controles *no se puede* calcular el riesgo relativo directamente. En esta sección veremos cómo se puede obtener otra medida de asociación, el *cociente (razón) de posibilidades* u *odds ratio (OR),* bien a partir de una cohorte o bien a partir de un estudio de casos y controles, y se puede utilizar en lugar del riesgo relativo. También veremos que, pese a que no se puede calcular un riesgo relativo a partir de un estudio de casos y controles, bajo muchas condiciones, es posible obtener una *estimación* muy buena del riesgo relativo a partir de un estudio de casos y controles utilizando el OR.

Definición del OR en estudios de cohortes y de casos y controles

En los capítulos anteriores se ha comentado la *proporción* de la población expuesta en la que se desarrolla la enfermedad y la *proporción* de la población no expuesta en la que se desarrolla la enfermedad en un estudio de cohortes. De forma similar, en los estudios de casos y controles se ha comentado la *proporción* de los casos que fueron expuestos y la *proporción* de los controles que fueron expuestos (tabla 12.8).

Tabla 12.8

Cálculo de las proporciones expuestas en un estudio de casos y controles

		Primero, seleccionar	
		Casos (con enfermedad)	***Controles (sin enfermedad)***
Después, medir la exposición previa	Estuvieron expuestos	a	b
	No estuvieron expuestos	c	d
	Totales	$a+c$	$b+d$
	Proporciones de expuestos	$\frac{a}{a+c}$	$\frac{b}{b+d}$

Una estrategia alternativa consiste en utilizar el concepto de *posibilidades (odds).* Al apostar en el hipódromo, normalmente se consideran las probabilidades de ganar de cada caballo, el llamado sistema de desventaja. Supongamos que estamos apostando a un caballo llamado Epi Beauty, que tiene un 60% de probabilidad de ganar la carrera *(P).* Por tanto, Epi Beauty tiene un 40% de probabilidad de perder (1 – *P*). Si estas son las probabilidades, ¿cuáles son las *posibilidades* de que el caballo gane la carrera? Para responder a esto, debemos tener en cuenta que *las posibilidades de un evento pueden definirse como el cociente del número de formas en las que el evento puede ocurrir entre el número de formas en las que el evento no puede ocurrir.* Por tanto, las posibilidades de que Epi Beauty gane, como se ha definido anteriormente, son las siguientes:

$$\textit{Posibilidades} = \frac{\text{Probabilidad de que Epi Beauty gane la carrera}}{\text{Probabilidad de que Epi Beauty pierda la carrera}}$$

Hay que recordar que, si *P* es la probabilidad de que Epi Beauty gane la carrera, 1 − *P* es la probabilidad de que Epi Beauty pierda la carrera. Por tanto, las posibilidades de que Epi Beauty gane son:

$$\text{Posibilidades} = \frac{P}{1-P} \quad \text{o} \quad \frac{60\%}{40\%} = 1{,}5{:}1 = 1{,}5$$

Es esencial tener en cuenta la distinción entre la probabilidad y las posibilidades. En el ejemplo anterior:

$$\text{Probabilidad de ganar} = 60\%$$

y

$$\text{Posibilidades de ganar } (Odds) = \frac{60\%}{40\%} = 1{,}5$$

OR en los estudios de cohortes

A continuación, se analizará cómo el concepto de posibilidad se puede aplicar tanto a los estudios de cohortes como a los estudios de casos y controles. Consideremos en primer lugar el diseño del estudio de cohortes que se muestra en la figura 12.5A. La primera pregunta que nos planteamos es: «¿Cuál es la *probabilidad (P)* de que la enfermedad se desarrolle en una persona expuesta?» La respuesta a esto es la incidencia de la enfermedad de la fila superior (personas expuestas), que equivale a $\left(\frac{a}{a+b}\right)$. A continuación, nos preguntamos: «¿Cuáles son *las posibilidades* de que la enfermedad se desarrolle en una persona expuesta?» De nuevo, basta con mirar la fila superior de la figura 12.5 para ver que hay ($a + b$) personas expuestas; las posibilidades de que la enfermedad se desarrolle en ellas son *a:b* o $\frac{a}{b}$. (Recordemos $\left(\frac{P}{1-P}\right)$.del ejemplo de Epi Beauty.) Del mismo modo, basta con mirar la fila inferior de esta tabla para ver que hay ($c + d$) personas no expuestas; la probabilidad de que la enfermedad se desarrolle en personas no expuestas es $\left(\frac{c}{c+d}\right)$ y las posibilidades de que la enfermedad se desarrolle en estas personas no expuestas son *c:d* o $\frac{c}{d}$.

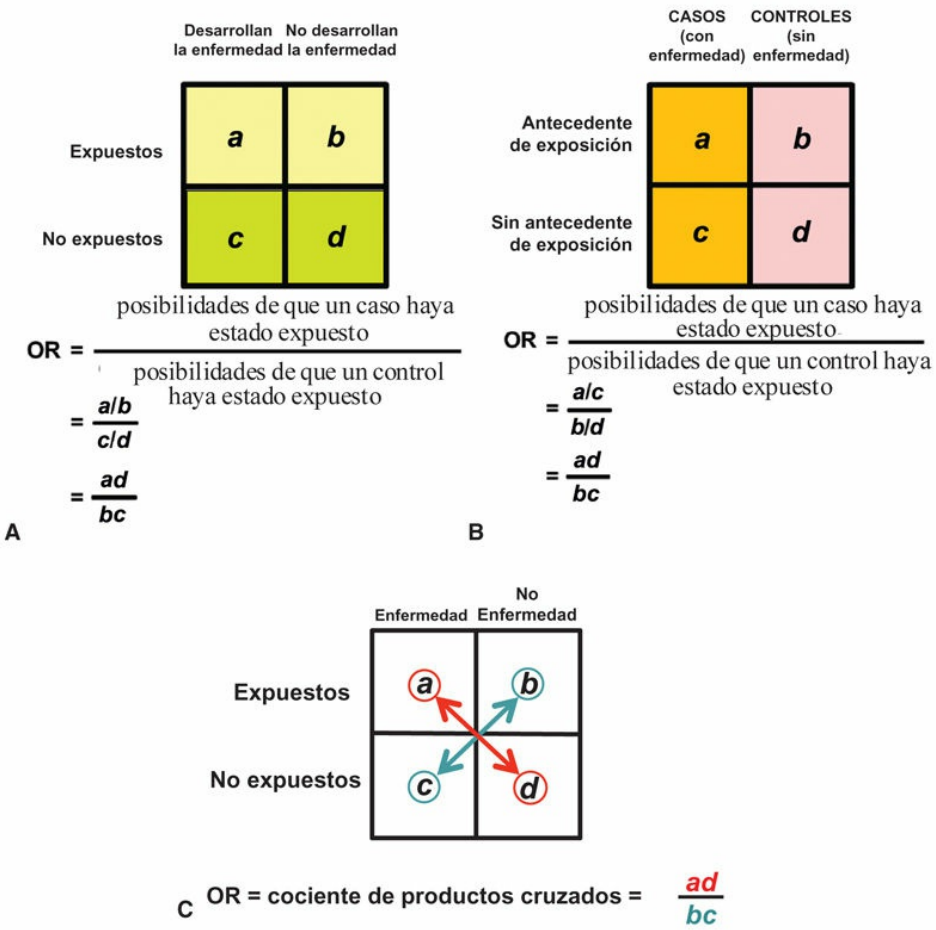


FIG. 12.5 (A) OR en un estudio de cohortes. (B) OR en un estudio de casos y controles. (C) Cociente de productos cruzados tanto en un estudio de cohortes como en uno de casos y controles.

Del mismo modo que el cociente de la incidencia en los expuestos entre la incidencia en los no expuestos se puede utilizar para medir una asociación entre la exposición y la enfermedad, también se puede considerar el cociente de las posibilidades de que se desarrolle la enfermedad en una persona expuesta entre las posibilidades de que se desarrolle en una persona no expuesta. Cualquiera de estas medidas de asociación es válida en un estudio de cohortes.

En un estudio de cohortes, para responder a la pregunta de si existe una asociación entre la exposición y la enfermedad, se puede utilizar el riesgo relativo comentado en la sección anterior o se puede emplear el OR (también denominado *posibilidades relativas).* En un *estudio de cohortes,* el OR se define como el *cociente de las posibilidades de que se desarrolle la enfermedad en las personas expuestas entre las posibilidades de que se desarrolle la enfermedad en las personas no expuestas,* y se puede calcular de la siguiente manera:

$$\frac{\left(\frac{a}{b}\right)}{\left(\frac{c}{d}\right)} = \frac{ad}{bc}$$

OR en un estudio de casos y controles

Como se acaba de exponer, en un *estudio de casos y controles* no se puede calcular el riesgo relativo directamente para determinar si existe una asociación entre la exposición y la enfermedad. Esto se debe a que, al haber comenzado con los casos y los controles en vez de con las personas expuestas y no expuestas, no se dispone de información sobre la incidencia de la enfermedad en las personas expuestas frente a las no expuestas. Sin embargo, aunque se puede utilizar el OR como una medida de la asociación entre la exposición y la enfermedad en un estudio de casos y controles, las preguntas que se plantearán serán algo diferentes: «¿Cuáles son las posibilidades de que un caso se haya expuesto?» Si miramos la columna de la izquierda de la figura 12.5B, veremos que las *posibilidades* de que un caso se haya expuesto son *a:c* o $\frac{a}{c}$. A continuación, nos preguntamos: «¿Cuáles son las posibilidades de que un control se haya expuesto?» Si miramos la columna de la derecha, veremos que las *posibilidades* de que un control se haya expuesto son *b:d* o $\frac{b}{d}$.

A continuación, se puede calcular el OR, que en un estudio de casos y controles se define como el cociente de las posibilidades de que los casos estuviesen expuestos entre las posibilidades de que los controles estuviesen expuestos. Esto se calcula del siguiente modo:

$$\frac{\left(\frac{a}{c}\right)}{\left(\frac{b}{d}\right)} = \frac{ad}{bc}$$

Por tanto, resulta interesante señalar que $\frac{ad}{bc}$ es el OR (o posibilidades

relativas), *tanto* en los estudios de cohortes (v. fig. 12.5A) *como* en los de casos y controles (v. fig. 12.5B). En ambos tipos de estudios, el OR es una medida excelente de si una determinada exposición se asocia con una enfermedad específica. El *OR* también se denomina *cociente de productos cruzados*, ya que se puede obtener multiplicando las celdas diagonales de una tabla de 2 × 2 y luego dividiendo ambos productos: $\frac{ad}{bc}$, como se observa en la figura 12.5C. Tenga en cuenta que, al calcular el cociente de posibilidades como una relación de productos cruzados, el formato de la tabla debe ser exactamente como el que se muestra en la tabla 12.5 (es decir, los casos en la columna del lado izquierdo y los individuos expuestos en la fila superior). De lo contrario obtendrá un cálculo incorrecto del cociente de posibilidades.

Tenga en cuenta también que, tanto en los estudios de cohortes como en los de casos y controles, el cociente de posibilidades se puede calcular como la relación de productos cruzados, lo que significa que el cociente de posibilidades de la exposición en un estudio de casos y controles es igual al cociente de posibilidades de la enfermedad. Por ejemplo, en un estudio de casos y controles de la enfermedad de Alzheimer, una patología bastante común en los ancianos, Tolppanen y cols.[1] considerando un intervalo de 5 años, encontraron que los casos tenían una probabilidad de lesión cerebral previa 1,23 veces más alta que los controles. Cuando se comunican estos resultados, sería correcto, aunque no muy útil, decir que las probabilidades de lesiones cerebrales previas (exposición) en los casos fueron 2,4 veces mayores que en los controles. Sin embargo, dado que la epidemiología siempre está interesada en la identificación de factores predictivos de la enfermedad y que el cociente de posibilidades de la exposición es el mismo que el cociente de posibilidades de la enfermedad, una descripción más útil de este resultado sería que las probabilidades de desarrollar la enfermedad de Alzheimer en individuos con una lesión cerebral (exposición) fue 1,23 veces mayor que en la población sin antecedentes de lesión cerebral. En otras palabras, incluso en los estudios de casos y controles, la interpretación del cociente de posibilidades es siempre prospectiva. (Tenga en cuenta que, dado que la enfermedad de Alzheimer es común en las personas de edad avanzada, el cociente de posibilidades en este ejemplo generalmente no se considera una buena estimación del riesgo relativo y, por lo tanto, el cociente de posibilidades debe describirse en términos de probabilidades, no de riesgos. Ver también la sección siguiente, titulada «¿Cuándo es el cociente de posibilidades (OR) una buena estimación del riesgo relativo?»)

El OR, o relación de productos cruzados, se puede considerar como el cociente del producto de las dos celdas que respaldan la hipótesis de una asociación (celdas *a* y *d*, correspondientes a las personas enfermas que estuvieron expuestas y a las personas no enfermas que no estuvieron expuestas, respectivamente) entre el producto de las dos celdas que niegan la hipótesis de una asociación (celdas *b* y *c*, correspondientes a las personas no

enfermas que estuvieron expuestas y a las personas enfermas que no estuvieron expuestas, respectivamente).

Interpretación del OR

Nosotros interpretamos el OR igual que el riesgo relativo. Si la exposición no está relacionada con la enfermedad, el OR será igual a 1. Si la exposición se relaciona positivamente con la enfermedad, dicho cociente será superior a 1, y si la exposición se relaciona negativamente con la enfermedad (es decir, si es protectora), será inferior a 1.

¿Cuándo es el OR una buena estimación del riesgo relativo?

En un estudio de casos y controles, solo se puede calcular el OR como medida de asociación, mientras que en un estudio de cohortes, tanto el riesgo relativo como el OR son una medida válida de la asociación. Sin embargo, muchas personas se sienten más cómodas usando el riesgo relativo, y esta es la medida de asociación más utilizada en la literatura cuando se publican los estudios de cohortes. Incluso cuando se utiliza el OR, las personas suelen mostrar interés en saber en qué medida supone una aproximación del riesgo relativo. Incluso revistas clínicas prestigiosas publican artículos de casos y controles en los que una columna de resultados lleva el encabezamiento de *riesgos relativos.* Después de leer la exposición en este capítulo, se puede advertir el error de dicho encabezamiento, porque ahora sabemos que los riesgos relativos no pueden calcularse directamente a partir de un estudio de casos y controles. Es evidente que hacen referencia a una *estimación* de riesgos relativos basada en los cocientes de posibilidades que se obtienen en los estudios de casos y controles.

¿Cuándo el OR (posibilidades relativas) obtenido en un estudio de casos y controles es una buena aproximación del riesgo relativo en la población? Cuando se cumplen las tres condiciones siguientes:

1. Cuando los *casos* estudiados son representativos, en lo que respecta a la historia de la exposición, de todas las personas con la enfermedad en la población de la que se extrajeron los casos.
2. Cuando los *controles* estudiados son representativos, en lo que respecta a la historia de la exposición, de todas las personas sin la enfermedad en la población de la que se extrajeron los casos.
3. Cuando la enfermedad que se está estudiando no ocurre con frecuencia.

Esta tercera condición, también conocida como la «suposición de rareza», tiene un significado estadístico, más que de salud pública. Por ejemplo, no se puede decir que una enfermedad con una incidencia del 4% en los individuos

expuestos y del 2% en los no expuestos sea infrecuente desde el punto de vista de la población. Sin embargo, desde el punto de vista de la suposición de rareza, la enfermedad es rara y, por lo tanto, el riesgo relativo (2,0) y el cociente de posibilidades (2,04) son prácticamente iguales.

La tercera condición (que la aparición de la enfermedad no sea frecuente) se puede explicar intuitivamente del siguiente modo:

Recordemos que hay $(a + b)$ personas expuestas. Dado que la mayoría de las enfermedades a las que nos enfrentamos no son frecuentes, muy pocas personas de una población expuesta desarrollarán en realidad la enfermedad; por tanto, a tiene un valor muy pequeño en comparación con b, y se puede aproximar $(a + b)$ como b, o $(a + b) \cong b$. De forma similar, muy pocas personas no expuestas $(c + d)$ desarrollan la enfermedad, y se puede aproximar $(c + d)$ como d, o $(c + d) \cong d$. Por tanto, es posible calcular un riesgo relativo de la siguiente manera:

$$\frac{\left(\frac{a}{a+b}\right)}{\left(\frac{c}{c+d}\right)} \cong \frac{\left(\frac{a}{b}\right)}{\left(\frac{c}{d}\right)}$$

$$\frac{ad}{bc}$$

Al realizar este cálculo, se obtiene $\frac{ad}{bc}$, que es el OR. Para los lectores interesados, se proporciona una derivación más elegante y más sofisticada en el apéndice de este capítulo.

En las figuras 12.6 y 12.7 se ofrecen dos ejemplos de estudios de cohortes que demuestran cómo el OR proporciona una buena aproximación del riesgo relativo cuando la aparición de una enfermedad es poco frecuente, pero no cuando es frecuente. En la figura 12.6, la aparición de la enfermedad es poco frecuente y se observa que el riesgo relativo es 2. Si ahora se calcula el OR (productos cruzados), se observa que es 2,02, un valor muy aproximado.

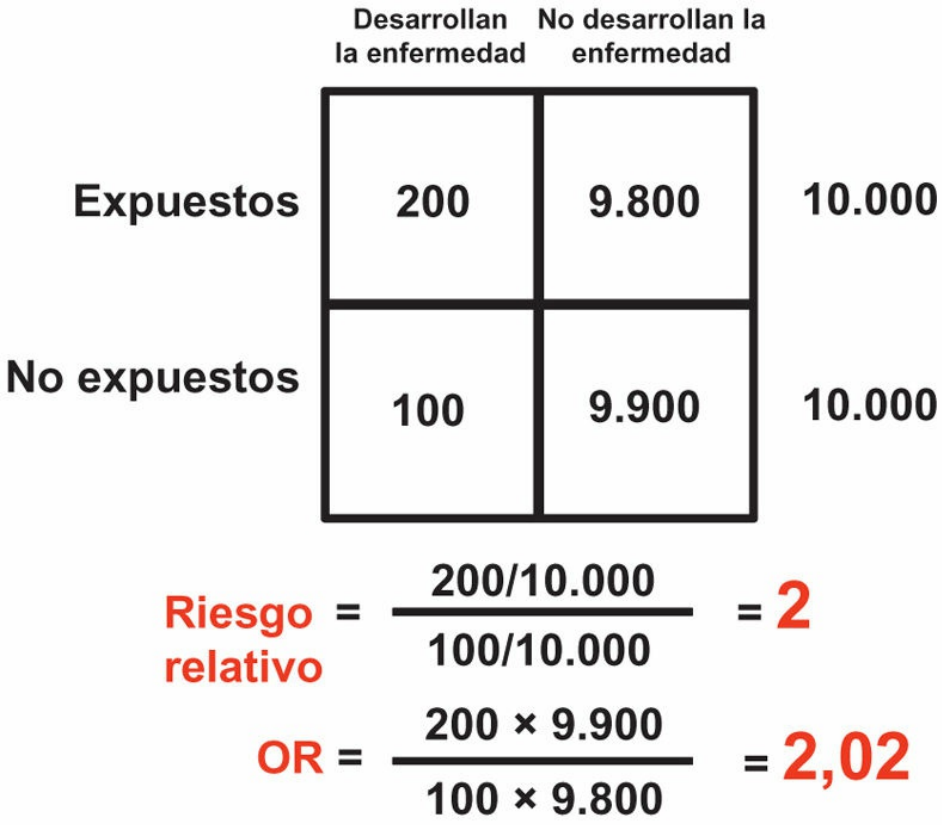


FIG. 12.6 Ejemplo: el OR es una buena estimación del riesgo relativo cuando una enfermedad es infrecuente.

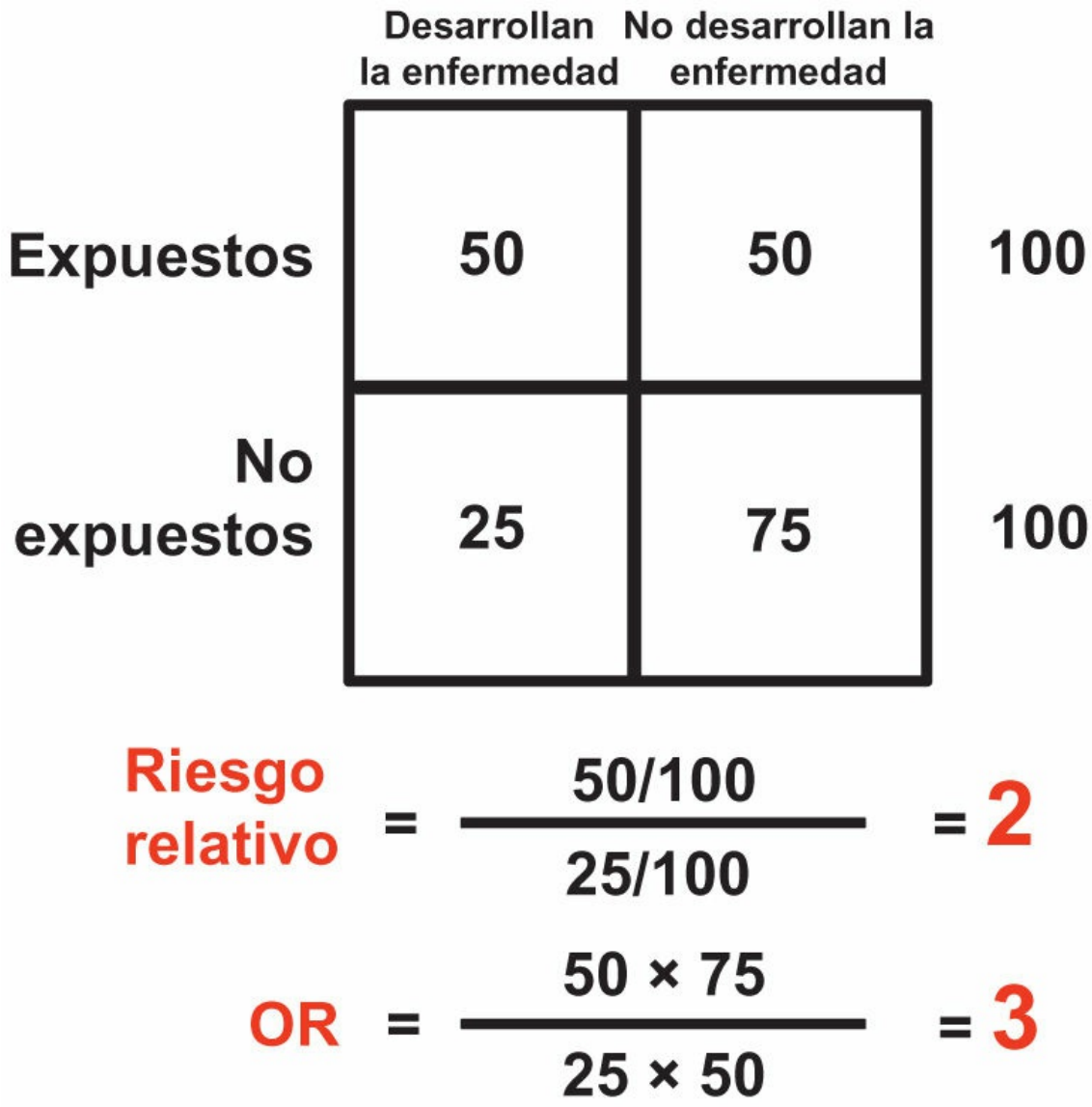


FIG. 12.7 Ejemplo: el OR *no* es una buena estimación del riesgo relativo cuando una enfermedad *no* es infrecuente.

Ahora, pasemos a la figura 12.7, en la que la aparición de la enfermedad es más frecuente. Aunque el riesgo relativo es de nuevo 2,0, el OR es 3,0, que es considerablemente diferente al riesgo relativo.

Se observa, por tanto, que el OR es en sí mismo una medida válida de la asociación sin considerar siquiera el riesgo relativo. Sin embargo, si se opta por utilizar el riesgo relativo como índice de asociación, cuando la aparición de la enfermedad es poco frecuente, el OR es una aproximación muy buena del riesgo relativo.

Hay que recordar que:

- El OR (posibilidades relativas) es una medida útil de asociación, por sí mismo, tanto en estudios de casos y controles como en estudios de cohortes.
- En un estudio de cohortes, el riesgo relativo puede calcularse directamente.

- En un estudio de casos y controles, el riesgo relativo no puede calcularse directamente, de manera que se utiliza el OR o posibilidades relativas (cociente de productos cruzados) como una estimación del riesgo relativo cuando el riesgo de la enfermedad es bajo.

El estudio de casos y controles realizado por Arvonen y cols.[2] es un ejemplo de cómo el cociente de posibilidades es una buena estimación del riesgo relativo. Estos autores encontraron un cociente de posibilidades de 2,4 para la asociación entre la alergia a la leche de vaca en la infancia (exposición) y el desarrollo de artritis idiopática juvenil. Debido a que la incidencia de esta enfermedad es muy baja (aproximadamente 13,9 por 100.000)[3], una interpretación precisa del cociente de posibilidades sería que el *riesgo* de la enfermedad en los individuos expuestos es más de dos veces mayor que en los no expuestos.

Ejemplos de cálculo del OR en estudios de casos y controles

En esta sección se calcularán los cocientes de posibilidades en dos estudios de casos y controles (uno en el que los controles *no* se han emparejado con los casos y otro en el que *sí* se han emparejado). A los efectos de estos ejemplos, se supondrá lo siguiente: el presupuesto de investigación es muy pequeño, por lo que se ha llevado a cabo un estudio de casos y controles de solo 10 casos y 10 controles. Es posible que se haya reconocido esta enfermedad recientemente y que haya muy pocos casos identificados hasta la fecha; sin embargo, ¡la enfermedad en sí es muy letal! *N* indica un individuo *no expuesto* y *E*, un individuo *expuesto*.

Cálculo del OR en un estudio de casos y controles sin emparejamiento

Supongamos que este estudio de casos y controles se realiza sin ningún tipo de emparejamiento de los controles con los casos y que se obtienen los resultados observados en la figura 12.8. Por tanto, 6 de los 10 casos estuvieron expuestos y 3 de los 10 controles estuvieron expuestos. Si se ordenan estos datos en una tabla de 2 × 2, se obtiene lo siguiente:

	Casos	**Controles**
Expuestos	6	3
No expuestos	4	7
Totales	10	10



FIG. 12.8 Estudio de casos y controles de 10 casos y 10 controles sin emparejar.

El OR en este estudio *sin emparejamiento* es igual al cociente de los productos cruzados:

$$\mathrm{OR} = \frac{\mathrm{ad}}{\mathrm{bc}}$$

$$OR = \frac{6 \times 7}{4 \times 3} = \frac{42}{12} = 3{,}5$$

En la tabla 12.9 se muestran los datos de un hipotético estudio de casos y controles sobre el tabaquismo y la AC. Las letras *a, b, c* y *d* se han insertado para identificar las celdas de la tabla de 2 × 2 que se utiliza para el cálculo. El OR, calculado a partir de estos datos, es el siguiente:

$$OR = \frac{\mathrm{ad}}{\mathrm{bc}} = \frac{112 \times 224}{176 \times 88} = 1{,}62$$

Tabla 12.9

Ejemplo del cálculo del OR en un estudio de casos y controles

		Primero, seleccionar	
		Casos de AC	***Controles***
Después, medir la exposición previa	Fumadores	112 *(a)*	176 *(b)*
	No fumadores	88 *(c)*	224 *(d)*
	Totales	200 $(a + c)$	400 $(b + d)$
	Proporciones de fumadores de cigarrillos	56%	44%

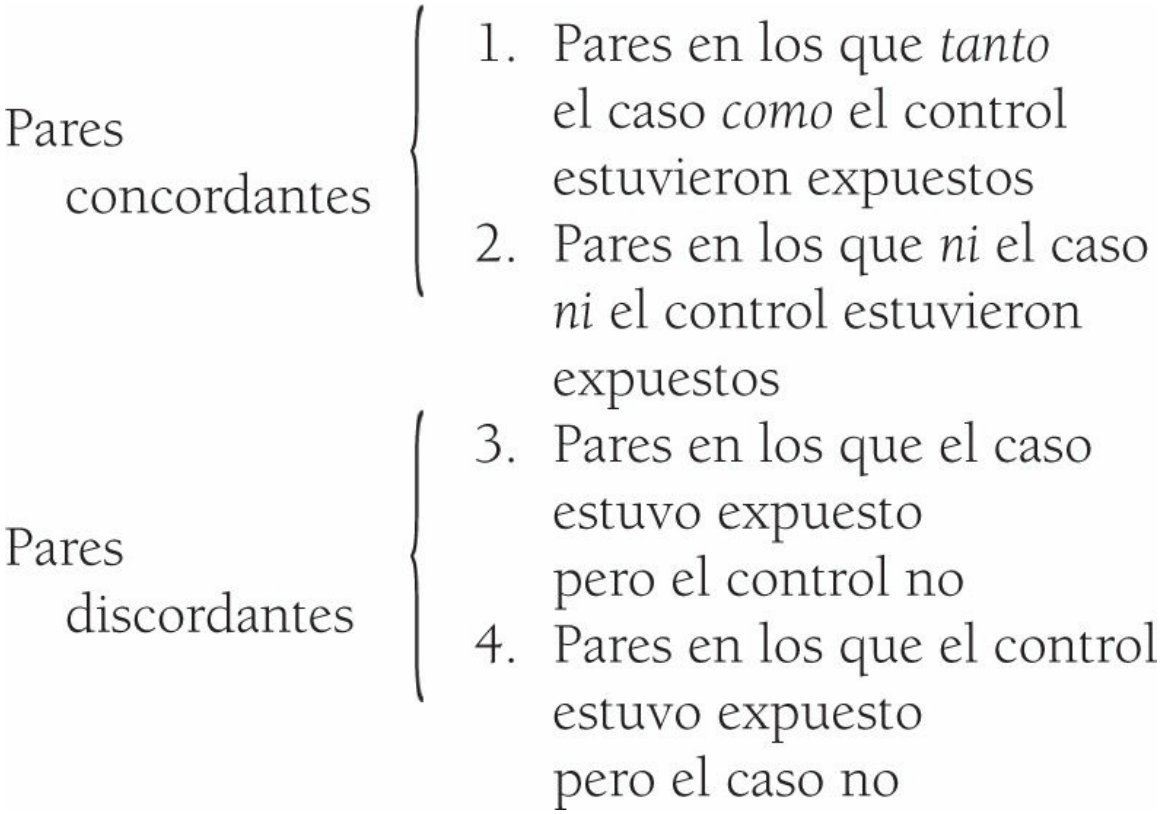
$$OR = \frac{ad}{bc} = \frac{112 \times 224}{176 \times 88} = 1{,}62$$

AC, arteriopatía coronaria.

Cálculo del OR en un estudio de casos y controles con datos emparejados

Como se comentó anteriormente, al seleccionar la población del estudio en los estudios de casos y controles, los controles se suelen seleccionar emparejando cada uno con un caso en función de las variables que se sabe que están relacionadas con el riesgo de padecer la enfermedad, como el sexo, la edad o la raza (empleando emparejamiento individual o datos emparejados). A continuación, los resultados se analizan en términos de pares de casos y controles en lugar de en función de los sujetos individuales.

¿Qué tipos de combinaciones de casos y controles son posibles en lo que respecta a la historia de la exposición? Es evidente que, si la exposición es dicotómica (una persona está expuesta o no expuesta), solo son posibles los siguientes cuatro tipos de pares de casos y controles:



Hay que señalar que los pares de casos y controles que tuvieron la misma experiencia de exposición se denominan *pares concordantes* y aquellos con una experiencia diferente de exposición, *pares discordantes.* Estas posibilidades se muestran esquemáticamente en una tabla de 2 × 2. Hay que tener en cuenta que, a diferencia de otras tablas de 2 × 2 que se han presentado anteriormente, la cifra de cada celda representa pares de sujetos (es decir, *pares de casos y controles*), *no* sujetos individuales. Por tanto, la tabla contiene *a* pares, en los que tanto el caso como el control estuvieron expuestos; *b* pares, en los que el caso estuvo expuesto y el control no; *c* pares, en los que el caso no estuvo expuesto y el control sí; y *d* pares, en los que ni el caso ni el control estuvieron expuestos.

		Control	
		Expuesto	*No expuesto*
Caso	*Expuesto*	*a*	*b*
	No expuesto	*c*	*d*

El cálculo del OR en un estudio de datos emparejados de este tipo se basa solo en los *pares discordantes* (*b* y *c*). Los pares concordantes (*a* y *d*, en los que los casos y los controles estuvieron ambos expuestos o no expuestos) se ignoran, ya que no contribuyen a nuestro conocimiento de la diferencia entre

los casos y los controles en lo que respecta a la historia previa de exposición.

Por tanto, el OR para datos emparejados es el cociente de los pares discordantes (es decir, el cociente del número de pares en los que el caso estuvo expuesto y el control no entre el número de pares en los que el control estuvo expuesto y el caso no). El OR para la tabla 2 × 2 precedente es el siguiente:

$$\text{OR de datos emparejados} = \frac{b}{c}$$

De nuevo, el OR de datos emparejados puede considerarse como el cociente del número de pares que respaldan la hipótesis de una asociación (pares en los que el caso estuvo expuesto y el control no) entre el número de pares que niegan la hipótesis de una asociación (pares en los que el control estuvo expuesto y el caso no).

Veamos ahora un ejemplo de un cálculo del OR en un estudio de casos y controles con datos emparejados (fig. 12.9). Volvamos a nuestro estudio de bajo presupuesto, que solo incluyó 10 casos y 10 controles: ahora nuestro estudio está diseñado para que cada control se haya emparejado de forma individual con un caso, lo que da lugar a 10 *pares* de casos y controles (las flechas horizontales indican el emparejamiento de los datos). Si se utilizan estos datos para construir una tabla de 2 × 2 para los *pares,* se obtiene lo siguiente:

		Control	
		Expuesto	***No expuesto***
Caso	***Expuesto***	2	4
	No expuesto	1	3



FIG. 12.9 Estudio de casos y controles de 10 casos y 10 controles emparejados.

Hay que tener en cuenta que hay dos pares en los que se expusieron *tanto* el caso *como* el control y tres pares en los que se no se expusieron *ni* el caso *ni* el control. Estos pares concordantes se ignoran en el análisis de datos emparejados.

Hay cuatro pares en los que el caso estuvo expuesto y el control no y un par en el que el control estuvo expuesto y el caso no.

Por tanto, el OR para datos emparejados es el siguiente:

$$\text{OR para datos emparejados} = \frac{b}{c} = \frac{4}{1} = 4$$

En las figuras 12.10 y 12.11 se presentan datos seleccionados del estudio de casos y controles de tumores cerebrales infantiles que se expuso en el capítulo anterior. Se muestran los datos para dos variables. En la figura 12.10 se presenta un análisis de datos emparejados para el peso al nacer. Varios estudios han sugerido que los niños con mayor peso al nacer tienen un riesgo mayor de ciertos cánceres infantiles. En este análisis, la *exposición* se define como un peso al nacer mayor de 3,6 kg. El resultado es un OR de 2,57.

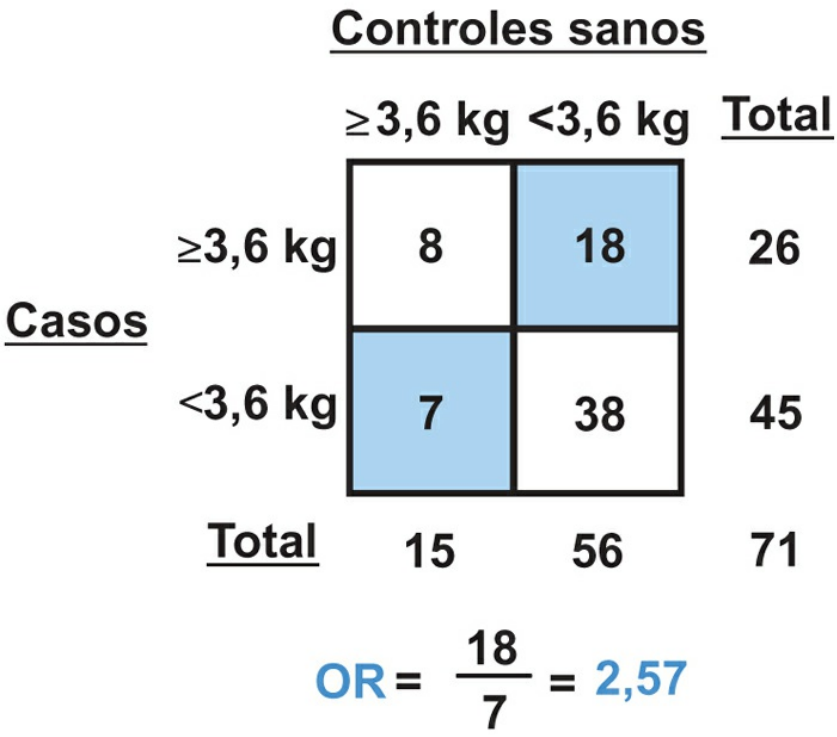


FIG. 12.10 Peso al nacer de un niño índice: comparación de datos emparejados entre casos y controles sanos (≥3,6 kg frente a <3,6 kg). (Datos de Gold E, Gordis L, Tonascia J, Szklo M. Risk factors for brain tumors in children. *Am J Epidemiol*. 1979;109:309–319.)

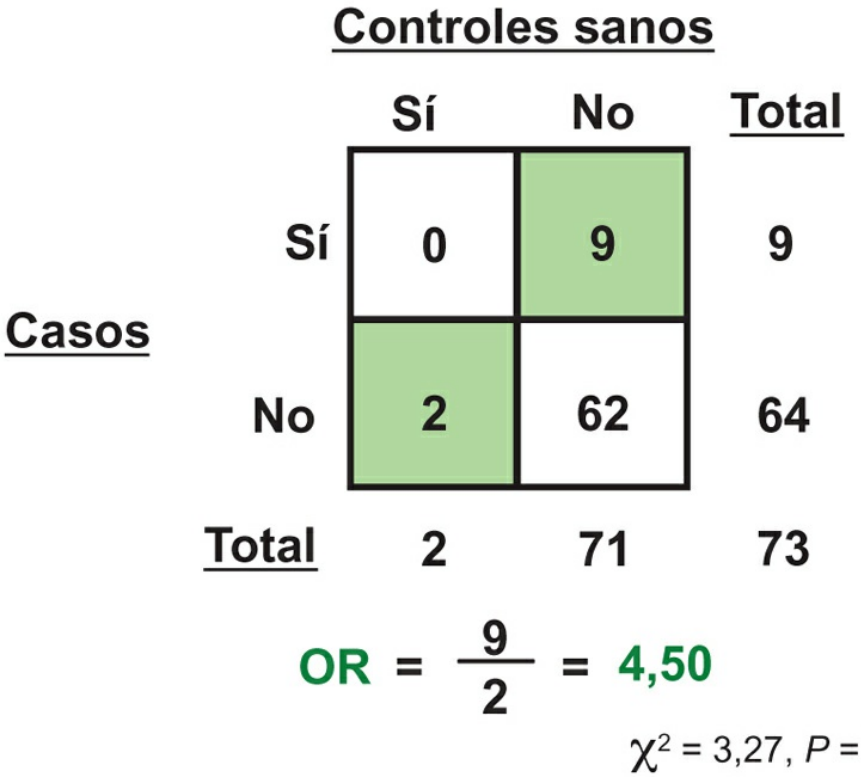


FIG. 12.11 Exposición de un niño índice a mascotas enfermas: comparación de

datos emparejados entre casos y controles sanos. (Datos de Gold E, Gordis L, Tonascia J, Szklo M. Risk factors for brain tumors in children. *Am J Epidemiol*. 1979;109:309–319.)

En la figura 12.11 se presenta un análisis de datos emparejados para la exposición a mascotas enfermas. Hace muchos años, el estudio Tri-State Leukemia demostró que más casos de leucemia tenían mascotas en casa en comparación con los controles. El interés reciente por los virus oncogénicos ha estimulado a su vez un interés por la exposición a mascotas enfermas como una posible fuente de tales agentes. Gold y cols. evaluaron esta cuestión en su estudio de casos y controles[1], cuyos resultados se muestran en la figura 12.11. Aunque el OR fue de 4,5, el número de pares discordantes fue muy pequeño.

Aunque los ejemplos mencionados anteriormente son un poco antiguos, la forma de realizar el cálculo del RR y del OR es la misma en los artículos de investigación más actuales. Sin embargo, la presentación actual de los resultados se centra más en el ajuste respecto a las posibles variables de confusión.

Conclusión

En este capítulo se han expuesto los conceptos de riesgo absoluto, riesgo relativo y OR. En el capítulo 13 se prestará atención a otro aspecto importante del riesgo: el riesgo atribuible. A continuación, se revisarán los diseños de los estudios y los índices de riesgo que se han descrito antes de abordar el uso de estos conceptos a la hora de deducir inferencias causales.

Bibliografía

1. Tolppanen AM, Taipale H, Hartikainen S. Head or brain injuries and Alzheimer's disease: a nested case-control regiser study. *Alzheimers Dement*. 2017;13:1371–1379.
2. Arvonen M, Virta LJ, Pokka T, et al. Cow's milk allergy in infancy and later development of juvenile idiopathic arthritis: a register-based case-control study. *Am J Epidemiol*. 2017;186:237–244.
3. Towner SR, Michet Jr CJ, O'Fallon WM, et al. The epidemiology of juvenile arthritis in Rochester, Minnesota 1960-1979. *Arthritis Rheum*. 1983;26:1208–1213.
4. Gold E, Gordis L, Tonascia J, et al. Risk factors for brain tumors in children. *Am J Epidemiol*. 1979;109:309–319.

Preguntas de repaso del capítulo 12

1. *De 2.872 personas que habían recibido radioterapia en la infancia debido a una hiperplasia del timo, se desarrolló un cáncer de tiroides en 24 y un tumor tiroideo benigno en 52. El grupo de comparación consistió en 5.055 niños que no habían recibido este tratamiento (hermanos y hermanas de los niños que sí habían recibido radioterapia). Durante el periodo de seguimiento, ninguno de los niños del grupo de comparación desarrolló cáncer de tiroides, pero en seis de ellos se desarrolló un tumor tiroideo benigno. Calcule el riesgo relativo para los tumores tiroideos benignos:*

_____

Las preguntas 2 y 3 se basan en la información de la siguiente tabla.

En un pequeño estudio piloto se contactó con 12 mujeres que tenían cáncer de útero y con 12 sin enfermedad aparente y se les preguntó si alguna vez habían tomado estrógenos. Cada mujer con cáncer se emparejó en función de la edad, la raza, el peso y la paridad con una mujer sin la enfermedad. Los resultados se muestran a continuación:

Par n.°	Mujer con cáncer de útero	Mujer sin cáncer de útero
1	Ha tomado estrógenos	No ha tomado estrógenos
2	No ha tomado estrógenos	No ha tomado estrógenos
3	Ha tomado estrógenos	Ha tomado estrógenos
4	Ha tomado estrógenos	Ha tomado estrógenos

5	Ha tomado estrógenos	No ha tomado estrógenos
6	No ha tomado estrógenos	No ha tomado estrógenos
7	Ha tomado estrógenos	No ha tomado estrógenos
8	Ha tomado estrógenos	No ha tomado estrógenos
9	No ha tomado estrógenos	Ha tomado estrógenos
10	No ha tomado estrógenos	Ha tomado estrógenos
11	Ha tomado estrógenos	No ha tomado estrógenos
12	Ha tomado estrógenos	No ha tomado estrógenos

2. ¿Cuál es el riesgo relativo estimado de cáncer cuando se analiza este estudio como un estudio de datos emparejados?

a. 0,25.

b. 0,33.

c. 1,00.

d. 3,00.

e. 4,20.

3. *A continuación, los datos se analizan sin emparejar. ¿Cuál es el riesgo relativo* estimado *de cáncer en el análisis de este estudio según un diseño de estudio* sin emparejamiento*?*

a. 0,70.

b. 1,43.

c. 2,80.

d. 3,00.

e. 4,00.

4. *En un estudio de una enfermedad en la que se evaluaron todos los casos que se desarrollaron, si el riesgo relativo para la asociación entre un factor y la enfermedad es igual o inferior a 1,0, entonces:*

a. No hay asociación entre el factor y la enfermedad.

b. El factor protege contra el desarrollo de la enfermedad.

c. El emparejamiento o la asignación aleatoria no se han realizado de forma satisfactoria.

d. El grupo de comparación utilizado no era adecuado, y no es posible establecer una comparación válida.

e. La asociación entre el factor y la enfermedad es inexistente o negativa.

Las preguntas 5 a 7 se basan en la siguiente información.

Talbot y cols. llevaron a cabo un estudio sobre la muerte súbita

inesperada en mujeres. Los datos sobre los antecedentes de tabaquismo se muestran en la siguiente tabla.

Antecedentes de tabaquismo para los casos de muerte súbita por cardiopatía aterosclerótica y controles (fumador actual, ≥1 paquete/día) [datos emparejados], condado de Allegheny, 1980

	Controles		
Casos	Fumador ≥1 paquete/día	Fumador <1 paquete/día	Totales
Fumador ≥1 paquete/día	2	36	38
Fumador <1 paquete/día	8	34	42
Totales	10	70	80

Modificada de Talbott E, Kuller LH, Perper J, Murphy PA. Sudden unexpected death in women: biologic and psychosocial origins. *Am J Epidemiol*. 1981;114:671–682.

5. *Calcular el OR de datos emparejados para estos datos:* _____

6. *Utilizando los datos de la tabla, desemparejar los pares y calcular el OR sin emparejamiento:* _____

7. *¿Cuáles son las posibilidades de que los controles fumen ≥1 paquete/día?* _____

Las preguntas 8 y 9 se basan en la información que aparece en la siguiente tabla.

Tasas de cardiopatía aterosclerótica (CPA) por 10.000 habitantes, por edad y sexo, Framingham, Massachusetts

	Varones		Mujeres	
Edad al comienzo del estudio (años)	Tasas de CPA en la exploración inicial	Exploraciones anuales de seguimiento (incidencia anual media)	Tasas de CPA en la exploración inicial	Exploraciones anuales de seguimiento (incidencia anual media)
29-34	76,7	19,4	0,0	0,0
35-44	90,7	40,0	17,2	2,1
45-54	167,6	106,5	111,1	29,4
55-62	505,4	209,1	211,1	117,8

8. *El riesgo relativo de desarrollar CPA después de entrar en este estudio* en los varones en comparación con las mujeres *es:*

a. Aproximadamente igual en todos los grupos de edad.

b. Máximo en el grupo de mayor edad.

c. Mínimo en los grupos de edad más joven y de mayor edad, y máximo en las edades de 35-44 y 45-54 años.

d. Máximo en los grupos de edad más joven y de mayor edad, y mínimo en las edades de 35-44 y 45-54 años.

e. Mínimo en el grupo de mayor edad.

9. *La explicación más probable de las diferencias en las tasas de CPA entre la exploración* inicial *y las exploraciones* de seguimiento *anuales en los varones es:*

a. La prevalencia y la incidencia de CPA aumentan con la edad en los varones.

b. La letalidad de CPA es mayor a edades más tempranas en los varones.

c. Un efecto de cohorte clásico explica estos resultados.

d. La letalidad en la CPA es máxima en las primeras 24 horas después de un ataque cardiaco.

e. En la exploración inicial se mide la prevalencia de CPA, mientras que en las exploraciones posteriores se mide principalmente la incidencia de CPA.

La pregunta 10 se basa en la siguiente información:

Se realiza un estudio de casos y controles emparejados para explorar la relación entre la proteína C reactiva (PCR) y la progresión a SIDA en Sudáfrica. Cada caso es emparejado con un control; se seleccionan de distintas clínicas de atención del VIH de los Estados Unidos. Los casos y los controles son emparejados por edad y sexo,

que son factores de confusión conocidos de la relación entre la PCR y la progresión al SIDA. La exposición fue un alto valor de PCR, definido como ≥2 mg/l. La siguiente distribución de exposición se observó entre 145 pares de casos y controles.

		Controles	
		+	-
Casos	+	25	31
	−	16	73

10. *Calcule el cociente de posibilidades (OR) del estudio de casos y controles.*

OR = _____

Las preguntas 11 y 12 se refieren a un estudio de casos y controles que explora la relación entre el uso de aspirina y las probabilidades de cáncer de próstata en un estudio con 375 casos y 407 controles.

Uso de aspirina	Casos	Controles
Nunca	198	174
Alguna vez	177	233

11. *¿Cuál es el cociente de posibilidades para el cáncer de próstata comparando el uso ocasional de aspirina con no usarla nunca? (El uso ocasional es la categoría de referencia).* _____

12. *Interprete en una oración el cociente de posibilidades que obtuvo en la pregunta 11.*

13. *¿Cuál de los siguientes no se puede estimar por lo general en un estudio de casos y controles?*

a. Riesgo relativo.

b. Tasa de incidencia.

c. Prevalencia de exposición en la población de origen.

d. Cociente de posibilidades de la enfermedad.

e. Cociente de posibilidades de la exposición.

Apéndice del capítulo 12

La derivación de la relación del OR y del riesgo relativo puede demostrarse por la siguiente fórmula. Debe recordarse que:

$$\text{Riesgo relativo}(RR) = \frac{\left(\frac{a}{a+b}\right)}{\left(\frac{c}{c+d}\right)}$$

$$OR = \frac{ad}{bc}$$

Así, la relación del riesgo relativo con el OR puede expresarse como el cociente de RR entre OR:

(1) FÓRMULA 1:

$$\frac{RR}{OR} = \frac{\left(\frac{a}{a+b}\right) \div \left(\frac{c}{c+d}\right)}{\left(\frac{ad}{bc}\right)}$$

$$= \frac{\left(\frac{a}{a+b}\right)}{\left(\frac{c}{c+d}\right)} \times \frac{bc}{ad}$$

$$= \frac{\left(\frac{abc}{a+b}\right)}{\left(\frac{cad}{c+d}\right)} = \frac{\left(\frac{b}{a+b}\right)}{\left(\frac{d}{c+d}\right)}$$

Dado que

$$\frac{b}{a+b} = \frac{a+b-a}{a+b} = \frac{a+b}{a+b} - \frac{a}{a+b} = 1 - \frac{a}{a+b}$$

y

$$\frac{d}{c+d} = \frac{c+d-c}{c+d} = \frac{c+d}{c+d} - \frac{c}{c+d} = 1 - \frac{c}{c+d}$$

la relación del riesgo relativo con el OR puede, por tanto, reducirse a la siguiente ecuación:

(2) FÓRMULA 2:

$$\frac{RR}{OR} = \frac{1-\left(\frac{a}{a+b}\right)}{1-\left(\frac{c}{c+d}\right)}$$

Si luego se multiplica la fórmula 2 por el OR:

(3) FÓRMULA 3:

$$RR = \frac{1-\left(\frac{a}{a+b}\right)}{1-\left(\frac{c}{c+d}\right)} \times OR$$

Si una enfermedad es rara, tanto $\frac{a}{a+b}$ como $\frac{c}{c+d}$ serán muy pequeños, por lo que los términos entre paréntesis en la fórmula 3 serán aproximadamente 1, y el OR se aproximará al riesgo relativo.

También resulta interesante analizar esta relación de una forma diferente. Si se recuerda la definición de *posibilidades* (es decir, el cociente del número de formas en que el evento puede ocurrir entre el número de formas en que el evento no puede ocurrir):

$$O = \frac{P}{1-P}$$

donde *O* representa las *posibilidades (odds)* de que la enfermedad se desarrolle y *P* es el *riesgo* de que la enfermedad se desarrolle. Debe tenerse en cuenta que, a medida que el valor de *P* se hace más pequeño, el denominador 1 − *P* se aproxima a 1, de modo que:

$$\frac{P}{1-P} \cong \frac{P}{1} = P$$

(es decir, las *posibilidades* se convierten en una buena aproximación del *riesgo*). Por tanto, si el riesgo es bajo (la enfermedad es rara), las *posibilidades* de que la enfermedad se desarrolle son una buena aproximación del *riesgo* de que se desarrolle.

Ahora, consideremos un grupo expuesto y un grupo no expuesto. Si el riesgo de una enfermedad es muy bajo, el *cociente* de las *posibilidades* en el grupo expuesto entre las *posibilidades* en el grupo no expuesto se aproxima estrechamente al *cociente* del *riesgo* en el grupo expuesto entre el *riesgo* en el grupo no expuesto *(el riesgo relativo):*

Es decir, cuando el valor de *P* es muy pequeño:

$$\frac{O_{exp}}{O_{noexp}} \cong \frac{P_{exp}}{P_{noexp}}$$

donde:

O_{exp} representa las posibilidades de desarrollar la enfermedad en la población expuesta,
O_{noexp} representa las posibilidades de desarrollar la enfermedad en la población no expuesta,
P_{exp} es la probabilidad (o riesgo) de desarrollar la enfermedad en la población expuesta, y
P_{noexp} es la probabilidad (o riesgo) de desarrollar la enfermedad en la población no expuesta.
Este cociente es el *OR* (posibilidades relativas).

CAPÍTULO 13

Conceptos adicionales sobre el riesgo: estimación de las posibilidades de prevención

Palabras clave

riesgo atribuible
riesgo atribuible poblacional
comparación de riesgo relativo y riesgo atribuible

Objetivos de aprendizaje

- Calcular e interpretar el riesgo atribuible para el grupo expuesto.
- Calcular e interpretar el riesgo atribuible poblacional.
- Describir cómo se utiliza el riesgo atribuible para estimar las posibilidades de prevención.

Riesgo atribuible

El capítulo 12 analizó el riesgo relativo y el cociente de posibilidades *(odds ratio)*, que suele utilizarse como un sustituto del riesgo relativo en un estudio de casos y controles. El *riesgo relativo* es importante como medida de la *fuerza de la asociación,* que es un aspecto destacado a la hora de obtener inferencias causales. Este capítulo se centra en una pregunta diferente: *¿qué proporción de la enfermedad se puede atribuir a una determinada exposición?* La respuesta procede de otra medida de riesgo, el *riesgo atribuible,* que se define como la cantidad o proporción de incidencia de la enfermedad (o riesgo de enfermedad) que se puede atribuir a una exposición específica. Por ejemplo, ¿qué proporción del riesgo de arteriopatía coronaria (AC) que presentan los fumadores se puede atribuir al tabaquismo? ¿Cuánto se puede atribuir a la genética del paciente? Mientras que el riesgo relativo es importante a la hora de establecer relaciones etiológicas, el riesgo atribuible es más importante en muchos aspectos en la práctica clínica y en salud pública, ya que aborda una cuestión diferente: ¿qué proporción del riesgo (incidencia) de la enfermedad es previsible que se pueda prevenir si somos capaces de eliminar la exposición al agente concreto?

Es posible calcular el riesgo atribuible para las personas expuestas (p. ej., el riesgo atribuible de AC en los fumadores) o el riesgo atribuible para la población total, lo que incluye tanto a personas expuestas como no expuestas (p. ej., el riesgo atribuible de AC en una población total, que consta tanto de fumadores como de no fumadores). Estos cálculos y sus usos e interpretaciones se describen en este capítulo.

Riesgo atribuible para el grupo expuesto

En la figura 13.1 se presenta una introducción esquemática a este concepto. Consideremos dos grupos: uno expuesto y otro no expuesto. En la figura 13.1A, el riesgo total de la enfermedad en el grupo expuesto se indica por toda la altura de la barra de la izquierda, y el riesgo total de enfermedad en el grupo no expuesto se indica por toda la altura de la barra de la derecha. Como se puede observar, el riesgo total de la enfermedad es mayor en el grupo expuesto que en el no expuesto. Se puede plantear la siguiente pregunta: en las personas expuestas, ¿qué proporción del riesgo total de la enfermedad se debe en realidad a la exposición? (p.ej., en un grupo de fumadores, ¿qué proporción del riesgo de AC se debe al tabaquismo?)

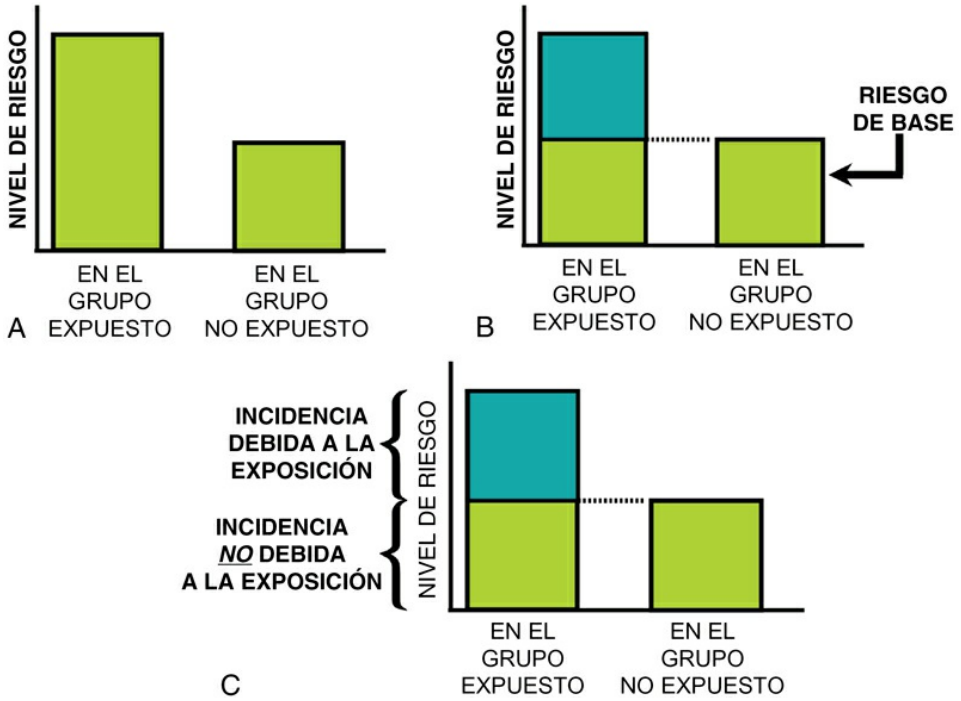


FIG. 13.1 (A) Riesgos totales en los grupos expuesto y no expuesto. (B) Riesgo de base. (C) Incidencia atribuible a la exposición e incidencia no atribuible a la exposición.

¿Cómo se puede responder a esta pregunta? Consideremos primero a las personas no expuestas, indicadas por la barra de la derecha. A pesar de que no están expuestas, tienen un cierto riesgo de presentar la enfermedad (aunque a un nivel mucho más bajo que el de las personas expuestas). Es decir, el riesgo de enfermedad no es cero incluso en las personas no expuestas. Por ejemplo, en este supuesto sobre el tabaquismo y la AC, incluso los no fumadores tienen un cierto riesgo de AC, debido posiblemente a la obesidad, la presión arterial, la concentración de colesterol y a otros factores. Este riesgo se denomina *riesgo de base.* Todas las personas comparten el riesgo de base, con independencia de si han tenido o no la exposición específica en cuestión (en este caso, el tabaquismo) (v. fig. 13.1B). Por consiguiente, tanto las personas no expuestas como las expuestas tienen este riesgo de base, de modo que el riesgo total de la enfermedad en los individuos expuestos es la suma del riesgo de base que cualquier persona tiene y el riesgo adicional debido a la exposición en cuestión. Si queremos saber qué proporción del riesgo total de las personas *expuestas se debe a la exposición,* hay que restar el riesgo de base del riesgo total (v. fig. 13.1C). Debido a que el riesgo en el grupo no expuesto es igual al riesgo de base, se puede calcular el riesgo en el grupo expuesto que es secundario a la exposición específica restando el riesgo en el grupo no expuesto (el riesgo de base) del riesgo total en el grupo expuesto.

Por tanto, la incidencia de una enfermedad que es atribuible a la exposición en el grupo expuesto se puede calcular de la siguiente manera:

Fórmula 13.1

$$\left(\begin{array}{c}\text{Incidencia en}\\ \text{el grupo expuesto}\end{array}\right)-\left(\begin{array}{c}\text{Incidencia en}\\ \text{el grupo no expuesto}\end{array}\right)$$

Otra alternativa sería plantear esta pregunta: «¿Qué *proporción* del riesgo en las personas expuestas se debe a la exposición?» A continuación, se podría expresar el riesgo atribuible como la *proporción* de la incidencia total en el grupo expuesto que es atribuible a la exposición simplemente dividiendo la fórmula 13.1 entre la incidencia en el grupo expuesto, de la siguiente manera:

Fórmula 13.2

$$\frac{\left(\begin{array}{c}\text{Incidencia en}\\ \text{el grupo expuesto}\end{array}\right)-\left(\begin{array}{c}\text{Incidencia en}\\ \text{el grupo no expuesto}\end{array}\right)}{\text{Incidencia en el grupo expuesto}}$$

El riesgo atribuible expresa la máxima reducción del riesgo de la enfermedad que sería previsible conseguir si se elimina por completo la exposición. Por ejemplo, si todos los fumadores fuesen inducidos a dejar de fumar, ¿qué proporción de reducción sería previsible en las tasas de AC? Desde un punto de vista programático práctico, el riesgo atribuible puede ser más relevante que el riesgo relativo. El riesgo relativo es una medida de la fuerza de la asociación, pero el riesgo atribuible indica el potencial de prevención si la exposición se pudiese eliminar.

El médico clínico está interesado principalmente en el riesgo atribuible *en el grupo expuesto:* por ejemplo, cuando un médico aconseja a un paciente que deje de fumar, está indicando al paciente que dejar de fumar reducirá el riesgo de AC. En este consejo está implícita la estimación del médico de que el riesgo del paciente se reducirá en una proporción determinada si se deja de fumar; esa reducción del riesgo motiva al médico a ofrecer el consejo. Aunque el médico a menudo no tiene en mente un valor específico del riesgo atribuible, sí tiene en cuenta que existe un riesgo atribuible en un grupo expuesto (fumadores) al que pertenece el paciente. El médico está planteando implícitamente esta pregunta: en una población de fumadores, ¿qué proporción de AC que presentan se debe al tabaquismo y, en consecuencia, qué proporción de AC se puede prevenir si no fumasen? El riesgo atribuible nos indica, por tanto, las posibilidades de prevención.

Si toda la incidencia de una enfermedad se debiese a un solo factor, el riesgo atribuible para esa enfermedad sería del 100%. Sin embargo, esto pocas veces sucede. Tanto el concepto como el cálculo del riesgo atribuible implican que no toda la incidencia de la enfermedad se debe a una sola exposición

específica, ya que la enfermedad se desarrolla incluso en algunos individuos no expuestos. En la figura 13.2 se recapitula este concepto.

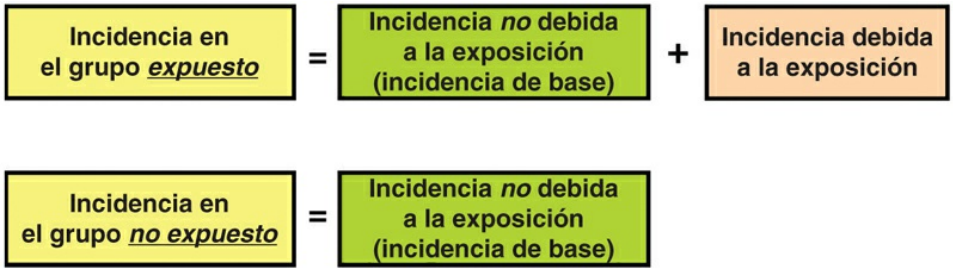


FIG. 13.2 Concepto de riesgo atribuible.

Riesgo atribuible para la población total: riesgo atribuible poblacional

Ahora se planteará una pregunta un tanto diferente en relación con el riesgo atribuible. Supongamos que sabemos cómo eliminar el tabaquismo. Le decimos al alcalde que tenemos una forma muy eficaz para eliminar el tabaquismo en la comunidad y queremos que nos proporcione los fondos para respaldar un programa de este tipo. El alcalde responde que está encantado de oír estas noticias, pero nos pregunta: «¿Cuál será el impacto de su programa para dejar de fumar sobre las tasas de incidencia de AC en nuestra ciudad?» Esta pregunta difiere de la que se acaba de comentar. Porque si tenemos en cuenta las tasas de AC en toda la población de una ciudad, y no solo en los individuos expuestos, estamos hablando de una población que se compone de fumadores y no fumadores. El alcalde no pregunta por el impacto que tendremos en los fumadores de esta ciudad, sino más bien por el impacto que tendrá sobre toda la población de la ciudad, que incluye tanto a fumadores como a no fumadores.

A continuación, ahondaremos más en esta pregunta. Además de suponer que tenemos un programa exitoso basado en la evidencia para dejar de fumar, supongamos también que todos los habitantes en la ciudad fuman. Ahora queremos calcular el riesgo atribuible. Es evidente que, debido a que todos los habitantes de la ciudad fuman, el riesgo atribuible para toda la población de la ciudad sería igual que el riesgo atribuible para la población expuesta. Si todo el mundo fuma, el riesgo atribuible para el grupo expuesto nos indica lo que sería previsible lograr con un programa para dejar de fumar dirigido a toda la población.

Ahora, supongamos que existe una situación ideal y que ningún habitante de la ciudad fuma. ¿Cuál será la posibilidad de prevenir la AC mediante el uso del programa para dejar de fumar completamente eficaz que se desea aplicar a la población de la ciudad? La respuesta es cero; dado que no hay personas expuestas en la ciudad, un programa dirigido a eliminar la

exposición no tiene sentido y, por tanto, no tendría ningún efecto sobre el riesgo de AC. Por tanto, el espectro del posible efecto se extiende desde un máximo (si todo el mundo fuma) hasta cero (si nadie fuma). Por supuesto, en realidad, la respuesta suele ser algún valor intermedio, debido a que algunos habitantes de la población fuman y otros no. Es evidente que el último grupo (todos los no fumadores) no se beneficiará de un programa para dejar de fumar, sin importar cuál sea su eficacia.

Hasta el momento, hemos comentado el concepto y el cálculo del riesgo atribuible para *un grupo expuesto.* Por ejemplo, en una población de fumadores, ¿qué proporción de los que tienen AC se debe al tabaquismo y, por tanto, qué proporción de la AC podría prevenirse si no fumasen? Sin embargo, para responder a la pregunta del alcalde sobre el efecto que tendría el programa para dejar de fumar sobre la población de la ciudad en su conjunto, hay que calcular el *riesgo atribuible para la población total:* ¿qué proporción de la incidencia de la enfermedad en el *total de la población* (incluidas tanto las personas expuestas como las no expuestas) se puede atribuir a una exposición específica? ¿Cuál sería el impacto total de un programa de prevención sobre la comunidad? Si queremos calcular el riesgo atribuible en la población total, el cálculo es similar al usado para las personas expuestas, pero se comienza con la *incidencia en la población total* y de nuevo se resta el riesgo de base, o la incidencia en la población no expuesta. La incidencia en la población total que se debe a la exposición[a] se puede calcular como se muestra en la fórmula 13.3.

Fórmula 13.3

$$\left(\begin{array}{c}\text{Incidencia en}\\ \text{la población total}\end{array}\right) - \left(\begin{array}{c}\text{Incidencia en}\\ \text{el grupo no expuesto}\\ \text{(riesgo de base)}\end{array}\right)$$

De nuevo, si se prefiere expresar esto como la *proporción* de la incidencia en la *población total* que es atribuible a la exposición, la fórmula 13.3 se puede dividir entre la incidencia en la población total:

Fórmula 13.4

$$\frac{\left(\begin{array}{c}\text{Incidencia en}\\ \text{la población total}\end{array}\right)-\left(\begin{array}{c}\text{Incidencia en}\\ \text{el grupo no expuesto}\end{array}\right)}{\text{Incidencia en la población total}}$$

El riesgo atribuible para la población total (riesgo atribuible poblacional [RAP]) es un concepto muy útil para quienes trabajan en salud pública. La pregunta que se plantea es: ¿qué *proporción* de AC en la *población total* se puede atribuir al tabaquismo? Esta pregunta podría reformularse de la siguiente manera: si se elimina el tabaquismo, ¿qué *proporción* de la incidencia de AC en la población total (que se compone de fumadores y no fumadores) se evitaría? La respuesta es: el riesgo atribuible *en la población total,* también denominado *riesgo atribuible poblacional* o *RAP* (como se ha comentado anteriormente)[b] .

Desde el punto de vista de la salud pública, esto suele ser tanto el aspecto crítico como la pregunta que plantean tanto los responsables políticos como los encargados de financiar los programas preventivos. Es posible que quieran saber lo que el programa propuesto va a aportar a la comunidad en su conjunto. ¿Cómo va a cambiar la carga sobre el sistema sanitario o la carga de sufrimiento en toda la comunidad, no solo en los individuos expuestos? Por ejemplo, si todos los fumadores de la comunidad dejasen de fumar, ¿cuál sería el impacto de este cambio en la incidencia de AC en toda la población de la comunidad (que incluye tanto a los fumadores como a los no fumadores)?

Ejemplo del cálculo del riesgo atribuible para el grupo expuesto

En esta sección se presenta un cálculo paso a paso del riesgo atribuible tanto en un grupo expuesto como en una población total. Se utilizará el ejemplo hipotético presentado previamente de un estudio de cohortes sobre tabaquismo y AC. Los datos se muestran de nuevo en la tabla 13.1.

Tabla 13.1

Tabaquismo y arteriopatía coronaria (AC): estudio de cohortes hipotético con 3.000 fumadore cigarrillos y 5.000 no fumadores

	Desarrollan AC	No desarrollan AC	Total
Fumadores	84	2.916	3.000
No fumadores	87	4.913	5.000
	$\text{Incidencia en los fumadores} = \frac{84}{3.000} = 28{,}0 \text{ por } 1.000$		

	Incidencia en los no fumadores $= \frac{87}{5.000} = 17,4$ por 1.000

La incidencia de AC en el grupo expuesto (fumadores) que es atribuible a la exposición se calcula utilizando la fórmula 13.1:

Fórmula 13.1

$$\left(\begin{array}{c}\text{Incidencia en}\\ \text{el grupo expuesto}\end{array}\right) - \left(\begin{array}{c}\text{Incidencia en}\\ \text{el grupo no expuesto}\end{array}\right) = \frac{28,0-17,4}{1.000} = \frac{10,6}{1.000}$$

¿Qué significa esto? Significa que 10,6 de los 28/1.000 casos incidentes en los fumadores son atribuibles al hecho de que estas personas fuman. Dicho de otra manera, si tuviéramos una campaña eficaz para dejar de fumar, sería previsible prevenir 10,6 de los 28/1.000 casos incidentes de AC que presentan los fumadores.

Si se prefiere, esto se puede expresar como una *proporción.* La proporción de la incidencia total en el grupo expuesto que es atribuible a la exposición se puede calcular dividiendo la fórmula 13.1 entre la incidencia en el grupo expuesto (fórmula 13.2):

Fórmula 13.2

$$\frac{\left(\begin{array}{c}\text{Incidencia en}\\ \text{el grupo expuesto}\end{array}\right) - \left(\begin{array}{c}\text{Incidencia en}\\ \text{el grupo no expuesto}\end{array}\right)}{\text{Incidencia en el grupo expuesto}} = \frac{28,0-17,4}{28,0} = \frac{10,6}{28,0} = 0,379 = 37,9\%$$

Por tanto, el 37,9% de la incidencia de AC en los fumadores puede ser atribuible al tabaquismo y, en teoría, podría prevenirse eliminando este hábito.

Ejemplo del cálculo del riesgo atribuible en la

población total (riesgo atribuible poblacional, RAP)

Usando el mismo ejemplo, a continuación se calculará el riesgo atribuible poblacional (RAP), que es el riesgo atribuible para la población total. La pregunta que se está planteando es: ¿qué es previsible lograr con nuestro programa para dejar de fumar *en la población total* (es decir, en toda la comunidad, que consta de fumadores y no fumadores)?

Hay que recordar que, en la población total, la incidencia que se debe al tabaquismo (la exposición) se puede calcular restando el riesgo de base (es decir, la incidencia en los no fumadores, o no expuestos) de la incidencia en la población total:

Fórmula 13.3

$$\left(\begin{array}{c}\text{Incidencia en}\\ \text{la población total}\end{array}\right)-\left(\begin{array}{c}\text{Incidencia en}\\ \text{el grupo no expuesto}\end{array}\right)$$

Para calcular la fórmula 13.3, se debe conocer *o bien* la incidencia de la enfermedad (AC) en la población total (que muchas veces se ignora) *o bien* la totalidad de los tres valores siguientes, a partir de lo cual es posible calcular la incidencia en la población total:

1. Incidencia en los fumadores.
2. Incidencia en los no fumadores.
3. Proporción de la población total que fuma.

En este ejemplo se sabe que la incidencia en los fumadores es de 28,0 por 1.000 y la incidencia en los no fumadores, de 17,4 por 1.000. Sin embargo, se desconoce la incidencia en la población total. Supongamos que, a partir de alguna otra fuente de información, se sabe que la proporción de fumadores en la población es del 44% (y, por tanto, la proporción de los no fumadores, del 56%). La incidencia en la población total se puede calcular entonces del siguiente modo:

$$\begin{pmatrix}\text{Incidencia}\\ \text{en fumadores}\end{pmatrix}\begin{pmatrix}\%\text{ de fumadores}\\ \text{en la población}\end{pmatrix} + \begin{pmatrix}\text{Incidencia}\\ \text{en no fumadores}\end{pmatrix}\begin{pmatrix}\%\text{ de no fumadores}\\ \text{en la población}\end{pmatrix}$$

(Simplemente se está ponderando el cálculo de la incidencia en la población total, teniendo en cuenta la proporción de la población que fuma y la proporción de la población que no fuma.)

Así, en este ejemplo, la incidencia en la población total se puede calcular de la siguiente manera:

$$\left(\frac{28,0}{1.000}\right)(0,44)+\left(\frac{17,4}{1.000}\right)(0,56)=\frac{22,1}{1.000}$$

Ahora tenemos los valores necesarios para el uso de la fórmula 13.3 con el fin de calcular el riesgo atribuible en la población total:

Fórmula 13.3

$$\begin{pmatrix}\text{Incidencia en}\\ \text{la población total}\end{pmatrix}-\begin{pmatrix}\text{Incidencia en}\\ \text{el grupo no expuesto}\end{pmatrix} = \frac{22,1}{1.000}-\frac{17,4}{1.000}=\frac{4,7}{1.000}$$

¿Qué nos dice esto? ¿Qué proporción del riesgo total de AC en esta población (que consta de fumadores y no fumadores) es atribuible al tabaquismo? Si tuviéramos un programa de prevención eficaz (dejar de fumar) en esta población, ¿qué proporción de reducción de la incidencia de AC sería previsible, en el mejor de los casos, en la población total (de fumadores y no fumadores)?

Si preferimos calcular la *proporción* de la incidencia en la *población total* que es atribuible a la exposición, puede hacerse dividiendo la fórmula 13.3 entre la incidencia en la población total como en la fórmula 13.4:

Fórmula 13.4

$$\frac{\left(\begin{array}{c}\text{Incidencia en}\\ \text{la población total}\end{array}\right)-\left(\begin{array}{c}\text{Incidencia en}\\ \text{el grupo no expuesto}\end{array}\right)}{\text{Incidencia en la población total}}$$

$$=\frac{22{,}1-17{,}4}{22{,}1}=21{,}3\%$$

Por tanto, el 21,3% de la incidencia de AC en la población total se puede atribuir al tabaquismo y, si un programa de prevención eficaz lograse que se deje de fumar, lo máximo que se podría esperar alcanzar sería una reducción del 21,3% de la incidencia de AC en la población total (que incluye tanto a fumadores como a no fumadores).

El *riesgo atribuible* es un concepto fundamental en casi todas las áreas de la salud pública y en la práctica clínica, en especial en relación con las preguntas sobre el potencial de las medidas preventivas. Por ejemplo, Lim y cols.[1] estimaron las causas reales de mortalidad mundial en 2010. Estas estimaciones emplearon los datos publicados y aplicaron cálculos del riesgo atribuible, así como otras estrategias. Sus estimaciones se muestran en la figura 13.3. Los autores indicaron que los factores de riesgo dietéticos y la inactividad física explicaban el 30% de todos los fallecimientos.

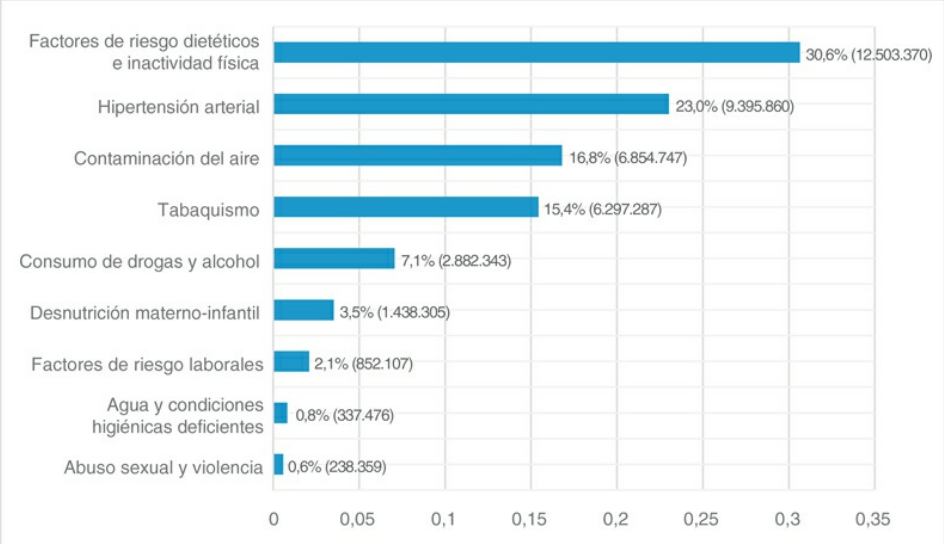


FIG. 13.3 Muertes atribuibles a determinados factores de riesgo o grupos de factores de riesgo en 2010, en todo el mundo. (Datos modificados de Lim SS, Vos T, Flaxman AD, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet.* 2012;380:2224–2260.)

También resulta interesante que, en el ámbito legal, en el que los pleitos por

perjuicios debidos a productos tóxicos se han convertido en algo cada vez más común, el concepto de riesgo atribuible en los individuos expuestos ha adquirido gran importancia. Uno de los criterios jurídicos utilizados en la búsqueda de una empresa responsable de un daño medioambiental, por ejemplo, es si es «más probable que no» que la empresa causase el daño. Se ha sugerido que un riesgo atribuible mayor del 50% podría representar una determinación cuantitativa de la definición legal de «más probable que no».

Comparación de riesgo relativo y riesgo atribuible

En capítulos anteriores se han analizado varias medidas de riesgo y de exceso de riesgo. El riesgo relativo y el cociente de posibilidades *(odds ratio)* son relevantes como medidas de la fuerza de la asociación, que es un aspecto fundamental a la hora de obtener una inferencia causal. El riesgo atribuible es una medida de la proporción del riesgo de enfermedad que es atribuible a una cierta exposición. Por tanto, el riesgo atribuible es útil para responder a la pregunta de qué proporción de enfermedad se puede prevenir si tenemos un medio eficaz para eliminar la exposición en cuestión. Así pues, el riesgo relativo es útil en los estudios etiológicos de la enfermedad, mientras que el riesgo atribuible tiene aplicaciones relevantes en la práctica clínica y la salud pública.

En la tabla 13.2 se muestra un ejemplo clásico de un estudio de Doll y Peto[2] que relaciona la mortalidad por cáncer de pulmón y AC en fumadores y no fumadores, y ofrece una comparación esclarecedora del riesgo relativo y del riesgo atribuible en el mismo conjunto de datos.

Tabla 13.2

Cáncer de pulmón y mortalidad por arteriopatía coronaria en médicos británicos varones: fumadores frente a no fumadores

	MORTALIDAD AJUSTADA POR EDAD POR 100.000				
	Fumadores	No fumadores	Riesgo relativo	Riesgo atribuible (fallecimientos por 100.000)	% de riesgo atribuible
Cáncer de pulmón	140	10	14,0	130	92,9
Arteriopatía coronaria	669	413	1,6	256	38,3

De Doll R, Peto R. Mortality in relation to smoking: Twenty years' observation on male British doctors. *Br Med J.* 1976;2:1525–1536.

Examinemos en primer lugar los datos del cáncer de pulmón. (Obsérvese que, en este ejemplo, se está utilizando la mortalidad como un sustituto del riesgo.) Se puede ver que el riesgo de mortalidad por cáncer de pulmón es de 140 para los fumadores y de 10 para los no fumadores. El riesgo relativo se puede calcular como 140/10 = 14.

Ahora, echemos un vistazo a los datos de la AC. La mortalidad por AC es de 669 en los fumadores y de 413 en los no fumadores. El riesgo relativo se

puede calcular como 669/413 = 1,6. Por tanto, el riesgo relativo es mucho mayor para el tabaquismo y el cáncer de pulmón que para el tabaquismo y la AC.

A continuación, pasemos a los riesgos atribuibles en los fumadores. ¿Qué proporción del riesgo total en los fumadores se puede atribuir al tabaquismo? Para calcular el riesgo atribuible, se resta el riesgo de base (que es el riesgo en el grupo no expuesto, en este caso los no fumadores) del riesgo en el grupo expuesto (los fumadores). Con los datos para el cáncer de pulmón que se han usado, esto es 140 – 10 = 130.

Para calcular el riesgo atribuible para la AC y el tabaquismo, se resta el riesgo en el grupo no expuesto (no fumadores) del riesgo en el grupo expuesto (fumadores), 669 – 413 = 256. Es decir, del total de 669 fallecimientos por 100.000 en los fumadores, 256 se pueden atribuir al tabaquismo.

Si se prefiere expresar el riesgo atribuible para el cáncer de pulmón y el tabaquismo como una proporción (es decir, la proporción del riesgo de cáncer de pulmón en los fumadores que se puede atribuir al tabaquismo), se divide el riesgo atribuible entre el riesgo en los fumadores:

$$\frac{(140-10)}{140} = 92,9\%$$

Si se prefiere expresar el riesgo atribuible para la AC y el tabaquismo como una proporción (la proporción del riesgo de AC en los fumadores que se puede atribuir al tabaquismo), se divide el riesgo atribuible entre el riesgo en los fumadores:

$$\frac{(669-413)}{669} = 38,3\%$$

¿Qué nos dice esta tabla? En primer lugar, se observa una gran diferencia entre los riesgos relativos para el cáncer de pulmón y la AC en relación con el tabaquismo: 14,0 para el cáncer de pulmón y 1,6 para la AC (es decir, existe una asociación mucho más fuerte para el tabaquismo y el cáncer de pulmón que para el tabaquismo y la AC). Sin embargo, el riesgo atribuible es casi el doble (256) para la AC que para el cáncer de pulmón (130). Si se opta por expresar el riesgo atribuible como proporción, se observa que el 92,9% de los fallecimientos por cáncer de pulmón en los fumadores pueden atribuirse al tabaquismo (y se pueden prevenir si se deja de fumar) en comparación con solo el 38,3% de los fallecimientos por AC en los fumadores que pueden atribuirse al tabaquismo.

Por tanto, el riesgo relativo es mucho mayor para el cáncer de pulmón que

para la AC y el riesgo atribuible expresado como proporción también es mucho mayor para el cáncer de pulmón. Sin embargo, si un programa eficaz para dejar de fumar estuviese disponible en la actualidad y el tabaquismo se eliminase, ¿sería el impacto preventivo mayor sobre la mortalidad por cáncer de pulmón o por AC? Si se analiza la tabla, se observa que, si se dejase de fumar, se podrían prevenir 256 fallecimientos por 100.000 debidos a AC, en contraste con solo 130 por cáncer de pulmón, a pesar de que el riesgo relativo es mayor para el cáncer de pulmón y a pesar del hecho de que la proporción de las muertes atribuibles al tabaquismo es mayor en el caso del cáncer de pulmón. ¿Cuál es la razón para ello? Esto se debe al hecho de que el nivel de mortalidad en los fumadores es mucho mayor por AC que por cáncer de pulmón (669 en comparación con 140) y a que el riesgo atribuible en los expuestos al tabaquismo (la diferencia entre el riesgo total en los fumadores y el riesgo de base) es mucho mayor para la AC que para el cáncer de pulmón.

Es importante enfatizar que, dado que el riesgo atribuible implica que se puede prevenir una cierta proporción del riesgo, solo debe estimarse cuando existe una certeza razonable de que la asociación del factor de riesgo con la enfermedad es causal[3]. Además, para los factores de riesgo con una exposición acumulada, es más apropiado definir los riesgos atribuibles como la proporción del riesgo (ya sea en la población expuesta o en toda la población) que puede atribuirse a una exposición determinada que como la proporción que puede eliminarse mediante el cese de la exposición. Por ejemplo, según el US Surgeon General's Report de 2010, el riesgo de cáncer de pulmón en los exfumadores 15 años después de haber dejado de fumar es la mitad que en los fumadores actuales, lo que significa que sigue siendo considerablemente más alto que el riesgo en los que nunca han fumado.

Conclusión

En este capítulo se ha introducido el concepto de riesgo atribuible y se ha descrito la forma de calcularlo e interpretarlo. El riesgo atribuible se resume en los cuatro cálculos que se muestran en la tabla 13.3.

Tabla 13.3

Resumen de los cálculos del riesgo atribuible

	En el grupo expuesto	En la población total
Incidencia atribuible a la exposición	$\left(\text{Incidencia en el grupo expuesto}\right) - \left(\text{Incidencia en el grupo no expuesto}\right)$	
Proporción de incidencia atribuible a la exposición	$\dfrac{\left(\text{Incidencia en el grupo expuesto}\right) - \left(\text{Incidencia en el grupo no expuesto}\right)}{\text{Incidencia en el grupo expuesto}}$	

Los conceptos de riesgo relativo y riesgo atribuible son esenciales para comprender la causalidad y el potencial de prevención. Por el momento, se han descrito varias medidas de riesgo: 1) riesgo absoluto, 2) riesgo relativo, 3) cociente de posibilidades *(odds ratio)* y 4) riesgo atribuible. En el capítulo 14 se repasarán brevemente los diseños de estudio y los conceptos de riesgo antes de proceder a una exposición sobre cómo se usan las estimaciones del exceso de riesgo para derivar inferencias causales.

Bibliografía

1. Lim SS, Vos T, Flaxman AD, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*. 2012;380:2224–2260.
2. Doll R, Peto R. Mortality in relation to smoking: twenty years' observations on male British doctors. *Br Med J*. 1976;2:1525–1536.
3. Rothman K, Greenland S. Modern Epidemiology. 3rd ed. Philadelphia: Wolters Kluwer Health/Lipincott; 2008.
4. Levin ML. The occurrence of lung cancer in man. *Acta Unio Int Contra Cancrum*. 1953;9:531.
5. Leviton A. Definitions of attributable risk. *Am J Epidemiol*. 1973;98:231.

Preguntas de revisión del capítulo 13

1. *En varios estudios se ha observado que alrededor del 85% de los casos de cáncer de pulmón se deben al tabaquismo. Esta medida es un ejemplo de:*

a. Tasa de incidencia.

b. Riesgo atribuible.

c. Riesgo relativo.

d. Riesgo de prevalencia.

e. Cociente de mortalidad proporcional.

Las preguntas 2 y 3 se basan en la siguiente información.

Los resultados de un estudio de cohortes de 10 años sobre el tabaquismo y la arteriopatía coronaria (AC) se muestran en la siguiente tabla:

	RESULTADO DESPUÉS DE 10 AÑOS	
Al comienzo del estudio	**Desarrollaron AC**	**No desarrollaron AC**
2.000 fumadores sanos	65	1.935
4.000 no fumadores sanos	20	3.980

2. *La incidencia de AC en los fumadores que se puede atribuir al tabaquismo es de: ____*

3. *La proporción de la incidencia total de AC en los fumadores que es atribuible al tabaquismo es: ___*

Las preguntas 4 y 5 se basan en la siguiente información.

En un estudio de cohortes sobre el tabaquismo y el cáncer de pulmón, la incidencia de cáncer de pulmón en los fumadores fue de 9/1.000 y la incidencia entre los no fumadores, de 1/1.000. A partir de otra fuente se sabe que el 45% de la población total eran fumadores.

4. *La incidencia de cáncer de pulmón atribuible al tabaquismo en la población total es de: ____*

5. *La proporción del riesgo en la población total que es atribuible al tabaquismo es: ____*

Apéndice del capítulo 13: fórmula de Levin para calcular el riesgo atribuible para la población total

Otra forma de calcular esta proporción para la población total es utilizar la fórmula de Levin[4]:

$$\frac{p(r-1)}{p(r-1)+1}$$

donde *p* es la proporción de la población con la característica o exposición y *r* es el riesgo relativo (o cociente de posibilidades).

Leviton[5] ha demostrado que la fórmula de Levin[4] y la siguiente fórmula son idénticas desde el punto de vista algebraico:

$$\frac{\left(\begin{array}{c}\text{Incidencia en}\\ \text{la población total}\end{array}\right)-\left(\begin{array}{c}\text{Incidencia en}\\ \text{el grupo no expuesto}\end{array}\right)}{\text{Incidencia en la población total}}$$

---

[a] La incidencia en la población debida a la exposición también se puede calcular de la siguiente manera: Riesgo atribuible para el grupo expuesto × proporción de la población expuesta.

[b] Otra forma de calcular el riesgo atribuible para la población total es utilizar la fórmula de Levin, que aparece en el apéndice de este capítulo. La fórmula de Levin permite la estimación del RAP utilizando datos de casos y controles al reemplazar el riesgo relativo con el cociente de posibilidades (si la enfermedad es relativamente rara). Sin embargo, requiere una estimación de la prevalencia de la exposición en la población de referencia. Tenga en cuenta que la fórmula de Levin solo se aplica a los RAP basados en los riesgos relativos no ajustados.

CAPÍTULO 14

De la asociación a la causalidad: derivación de inferencias a partir de estudios epidemiológicos

Palabras clave

asociación falsa
etiología
confusión
asociación causal
vía causal directa o vía causal indirecta
directrices del Surgeon General estadounidense sobre causalidad
grados del Grupo de Trabajo de los Servicios Preventivos de EE. UU.

No todo lo que se puede contar cuenta, y no todo lo que cuenta se puede contar.

—William Bruce Cameron, 1963[1]

Objetivos de aprendizaje

- Describir una secuencia frecuente de diseños de estudio utilizados para abordar cuestiones de etiología en las poblaciones humanas.
- Diferenciar entre las asociaciones verdaderas y falsas en los estudios observacionales.
- Definir los conceptos de «necesario» y «suficiente» en el contexto de las relaciones causales.
- Presentar directrices para juzgar si una asociación es causal basándose en las directrices establecidas por el Surgeon General estadounidense y describir la aplicación de estas directrices a cuestiones más amplias de inferencia causal.

- Describir cómo las directrices referentes a la causalidad propuestas originalmente por el Surgeon General estadounidense han sido modificadas y utilizadas por el Servicio Público de Salud y el Grupo de Trabajo de los Servicios Preventivos de EE. UU.

En los capítulos anteriores se han descrito los distintos diseños de los estudios epidemiológicos que se utilizan para determinar si existe una asociación entre una exposición y el resultado de una enfermedad (fig. 14.1A). A continuación, se han expuesto diferentes tipos de medición del riesgo que se utilizan para expresar cuantitativamente un exceso de riesgo. Si se determina que una exposición se asocia con una enfermedad, la siguiente pregunta es si la asociación observada refleja una relación causal (v. fig. 14.1B).

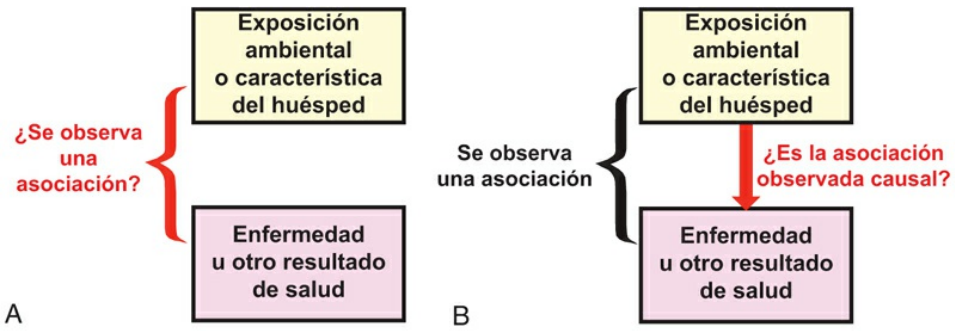


FIG. 14.1 (A) ¿Estamos observando una asociación entre exposición y enfermedad? (B) ¿La asociación observada entre exposición y enfermedad es causal?

Aunque las figuras 14.1A y B se refieren a una exposición ambiental, también podrían corresponder a una característica genética o alguna otra característica de riesgo o a una combinación específica de factores ambientales y genéticos. Como se verá en el capítulo sobre genética y factores ambientales, los estudios sobre la etiología de las enfermedades suelen analizar las contribuciones de los factores genéticos y ambientales, así como sus interacciones.

En este capítulo se describe la derivación de inferencias causales en epidemiología. Para empezar, se planteará la siguiente pregunta: «¿De qué estrategias se dispone para estudiar la etiología de las enfermedades?»

Estrategias para el estudio de la etiología de las enfermedades

Si estamos interesados en saber si una determinada sustancia es carcinógena para el ser humano, un primer paso en el estudio de su efecto podría consistir en exponer a animales a dicha sustancia en un entorno de laboratorio controlado. Aunque tales estudios con animales permiten controlar la dosis de exposición y otras condiciones ambientales, así como los factores genéticos con precisión, además de minimizar las pérdidas durante el seguimiento, al finalizar el estudio queda el problema de tener que extrapolar los datos de una especie a otra (es decir, de la población de animales a la de seres humanos). Ciertas enfermedades observadas en los seres humanos ni se han producido ni se han inducido en animales. También es difícil extrapolar las dosis empleadas en los animales a las usadas en el ser humano, y las especies difieren en sus respuestas. Por tanto, aunque tales estudios toxicológicos pueden ser útiles, no despejan todas las incertidumbres en cuanto a si los hallazgos en animales se pueden generalizar a los seres humanos.

También se pueden usar sistemas *in vitro*, como el cultivo celular o de órganos. Sin embargo, debido a que estos son sistemas artificiales, de nuevo surge la dificultad de extrapolar los resultados obtenidos en sistemas artificiales a los organismos humanos íntegros e intactos.

A la vista de estas limitaciones, si queremos ser capaces de llegar a una conclusión en cuanto a si una sustancia causa la enfermedad en los seres humanos, es necesario hacer *observaciones en las poblaciones humanas.* Dado que, desde el punto de vista ético o práctico, no es posible distribuir de forma aleatoria a los seres humanos a la exposición a un carcinógeno, dependemos de observaciones no aleatorizadas, como las que provienen de los estudios de casos y controles y de cohortes.

Enfoques de la etiología en las poblaciones humanas

La epidemiología a menudo aprovecha lo que se ha venido en llamar experimentos «no planificados» o «naturales». (Hay quien piensa que esta frase es una contradicción en los términos, puesto que la palabra «experimento» implica una exposición planeada.) Lo que queremos expresar cuando decimos experimentos *no planificados* o *naturales* es que utilizamos grupos de personas que han estado expuestas por fines distintos a un estudio, como las cohortes laborales de industrias específicas o personas expuestas a productos químicos tóxicos. Algunos ejemplos son los afectados por la catástrofe de la fuga de gas venenoso en una planta de fabricación de pesticidas en Bhopal, India, en 1984, y los habitantes de Hiroshima y Nagasaki, Japón, que fueron expuestos a la radiación de las bombas atómicas

lanzadas sobre ambas ciudades en 1945 por fuerzas estadounidenses. Cada uno de estos grupos expuestos se puede comparar con un grupo no expuesto (p.ej., residentes de Chennai, India, o de Tokio, Japón) para determinar si hay un aumento del riesgo de un cierto efecto adverso en las personas que han estado expuestas.

Cuando se realizan estudios en seres humanos, suele seguirse la secuencia que se muestra en la figura 14.2.



FIG. 14.2 Secuencia frecuente de estudios en poblaciones humanas.

El primer paso puede consistir en la *observación clínica* a la cabecera del paciente. Por ejemplo, cuando el cirujano Alton Ochsner observó que prácticamente todos los pacientes a quienes operó por cáncer de pulmón tenían antecedentes de tabaquismo, fue uno de los primeros en sugerir una posible relación causal[2]. Un segundo paso consiste en tratar de identificar los *datos disponibles de forma rutinaria,* cuyo análisis podría arrojar luz sobre la

pregunta. A continuación, es posible llevar a cabo *nuevos estudios*, como los estudios de cohortes y de casos y controles descritos en capítulos anteriores, que están diseñados específicamente para determinar si existe una asociación entre una exposición y una enfermedad, y si existe una relación causal.

El primer paso habitual cuando se realizan nuevos estudios para explorar una relación suele ser un *estudio de casos y controles.* Por ejemplo, si Ochsner hubiese querido explorar más su sugerencia de que el consumo de cigarrillos puede estar asociado con el cáncer de pulmón, habría comparado los antecedentes de tabaquismo de un grupo de sus pacientes con cáncer de pulmón con los de un grupo de pacientes sin dicho cáncer, es decir, habría realizado un estudio de casos y controles.

Si un estudio de casos y controles proporciona evidencia de que una cierta exposición es sospechosa, el siguiente paso podría ser hacer un *estudio de cohortes* (p. ej., comparar a los fumadores y no fumadores y determinar la tasa de cáncer de pulmón en cada grupo o comparar los trabajadores expuestos a una toxina industrial con los trabajadores sin tal exposición). Aunque, en teoría, un ensayo aleatorizado podría ser el siguiente paso, como se ha explicado anteriormente, los ensayos aleatorizados casi nunca se usan para estudiar los efectos de las supuestas toxinas o agentes cancerígenos y, por lo general, solo se emplean para la evaluación de agentes potencialmente beneficiosos.

Desde el punto de vista conceptual, se sigue un proceso de dos pasos al llevar a cabo estudios y evaluar la evidencia. Sin embargo, en la práctica, este proceso a menudo se convierte en interactivo y no sigue una secuencia fija:

1. Se determina si existe una asociación o correlación entre una exposición o característica y el riesgo de una enfermedad (fig. 14.3). Para ello, se utilizan:
    a. Estudios de las características del grupo: estudios ecológicos (descritos en el cap. 7).
    b. Estudios de las características individuales: cohortes, casos y controles y otros tipos de estudios.
2. Si se demuestra una asociación, se determina si es probable que la asociación observada sea causal.



FIG. 14.3 «Me llega un correo como ese cada vez que salgo de tu despacho. ¿Pensabas que no iba a darme cuenta de la correlación?» Otro ejemplo de

asociación o causalidad. (DILBERT © 2011 Scott Adams. Usado con autorización de ANDREWS MCMEEL SYNDICATION. Todos los derechos reservados.)

Tipos de asociaciones

Asociaciones reales o espurias

A continuación, se describirán los tipos de asociaciones que pueden observarse en un estudio de cohortes o de casos y controles. Si se observa una asociación, la pregunta inicial será: «¿Es una asociación verdadera (real) o falsa (espuria)?» Por ejemplo, si se diseña un estudio para seleccionar los controles de tal manera que tiendan a no presentar la exposición, podría observarse una asociación entre la exposición y la enfermedad (es decir, una exposición más frecuente en los casos que en los controles). Esto no sería una verdadera asociación, sino solo un resultado del diseño del estudio. Debe recordarse que esta cuestión se planteó en el capítulo 7 en relación con un estudio sobre el consumo de café y el cáncer de páncreas. Se sugirió la posibilidad de que los controles seleccionados para el estudio tuviesen una menor tasa de consumo de café de la que se observaba en la población general.

Interpretación de asociaciones reales

Si la asociación observada es real, lo siguiente que hay que plantearse es si es causal. En la figura 14.4 se muestran dos posibilidades. En la figura 14.4A se presenta una asociación causal: se observa una asociación de la exposición y la enfermedad, como se indica con la llave, y la exposición induce el desarrollo de la enfermedad, como se indica con la flecha. En la figura 14.4B se muestra la misma asociación observada de exposición y enfermedad, pero esta asociación se debe únicamente a que ambas están vinculadas a un tercer factor, conocido como variable de confusión y denominado aquí *factor X*. Esta asociación se debe a un factor de confusión y no es causal. Los factores de confusión se describen con mayor detalle en el capítulo 15.

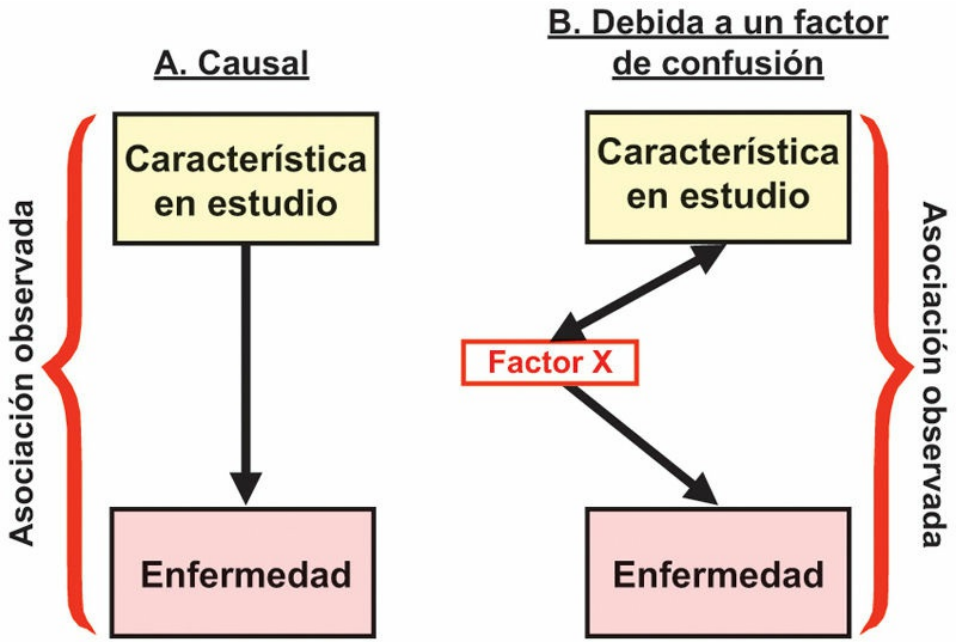


FIG. 14.4 Tipos de asociaciones.

En el capítulo 7 se expuso esta cuestión en relación con el estudio de McMahon sobre el café y el cáncer de páncreas. McMahon observó una asociación entre el consumo de café y el riesgo de cáncer de páncreas. Se sabe que el tabaquismo se asocia con el cáncer de páncreas, y el consumo de café y el tabaquismo están estrechamente asociados (en la época del estudio había pocos fumadores que no bebiesen café) (fig. 14.5). Por tanto, surge la pregunta de si la asociación observada entre el consumo de café y el cáncer de páncreas era probablemente de tipo causal, o si dicha asociación podría deberse al hecho de que el café y el tabaquismo están asociados, y a que el tabaquismo es un factor de riesgo conocido para el cáncer de páncreas.

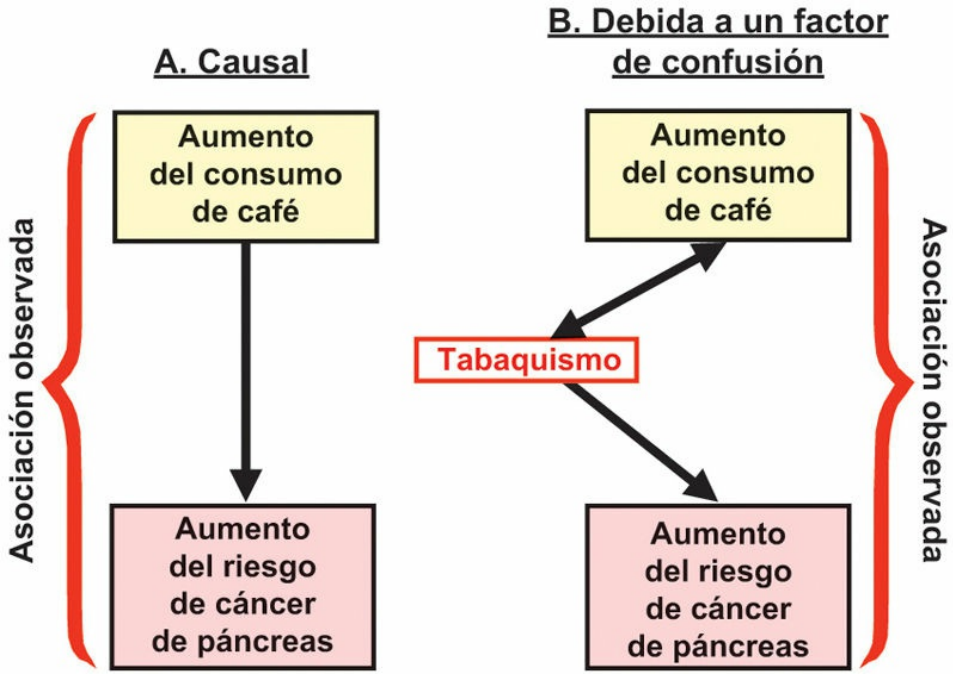


FIG. 14.5 Interpretación de una asociación observada entre el aumento del consumo de café y el aumento del riesgo de cáncer de páncreas.

El mismo problema se ejemplifica con la asociación observada entre la hipercolesterolemia y el riesgo de arteriopatía coronaria (AC) (fig. 14.6). ¿Es la inactividad física un factor causal de un mayor riesgo de cáncer de colon, o la asociación observada se debe a un factor de confusión? Es decir, ¿estamos observando una asociación entre la inactividad física y el cáncer de colon debido a que ambas se asocian con un factor X (como el tabaquismo), que podría hacer que las personas fueran sedentarias y tuviesen un mayor riesgo de cáncer de colon?

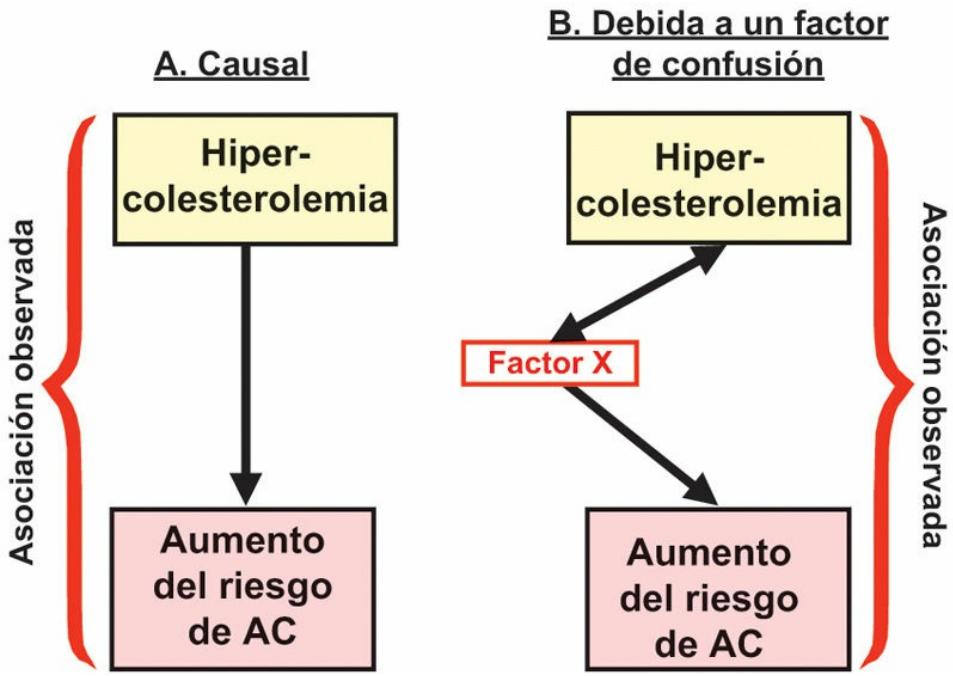


FIG. 14.6 Interpretación de una asociación observada entre la hipercolesterolemia y el aumento del riesgo de arteriopatía coronaria *(AC)*.

¿Es realmente importante esta distinción? ¿Qué diferencia supone? La respuesta es que conlleva una gran diferencia desde los puntos de vista clínico y de salud pública. Si la relación es causal, se logrará reducir el riesgo de cáncer de colon si se fomenta la actividad física, tanto a nivel individual como en toda la población. Sin embargo, si la relación se debe a un factor de confusión, el aumento del riesgo de cáncer de colon se debe al factor X. Por tanto, el aumento de la actividad física no tendrá ningún efecto sobre el riesgo de cáncer de colon. Así pues, es crucial poder distinguir si la asociación se debe a una relación causal o a un factor de confusión (que no es causal).

A continuación se presenta otro ejemplo. Desde hace muchos años se sabe que el tabaquismo de las mujeres embarazadas se asocia con bajo peso al nacer en sus lactantes. Como se ve en la figura 14.7, el efecto no es solo el resultado del nacimiento de unos bebés con bajo peso al nacer en este grupo de mujeres, sino que toda la curva de distribución del peso se desplaza hacia la izquierda en los bebés nacidos de fumadoras. La reducción del peso al nacer tampoco se debe a embarazos más cortos. Los bebés de las fumadoras son más pequeños que los de las no fumadores a cualquier edad gestacional (fig. 14.8). Asimismo, se observa una relación dosis-respuesta (fig. 14.9). Cuanto más fuma una mujer, mayor es su riesgo de tener un bebé con bajo peso al nacer. Durante muchos años, la interpretación de esta asociación ha sido objeto de gran controversia. Muchos autores creían que la asociación reflejaba una relación causal. Otros, como un estadístico de vanguardia, Jacob Yerushalmy, creían que la asociación se debía a un factor de confusión y que no era causal. Este autor escribió lo siguiente:

Una comparación entre las fumadoras y no fumadoras muestra que ambas difieren notablemente en lo que respecta a muchas variables ambientales, conductuales y biológicas. Por ejemplo, es menos probable que las fumadoras usen anticonceptivos y que planifiquen el embarazo. Es más probable que las fumadoras beban café, cerveza y whisky y que las no fumadoras beban té, leche y vino. Las fumadoras tienen más probabilidades que las no fumadoras de incurrir en el exceso en dichos hábitos. En general, las no fumadoras se muestran más moderadas que las fumadoras, que resultan ser más extremas y descuidadas en su modo de vida. También se han señalado algunas diferencias biológicas entre ellas: por ejemplo, las fumadoras tienen una mayor tasa de partos gemelares (solo en mujeres de raza blanca) y la edad de la menarquia es inferior a la de las no fumadoras[3].

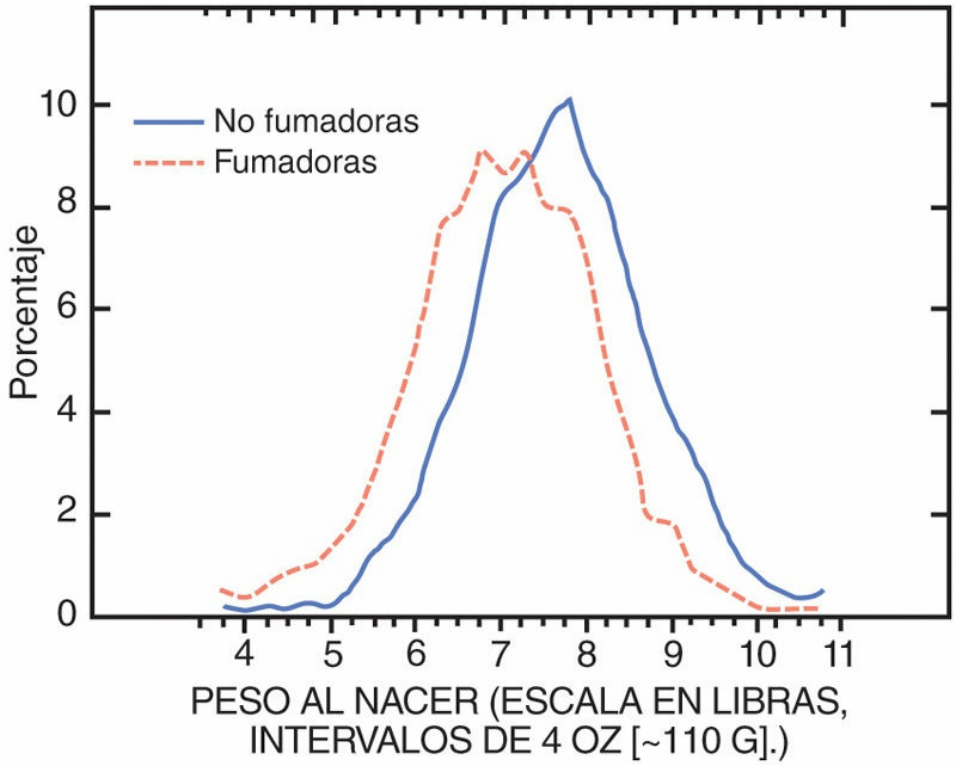


FIG. 14.7 Distribución porcentual por peso al nacer de los lactantes de madres que no fumaban durante el embarazo y de madres que fumaban 1 paquete de cigarrillos o más al día. (Del U.S. Department of Health, Education, and Welfare. *The Health Consequences of Smoking*. Washington, DC: Public Health Service; 1973:105.)

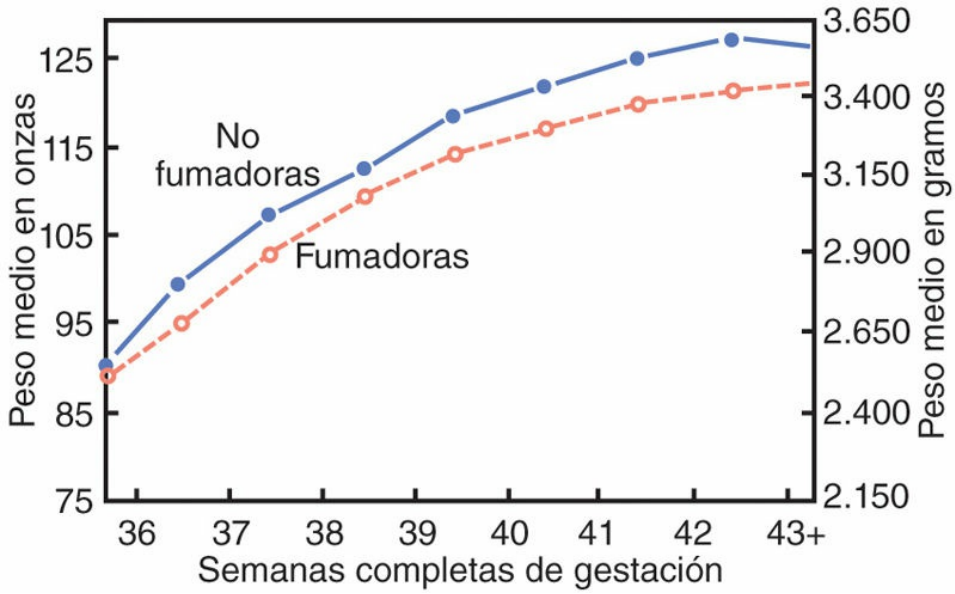


FIG. 14.8 Peso medio al nacer por semana de gestación en función del tabaquismo materno. (Del U.S. Department of Health, Education, and Welfare. *The Health Consequences of Smoking.* Washington, DC: Public Health Service; 1973:104.)

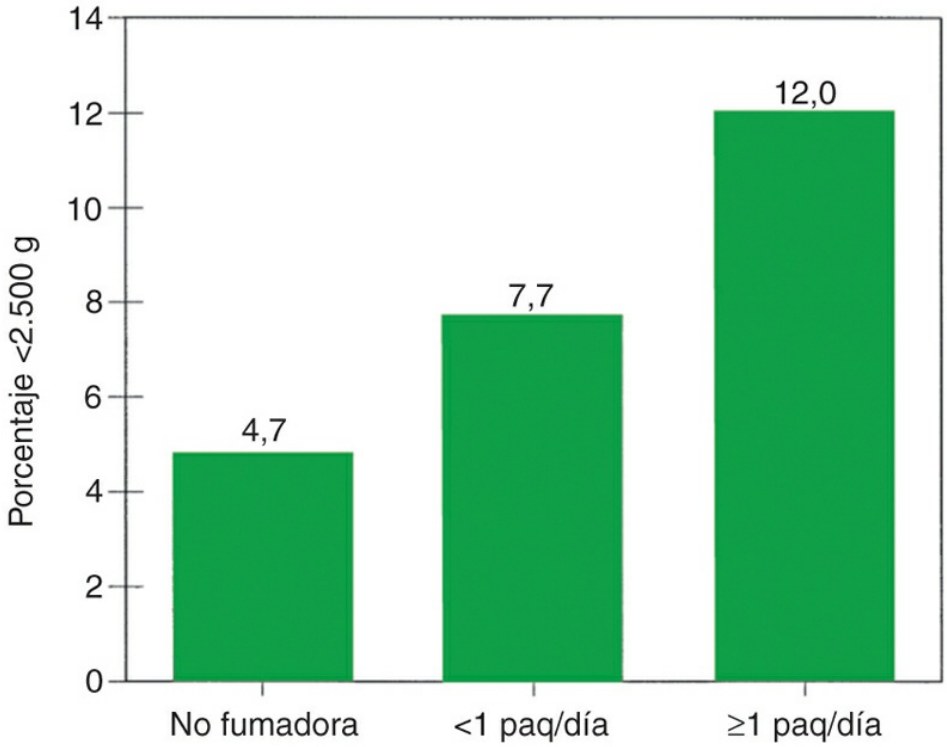


FIG. 14.9 Porcentaje de embarazos (*n* = 50.267) con un peso del lactante menor de 2.500 g, por categoría de tabaquismo materno. (Reproducida del Ontario Department of Health. *Second Report of the Perinatal Mortality Study in Ten University Teaching Hospitals.* Vol. I. Toronto, Ontario: Department of Health, Ontario Perinatal Mortality Study Committee; 1967:275.)

A la vista de estas numerosas diferencias entre fumadoras y no fumadoras, Yerushalmy pensó que no era el *tabaquismo* el factor causante del bajo peso al nacer, sino que este bajo peso era atribuible a *otras características de los fumadores.* Es interesante analizar un estudio llevado a cabo por Yerushalmy para apoyar su opinión en aquel momento (fig. 14.10)[3].

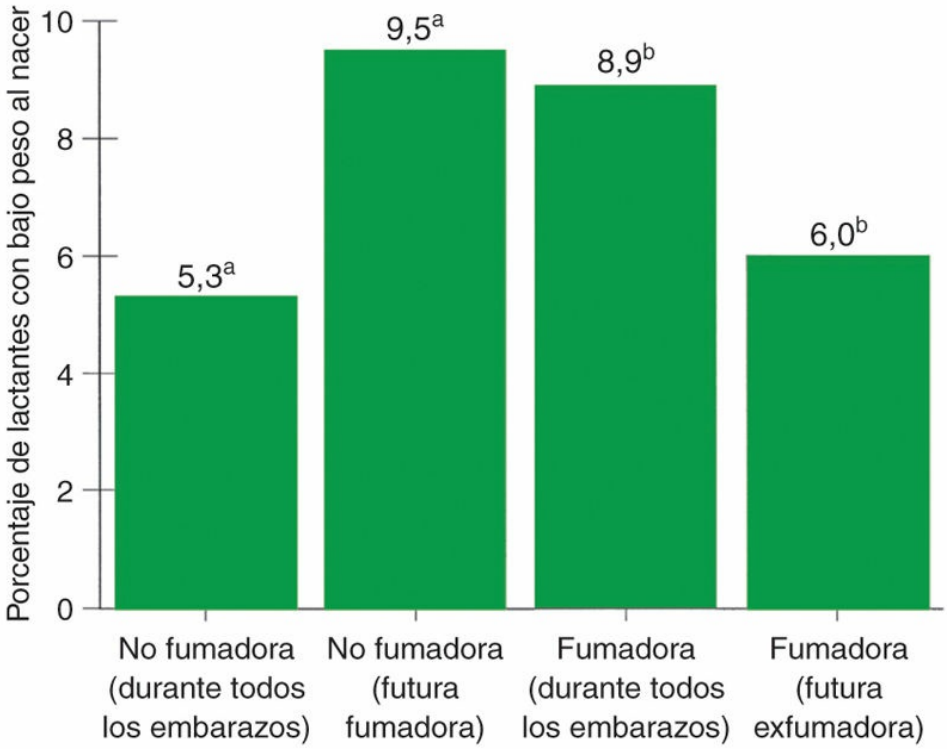


FIG. 14.10 Porcentaje de lactantes con bajo peso al nacer según el tabaquismo materno. ([a]*P* <0,01; [b]*P* <0,02). (Reproducida de Yerushalmy J. Infants with low birth weight born before their mothers started to smoke cigarettes. *Am J Obstet Gynecol*. 1972;112:277–284.)

Yerushalmy analizó los resultados de un embarazo (el embarazo de estudio) en una población de mujeres que habían tenido varios embarazos. La tasa de bebés con bajo peso al nacer en el embarazo de estudio fue del 5,3% para las mujeres que eran no fumadoras en *todos* sus embarazos. Sin embargo, si eran fumadoras en todos sus embarazos, la tasa de bajo peso al nacer en el embarazo de estudio era casi del 9%. Cuando analizó los embarazos de mujeres que eran no fumadoras durante el embarazo de estudio, pero que más tarde se convirtieron en fumadoras, se observó que su tasa de bebés con bajo peso al nacer era aproximadamente igual a la de las mujeres que fumaban en todos los embarazos. Cuando analizó los embarazos de mujeres que eran fumadoras en el embarazo de estudio, pero que posteriormente dejaron de fumar, observó que su tasa de bebés con bajo peso al nacer en el embarazo de estudio era similar a la de las mujeres que eran no fumadoras en todos sus embarazos.

Basándose en estos datos, Yerushalmy llegó a la conclusión de que no era el tabaquismo, sino más bien alguna característica de la fumadora, lo que causaba el bajo peso al nacer. En la actualidad, sin embargo, se acepta de forma casi universal que el tabaquismo es una causa de bajo peso al nacer. La naturaleza causal de esta relación también se ha demostrado en ensayos aleatorizados que han reducido la frecuencia de bajo peso al nacer mediante el inicio de los programas para que las mujeres embarazadas dejen de fumar. Aunque este tema está resuelto en gran medida en la actualidad, es esclarecedor revisar tanto la controversia como el estudio, ya que ejemplifican

el razonamiento que es necesario para tratar de distinguir las interpretaciones causales de las no causales de las asociaciones observadas.

Tipos de relaciones causales

Una vía causal puede ser *directa* o *indirecta* (fig. 14.11). En la relación de causalidad *directa,* un factor provoca directamente una enfermedad sin ningún paso intermedio. En la causalidad *indirecta,* un factor provoca una enfermedad, pero solo a través de un paso o pasos intermedios. En la biología humana, los pasos intermedios están presentes casi siempre en cualquier proceso causal.

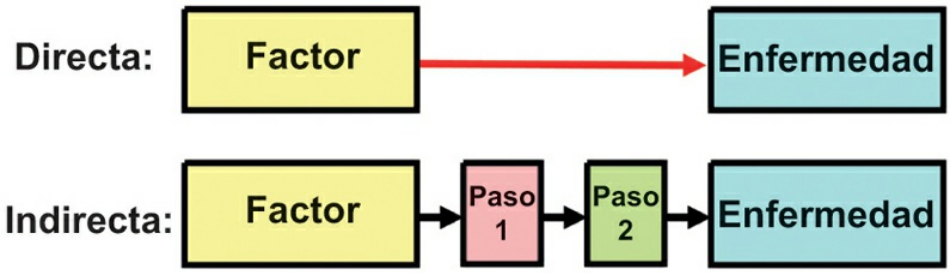


FIG. 14.11 Causas directas frente a causas indirectas de enfermedad.

Si una relación es causal, puede haber cuatro tipos de relaciones causales: 1) necesaria y suficiente, 2) necesaria, pero no suficiente, 3) suficiente, pero no necesaria, y 4) ni suficiente ni necesaria.

Necesaria y suficiente

En el primer tipo de relación causal, un factor es tanto necesario como suficiente para producir la enfermedad. Sin ese factor, la enfermedad nunca se desarrolla (el factor es necesario), y, en presencia de ese factor, la enfermedad siempre se desarrolla (el factor es suficiente) (fig. 14.12). Esta situación se produce pocas veces o nunca. Por ejemplo, en la mayoría de las enfermedades infecciosas, un número de personas están expuestas, algunas de las cuales manifiestan la enfermedad y otras no. Los miembros de los hogares de una persona con tuberculosis no adquieren de manera uniforme la enfermedad a partir del caso índice. Si se supone que la dosis de exposición es igual, es probable que existan diferencias en el estado inmunitario, la susceptibilidad genética u otras características que determinan quién desarrolla la enfermedad y quién no. Una relación uno-a-uno entre la exposición y la enfermedad, que es consecuencia de una relación necesaria y suficiente, se produce pocas veces o nunca.



FIG. 14.12 Tipos de relaciones causales: I. El factor A es necesario y suficiente.

Necesaria, pero no suficiente

En otro modelo, cada factor es necesario, pero no suficiente por sí mismo, para causar la enfermedad (fig. 14.13). Se requieren, por tanto, múltiples factores, a menudo en una secuencia temporal específica. Por ejemplo, se considera que la carcinogénesis es un proceso de múltiples etapas que implica tanto la iniciación como la promoción. Para que se produzca un cáncer, un promotor debe actuar después de que lo haya hecho un iniciador. La acción de un iniciador o de un promotor por sí sola no producirá un cáncer.

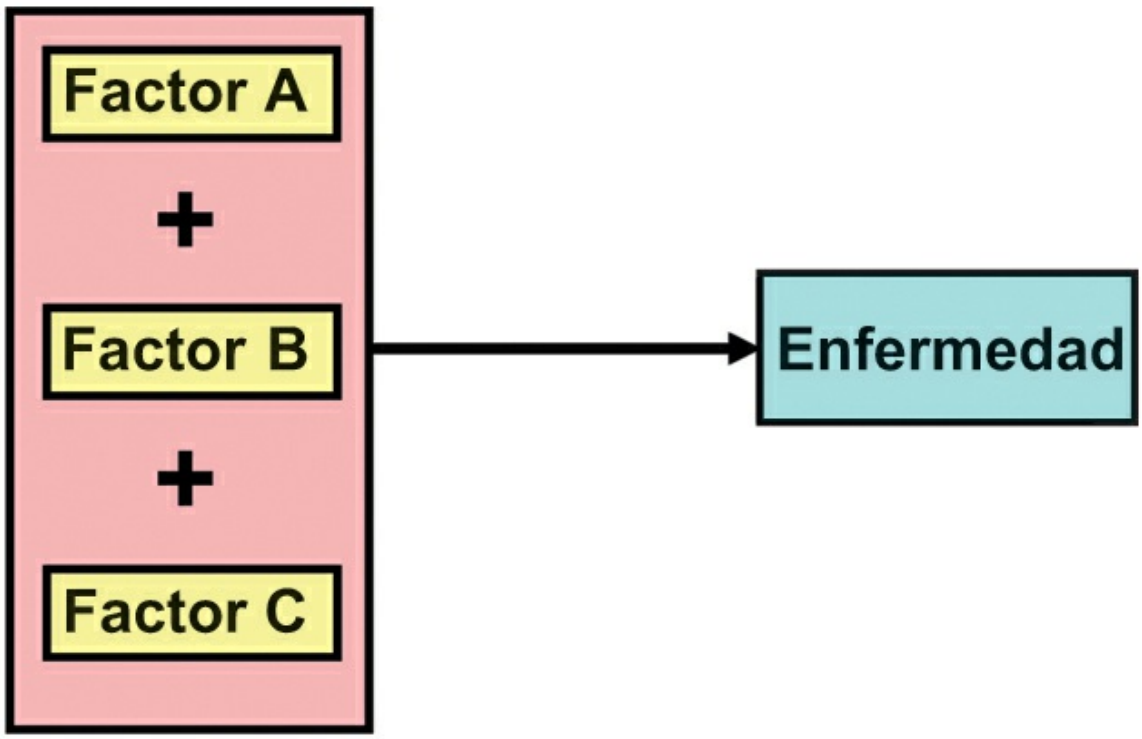


FIG. 14.13 Tipos de relaciones causales: II. Cada factor es necesario, pero no suficiente.

Asimismo, en la tuberculosis, el bacilo tuberculoso es claramente un factor necesario, a pesar de que su presencia puede no ser suficiente para producir la enfermedad en todas las personas infectadas. Otro ejemplo es la relación de *Helicobacter pylori* con el adenocarcinoma gástrico que no compromete el cardias[4]. Aunque esta bacteria es considerada como una causa necesaria de este cáncer, la mayoría de las personas infectadas con *H. pylori* no desarrollan

este tipo de cáncer, un fenómeno que explica por qué la prevalencia de infección por *H. pylori* es muy alta en muchas poblaciones, aunque, el adenocarcinoma gástrico sigue siendo relativamente raro en estas mismas poblaciones. Por lo tanto, además de *H. pylori,* los individuos deben estar expuestos a otros factores de riesgo (p.ej., tabaquismo y la ingesta de alimentos que contienen nitratos) para desarrollar cáncer gástrico. (La relación de *H. pylori* con las úlceras gástricas se estudia más adelante en este capítulo).

Suficiente, pero no necesaria

En este modelo, el factor por sí solo puede producir la enfermedad, pero también pueden hacerlo otros factores que actúen solos (fig. 14.14). Por consiguiente, tanto la exposición a la radiación como la exposición al benceno pueden producir leucemia sin la presencia de la otra. Incluso en esta situación, sin embargo, el cáncer no se desarrolla en todas las personas que hayan estado expuestas a la radiación o al benceno, por lo que, aunque ambos factores no son necesarios, es probable que otros cofactores sí lo sean. Así, el criterio de *suficiente* pocas veces lo cumple un solo factor.

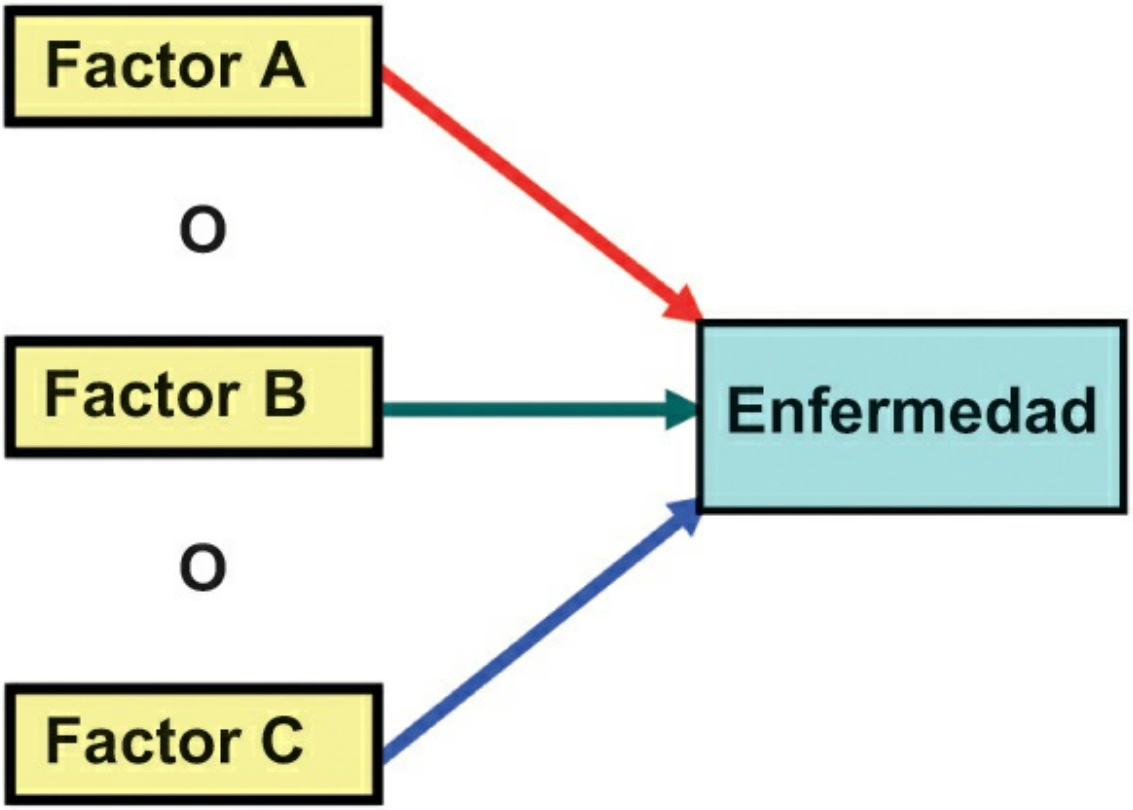


FIG. 14.14 Tipos de relaciones causales: III. Cada factor es suficiente, pero no necesario.

Ni suficiente ni necesaria

En el cuarto modelo, un factor, por sí mismo, no es suficiente ni necesario para producir la enfermedad (fig. 14.15). Éste es un modelo más complejo y representa probablemente con mayor precisión las relaciones causales que intervienen en la mayoría de las enfermedades crónicas. Un ejemplo es el de la naturaleza a menudo no superpuesta de los grupos de factores de riesgo para el desarrollo de AC; por ejemplo, los individuos pueden desarrollar AC si están expuestos al tabaquismo, a la diabetes y a la concentración baja de lipoproteínas de alta densidad (HDL) o a una combinación de hipercolesterolemia, hipertensión y sedentarismo. Cada uno de estos factores de riesgo de AC no es suficiente ni necesario. Curiosamente, al reconocer que muchos factores de riesgo individuales, si no la mayoría, no son suficientes ni necesarios, Rothman ha propuesto un modelo consistente con la figura 14.15, en el que una «causa suficiente» está formada por una *constelación* de factores de riesgo, denominada por él «causas componentes». En la conceptualización de Rothman, un gráfico circular formado por una serie de «causas componentes» representa la «causa suficiente». Por lo tanto, la «causa suficiente» de Rothman es en realidad un conjunto de «causas componentes»[5]. En la figura 14.16 se exponen ejemplos hipotéticos (aunque no ilógicos) de dos «causas suficientes» definidas por Rothman para la enfermedad aterosclerótica.

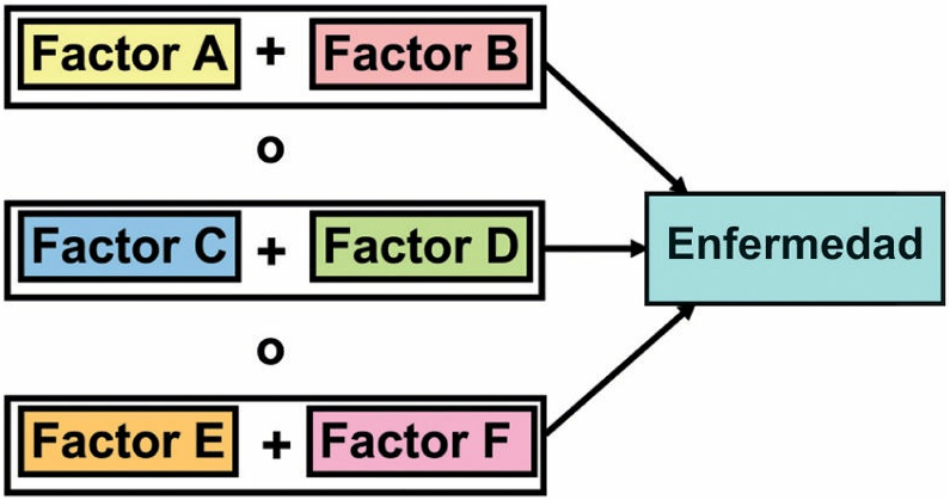


FIG. 14.15 Tipos de relaciones causales: IV. Cada factor no es ni suficiente ni necesario.

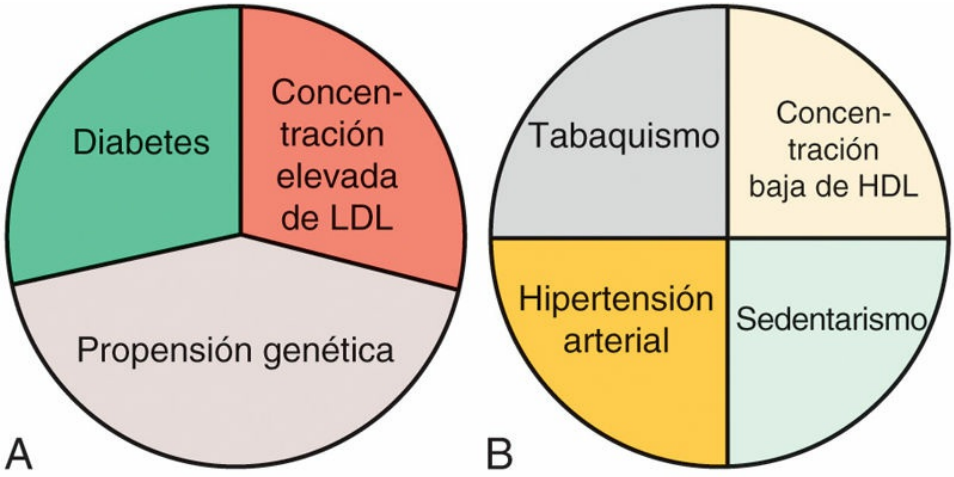


FIG. 14.16 (A-B) Ejemplos hipotéticos de causas suficientes de enfermedad aterosclerótica. *HDL,* lipoproteína de alta densidad; *LDL,* lipoproteína de baja densidad.

Evidencia de una relación causal

Hace muchos años, cuando las principales enfermedades a las que se enfrentaba el ser humano eran de origen fundamentalmente infeccioso, se planteó la cuestión de qué evidencia sería necesaria para demostrar que un microorganismo produce una enfermedad. En 1840, Henle propuso los postulados de causalidad, que fueron aumentados por Koch en 1880[6]. Los postulados de causalidad fueron los siguientes:

1. El microorganismo se encuentra *siempre* en los casos de enfermedad.
2. El microorganismo *no* se encuentra en ninguna otra enfermedad.
3. El microorganismo, cuando se aísla de alguien que tiene la enfermedad y se cultiva a través de varias generaciones, produce la enfermedad (en animales de experimentación).

Koch añadió que «incluso cuando una enfermedad infecciosa no puede contagiarse a los animales, la presencia "habitual" y "exclusiva" del microorganismo [postulados 1 y 2] demuestra una relación causal»[6].

Estos postulados, aunque no son perfectos, resultaron muy útiles para las enfermedades infecciosas. Sin embargo, dado que las enfermedades aparentemente no infecciosas asumieron una importancia creciente hacia mediados del siglo XX, se planteó la cuestión de qué sería una fuerte evidencia de causalidad en enfermedades que por lo general no eran de origen infeccioso. En estas enfermedades no había un microorganismo que pudiese aislarse, cultivarse y desarrollarse en animales. En concreto, cuando la atención se dirigió a una posible relación entre el tabaquismo y el cáncer de pulmón, el Surgeon General de EE. UU. nombró un comité de expertos para revisar la evidencia. El comité elaboró un conjunto de directrices[7], que se han revisado a lo largo de los años. En las páginas siguientes se presenta una lista modificada de esas directrices (cuadro 14.1) con unos breves comentarios.

Cuadro 14.1 Directrices para juzgar si una asociación observada es causal

1. Relación temporal
2. Fuerza de la asociación
3. Relación dosis-respuesta
4. Replicación de los hallazgos
5. Plausibilidad biológica
6. Consideración de explicaciones alternativas
7. Interrupción de la exposición
8. Concordancia con otros conocimientos

9. Especificidad de la asociación

Directrices para juzgar si una asociación observada es causal

1. ***Relación temporal.*** Está claro que, si se cree que un factor es la causa de una enfermedad, la exposición a dicho factor debe haber ocurrido antes de que se desarrolle la enfermedad. En la figura 14.17 se muestra el número de fallecimientos por día y la concentración media de partículas en suspensión en Londres a principios de diciembre de 1952[8]. El patrón de un aumento de la concentración de partículas seguido del aumento de la mortalidad y del descenso posterior de la concentración de partículas seguido de la disminución de la mortalidad apoya firmemente que el incremento de la mortalidad se debía al aumento de la contaminación del aire. Este ejemplo demuestra el uso de datos ecológicos para explorar una relación temporal. Investigaciones posteriores revelaron que el aumento de la mortalidad consistía casi en su totalidad en fallecimientos por enfermedades cardiovasculares y respiratorias, y era máximo en ancianos.

A menudo, es más fácil establecer una relación temporal en un estudio de cohortes prospectivo que en un estudio de casos y controles o en un estudio de cohortes retrospectivo (no concurrente). En los dos últimos tipos de estudio, puede ser necesario obtener o recrear información de la exposición a partir de los registros anteriores, por lo que la secuencia cronológica puede ser imprecisa.

La relación temporal entre la exposición y la enfermedad es importante no solo para aclarar el orden en el que se producen ambas, sino también en lo que respecta a la duración del intervalo entre la exposición y la enfermedad. Por ejemplo, el amianto se ha relacionado claramente con un mayor riesgo de cáncer de pulmón, pero el periodo de latencia entre la exposición y la aparición del cáncer de pulmón es de al menos 15-20 años. Por tanto, si, por ejemplo, el cáncer de pulmón se desarrolla después de tan solo 3 años desde la exposición al amianto, probablemente sea seguro concluir que el cáncer de pulmón no se debió a esta exposición.

2. ***Fuerza de la asociación.*** La fuerza de la asociación se mide por el riesgo relativo (u *odds ratio*). Cuanto más fuerte sea la asociación, más probable será que la relación sea causal. Por ejemplo, el riesgo relativo de la relación entre la presión arterial alta (exposición) y el accidente cerebrovascular (resultado) es muy alto. En un estudio de base poblacional realizado en Suecia, se encontró que el riesgo relativo era mayor de 5,0 en individuos con hipertensión grave[9]. Hay pocas o ninguna duda de que los niveles de presión arterial alta

causan accidentes cerebrovasculares.

3. ***Relación dosis-respuesta.*** A medida que la dosis de exposición aumenta, el riesgo de la enfermedad también lo hace. En la figura 14.18 se muestra un ejemplo de la relación dosis-respuesta para el consumo de cigarrillos y el cáncer de pulmón. Otro ejemplo lo proporciona el estudio sueco mencionado anteriormente, en el cual, al utilizar los niveles normales de presión arterial como categoría de referencia, el riesgo relativo ajustado de accidente cerebrovascular aumentó de manera gradual de 2,84 en individuos con prehipertensión, a 3,90 en aquellos con hipertensión moderadamente grave y a 5,43 en aquellos con niveles compatibles con hipertensión grave. La existencia de una relación dosis-respuesta es una fuerte evidencia de una relación causal. Sin embargo, la ausencia de una relación dosis-respuesta no excluye necesariamente una relación causal. En algunos casos en los que puede existir un umbral, ninguna enfermedad puede desarrollarse hasta un cierto nivel de exposición (un umbral); por encima de este nivel, la enfermedad puede aparecer.
4. ***Replicación de los hallazgos.*** Si la relación es causal, sería de esperar observarla de forma constante en los diferentes estudios y en distintas poblaciones. La replicación de los resultados es especialmente importante en epidemiología. Si se observa una asociación, también sería de esperar que apareciese de forma constante en los subgrupos de la población y en diferentes poblaciones, a menos que exista una razón clara para esperar resultados diferentes.
5. ***Plausibilidad biológica.*** La plausibilidad biológica hace referencia a la coherencia con el corpus actual de los conocimientos biológicos. Se pueden citar varios ejemplos para demostrar que las observaciones epidemiológicas a veces han precedido al conocimiento biológico. Así, como se comenta en un capítulo anterior, las observaciones de Gregg sobre la rubéola y las cataratas congénitas precedieron a cualquier conocimiento de los virus teratogénicos. De forma similar, la implicación de una concentración de oxígeno elevada en la causalidad de la retinopatía de la prematuridad, una forma de ceguera que se produce en los lactantes prematuros, precedió a cualquier conocimiento biológico que apoyase tal relación. Sin embargo, hay que buscar la coherencia entre los hallazgos epidemiológicos y el conocimiento biológico existente; cuando esto no sucede, puede ser difícil interpretar el significado de la asociación observada. En tal caso, se puede ser más exigente en cuanto a los requisitos sobre el tamaño y la importancia de las diferencias observadas y en la replicación del estudio por otros investigadores en otras poblaciones.
6. ***Consideración de explicaciones alternativas.*** Ya se ha comentado el problema existente al interpretar una asociación observada en cuanto a si una relación es causal o si es el resultado de un factor de

confusión. A la hora de juzgar si una asociación descrita es causal, el grado en el que los investigadores han tenido en cuenta otras explicaciones posibles y el grado en el que se han descartado esas explicaciones son consideraciones importantes.

7. ***Interrupción de la exposición.*** Si un factor es una causa de una enfermedad, sería de esperar que el riesgo de la enfermedad declinase cuando la exposición al factor se reduce o se elimina. En la figura 14.19 se muestran estos datos históricos para el tabaquismo y el cáncer de pulmón. Otro ejemplo fue el rápido descenso de las tasas de AC, diabetes y accidentes cerebrovasculares después de la disminución drástica de la ingesta calórica y, por lo tanto, de la obesidad, debido a la crisis económica cubana de 1989-2000[10].

El síndrome eosinofilia-mialgia (SEM) alcanzó proporciones epidémicas en 1989. Este síndrome se caracteriza por dolor muscular intenso y un recuento sanguíneo elevado de eosinófilos, y se ha observado su asociación con preparados fabricados de L-triptófano. En noviembre de 1989, la retirada nacional por la Food and Drug Administration de EE. UU. de las preparaciones sin receta de L-triptófano se siguió de una reducción drástica del número de casos de SEM notificados cada mes (fig. 14.20). Éste es otro ejemplo de una reducción de la incidencia que se relaciona con la interrupción de la exposición, lo que se suma a la fuerza de la inferencia causal respecto a la exposición.

Cuando se dispone de datos de la interrupción, estos proporcionan una evidencia útil en apoyo de una asociación causal. Sin embargo, en ciertos casos, el proceso patogénico puede haber sido iniciado de forma irreversible, y la aparición de la enfermedad puede que ya esté determinada cuando se elimina la exposición. El enfisema no se revierte al dejar de fumar, pero su progresión se reduce.

8. ***Concordancia con otros conocimientos.*** Si una relación es causal, sería de esperar que los resultados fuesen concordantes con otros datos. Por ejemplo, en la figura 14.21 se muestran los datos relativos a las tasas de cáncer de pulmón en varones y mujeres y el tabaquismo en ambos sexos.

Se puede observar una dirección concordante en las curvas, con un aumento de las tasas de cáncer de pulmón tras el incremento de las ventas de cigarrillos, tanto en varones como en mujeres. Estos datos concuerdan con lo que cabría esperar si la relación entre el tabaquismo y el cáncer de pulmón se estableciese como causal. Aunque la ausencia de dicha concordancia no descartaría por completo esta hipótesis, si se observase el aumento de las tasas de cáncer de pulmón después de un periodo de disminución de las ventas de cigarrillos, por ejemplo, habría que explicar cómo esta observación podría concordar con una hipótesis causal.

9. ***Especificidad de la asociación.*** Una asociación es específica cuando una cierta exposición se asocia con una sola enfermedad; esta es la

más débil de todas las directrices y probablemente debería suprimirse de la lista. Los fabricantes de cigarrillos han señalado que las enfermedades atribuidas al tabaquismo no cumplen con los requisitos de esta directriz, ya que el consumo de cigarrillos se ha relacionado con los cánceres de pulmón, páncreas y vejiga, cardiopatías, enfisema y otras afecciones.

La posibilidad de esta multiplicidad de efectos debida a un solo factor no es en realidad sorprendente: con independencia del tejido donde se encuentren, todas las células tienen características comunes, como el ADN, el ARN y las diferentes estructuras subcelulares, por lo que un único agente podría tener efectos en múltiples tejidos. Además, los cigarrillos no son un único factor, sino que constituyen una mezcla de un gran número de compuestos y, en consecuencia, sería previsible un gran número de efectos.

Cuando se observa la especificidad de una asociación, proporciona apoyo adicional para una inferencia causal. Sin embargo, como sucede con la relación dosis-respuesta, la falta de especificidad no excluye en modo alguno una relación causal.

Cualquier conclusión de que una asociación observada es causal se ve reforzada en gran medida cuando diferentes tipos de evidencia procedentes de múltiples fuentes apoyan este razonamiento. Así pues, no es tanto el recuento del número de directrices presente lo que es relevante para la inferencia causal, sino más bien la *evaluación del patrón total de la evidencia observada* que pueda concordar con una o más de las directrices. Sir Austin Bradford Hill expresó con elocuencia esta observación en un ensayo escrito en 1965:

He aquí nueve puntos de vista [directrices] diferentes a partir de los que debería estudiarse la existencia de una asociación antes de afirmar que existe una causalidad. De todos modos, yo no creo (y esto se ha sugerido) que resulte útil establecer una serie de reglas rígidas sobre la evidencia que deban cumplirse antes de poder aceptar la relación de causa y efecto. Ninguno de mis nueve puntos de vista puede aportar pruebas irrefutables a favor o en contra de la hipótesis de causa y efecto, y ninguno puede ser exigido como una condición sine qua non. Para lo que sirven, con mayor o menor fuerza, es para ayudarnos a decidir sobre la pregunta fundamental: ¿hay alguna otra manera de explicar el conjunto de hechos que tenemos ante nosotros, hay alguna otra respuesta igual de probable, o más, que la relación de causa y efecto?[11]

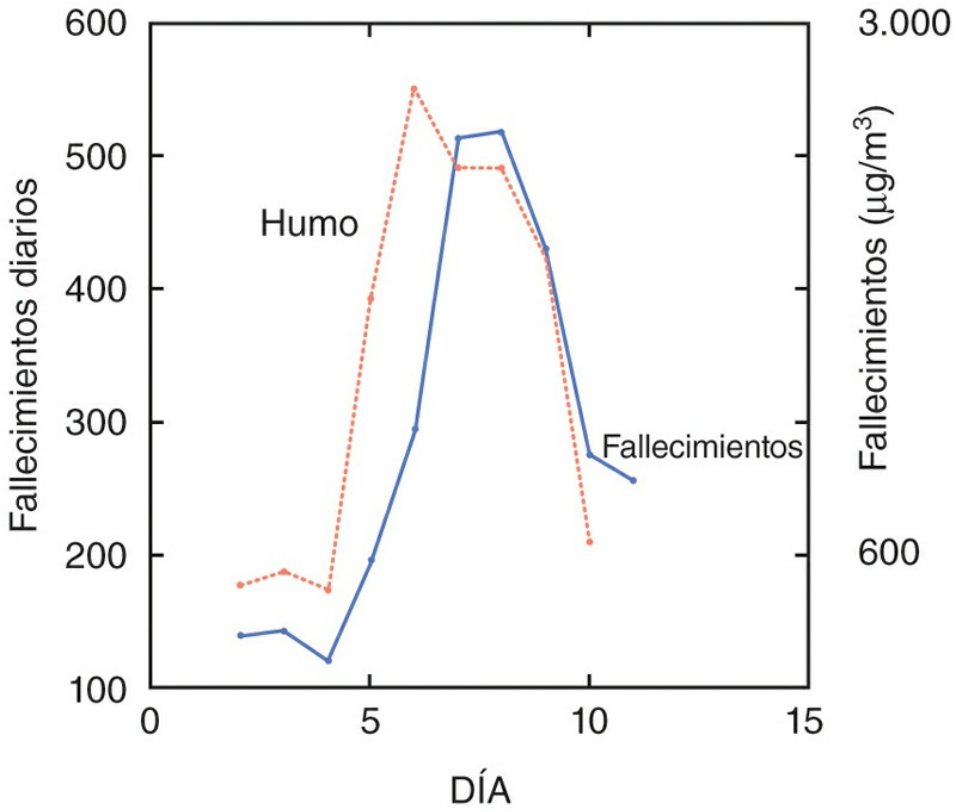


FIG. 14.17 Concentración media de partículas en suspensión (µg/m^3) de las cuatro estaciones interiores de monitorización de Londres y recuento de los fallecimientos diarios en el Condado Administrativo de Londres durante principios de diciembre de 1952. (De Schwartz J: Air pollution and daily mortality: A review and metaanalysis. *Environ Res*. 1994;64:36–52.)

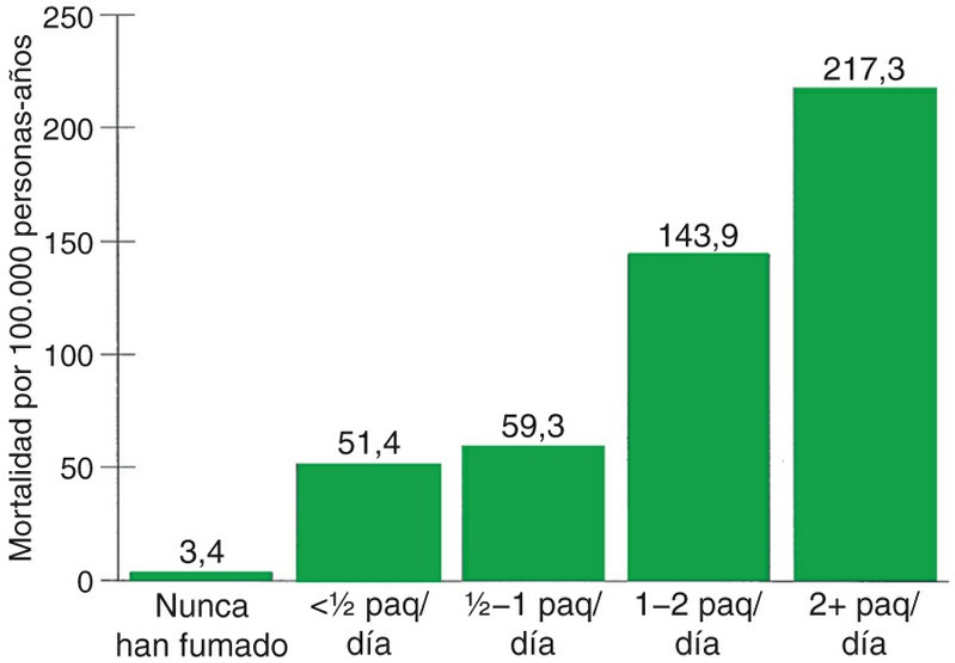


FIG. 14.18 Mortalidad estandarizada por edad debida a casos bien establecidos de carcinoma broncogénico (sin incluir el adenocarcinoma) por grado de tabaquismo actual. (Modificada de Hammond EC, Horn D. Smoking and death

rates: Report on 44 months of follow-up of 187,783 men: II. Death rates by cause. *JAMA*. 1958;166:1294–1508. Copyright 1958, American Medical Association.)

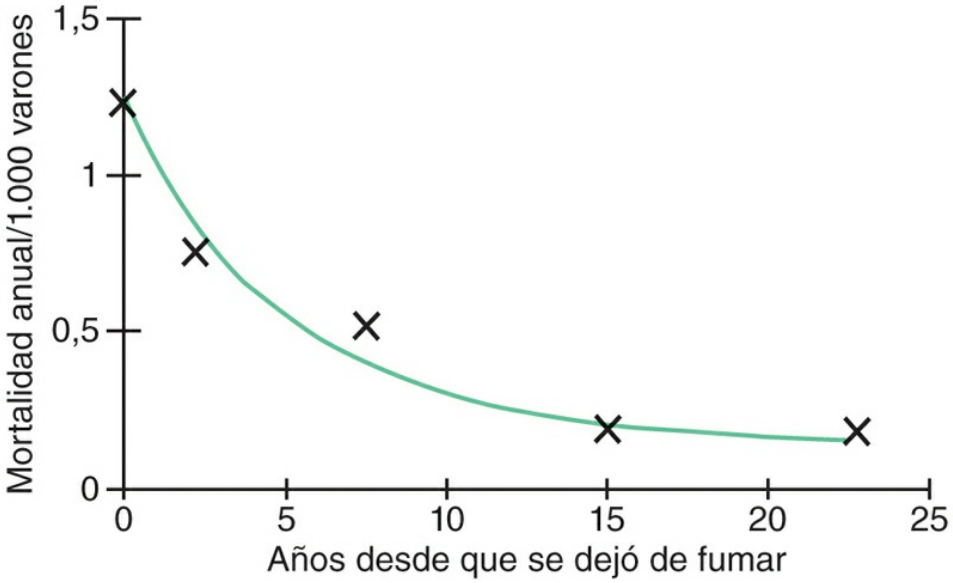


FIG. 14.19 Efectos del fin de la exposición: mortalidad por cáncer de pulmón, estandarizada por edad y cantidad fumada, entre los varones que continúan fumando cigarrillos y que dejan de fumar durante diferentes periodos. La tasa correspondiente a los no fumadores fue de 0,07 por 1.000. (Modificada de Doll R, Hill AB. Mortality in relation to smoking: Ten years' observations of British doctors. *BMJ*. 1964;1:1399–1410.)

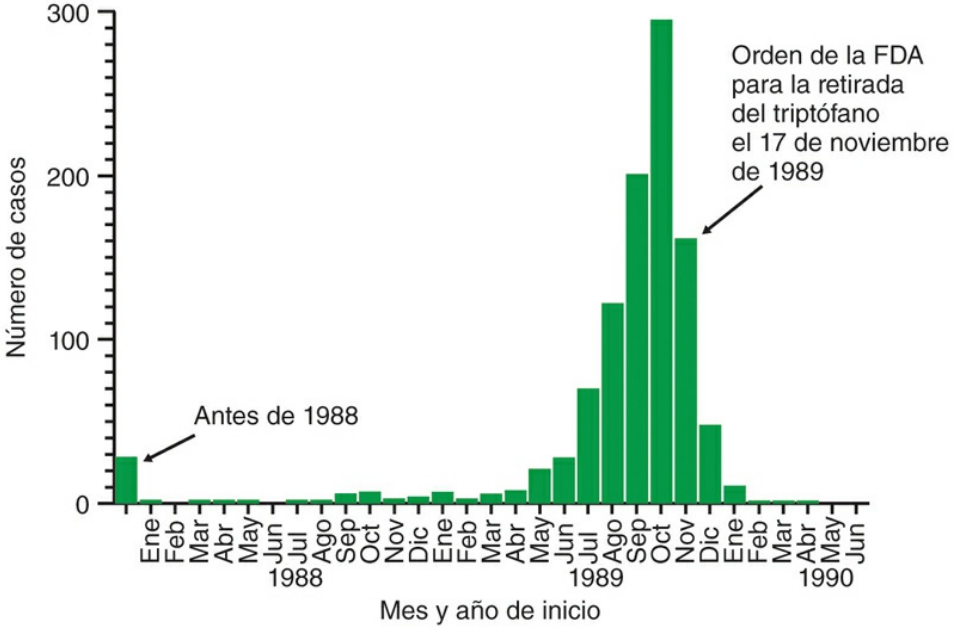


FIG. 14.20 Fechas notificadas de inicio de enfermedad por meses y años para los casos de síndrome de eosinofilia-mialgia, según se notificaron a los Centros para el Control y la Prevención de Enfermedades, Atlanta, hasta el 10 de julio de 1990. (Modificada de Swygert LA, Maes EF, Sewell LE, et al. Eosinophilia-myalgia

syndrome: Results of national surveillance. *JAMA*. 1990;264:1698–1703. Copyright 1990, American Medical Association.)

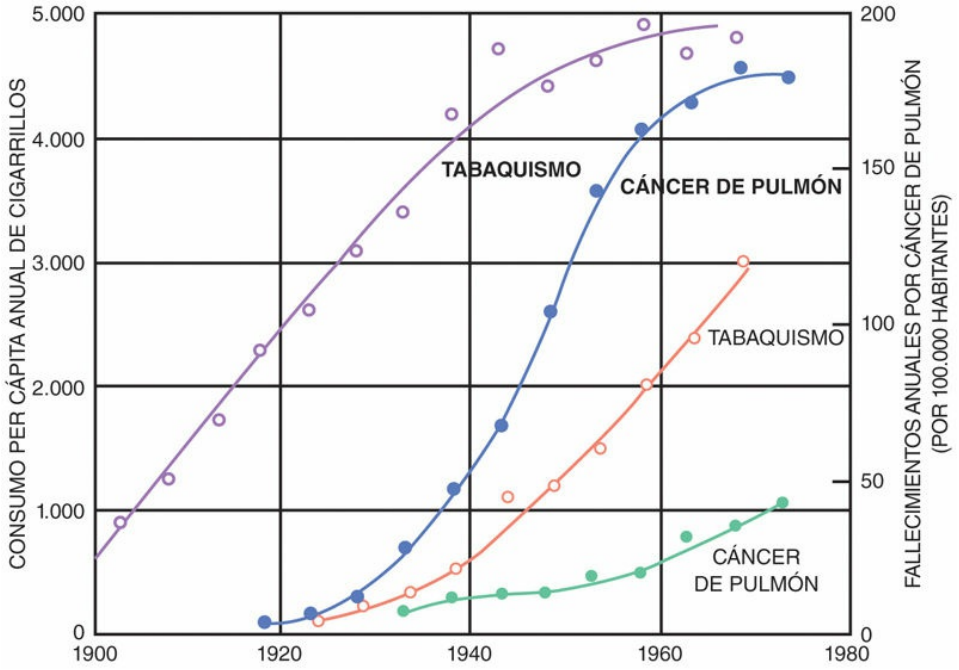


FIG. 14.21 Tendencias paralelas entre el tabaquismo y el cáncer de pulmón en varones *(dos curvas de la izquierda)* y en mujeres (dos curvas de la derecha), en Inglaterra y Gales. (De Cairns J. The cancer problem. *Sci Am*. 1975;233:64–72, 77–78.)

Derivación de inferencias causales: dos ejemplos

Úlcera péptica y cáncer gástrico en relación con la infección por *Helicobacter pylori*

Aunque las directrices anteriores no permiten una estimación cuantitativa de si una asociación es o no causal, sí que pueden ser muy útiles, como se verá en los siguientes ejemplos:

Hasta la década de 1980, se consideraba que las principales causas de la enfermedad ulcerosa péptica eran el estrés y los factores del estilo de vida, como el tabaquismo. La enfermedad ulcerosa péptica se había atribuido durante mucho tiempo a los efectos del ácido gástrico. La susceptibilidad a los ácidos gástricos se había relacionado con el tabaquismo, el consumo de alcohol y el uso de antiinflamatorios no esteroideos. El tratamiento se dirigía principalmente a inhibir la secreción de ácido y a proteger las superficies mucosas contra el ácido. Aunque estos tratamientos ayudaban a la curación, las recidivas eran frecuentes.

En 1984, los médicos australianos Barry J. Marshall y J. Robin Warren publicaron su observación de pequeñas bacterias curvadas que colonizaban la parte inferior del estómago en pacientes con gastritis y úlceras pépticas[12]. Después de varios intentos, Marshall logró cultivar una especie bacteriana hasta entonces desconocida (después denominada *H. pylori)* a partir de varias de estas biopsias (fig. 14.22). Juntos, observaron que el microorganismo estaba presente en casi todos los pacientes con inflamación gástrica o úlcera péptica. En muchos de estos pacientes se realizaron después biopsias que mostraron signos de inflamación en la mucosa gástrica cerca de donde se observaron las bacterias. Basándose en estos resultados, propusieron que *H. pylori* estaba implicada en la etiología de estas enfermedades. Posteriormente se demostró que la úlcera a menudo no se curaba hasta la erradicación de *H. pylori*.



FIG. 14.22 *Helicobacter pylori*. (Encyclopædia Britannica Online. http://www.britannica.com/EBchecked/topic/450889/peptic-ulcer?overlay=true&assemblyId=94921. Consultado en noviembre de 2017.)

En la actualidad, está firmemente establecido que *H. pylori* causa más del 90% de las úlceras duodenales y hasta el 80% de las úlceras gástricas. La relación entre la infección por *H. pylori* y la gastritis y la úlcera péptica subsiguientes se ha establecido mediante estudios de voluntarios humanos, estudios de tratamiento con antibióticos y estudios epidemiológicos. Por tanto, muchos de los diseños de estudio analizados en los capítulos anteriores

y muchas de las directrices para las inferencias causales descritas con anterioridad en este capítulo estuvieron implicados en el esclarecimiento del papel de *H. pylori* en la úlcera péptica y la gastritis. En 2005, los Dres. Marshall y Warren compartieron el Premio Nobel de Fisiología o Medicina «por su descubrimiento de la bacteria *H. pylori* y su papel en la gastritis y la úlcera péptica».

En el cuadro 14.2 se clasifica esta evidencia según varias de las directrices para la relación de causalidad que se acaban de exponer. Como se ve aquí, las directrices pueden ser extremadamente útiles a la hora de caracterizar la evidencia que apoya una relación causal.

Cuadro 14.2 Evaluación de la evidencia sugestiva de que *Helicobacter pylori* es un agente causal de úlceras duodenales

1. Relación temporal:
    - *Helicobacter pylori* está claramente relacionado con la gastritis crónica. Alrededor del 11% de los pacientes con gastritis crónica desarrollarán úlceras duodenales en un periodo de 10 años.
    - En un estudio de 454 pacientes que se sometieron a una endoscopia hace 10 años, 34 de 321 pacientes que habían sido positivos para *Helicobacter pylori* (11%) tenían úlcera duodenal, en comparación con 1 de 133 pacientes negativos para *Helicobacter pylori* (0,8%).
2. Fuerza de la asociación:
    - *Helicobacter pylori* se encuentra en al menos el 90% de los pacientes con úlcera duodenal. En al menos una población en la que se ha descrito la ausencia de úlceras duodenales (una tribu aborigen del norte de Australia que está aislada de otras personas) nunca se ha encontrado.
3. Relación dosis-respuesta:
    - La densidad de *Helicobacter pylori* por milímetro cuadrado de mucosa gástrica es mayor en pacientes con úlcera duodenal que en aquellos sin úlcera duodenal. Véase también el punto 2.
4. Replicación de los hallazgos:
    - Muchas de las observaciones relativas a *Helicobacter pylori* se han replicado en varias ocasiones.
5. Plausibilidad biológica:
    - Aunque en un principio era difícil imaginar que una bacteria que infectase el antro gástrico causase úlceras en el duodeno, ahora se acepta que *Helicobacter pylori* presenta sitios de unión en las células antrales y puede seguir a estas células

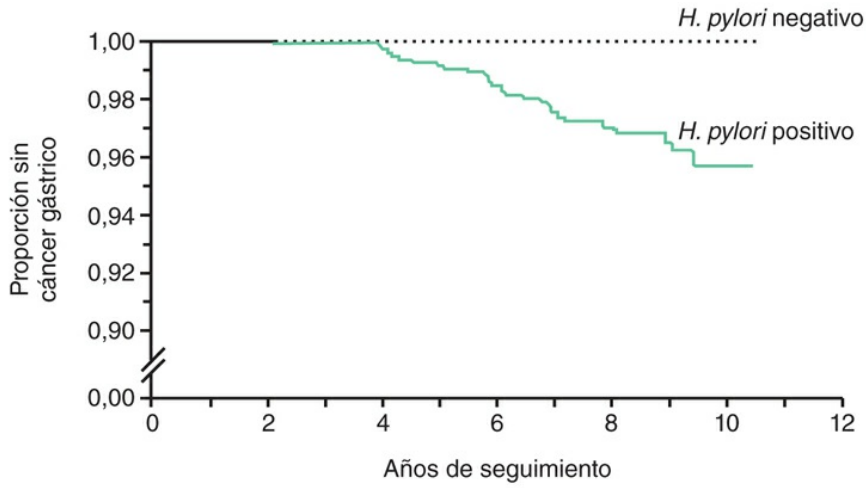
hasta el duodeno.
- *Helicobacter pylori* también induce mediadores de la inflamación.
- La mucosa infectada por *Helicobacter pylori* se debilita y es susceptible a los efectos lesivos del ácido.

6. Consideración de explicaciones alternativas:
   - Los datos sugieren que fumar puede aumentar el riesgo de úlcera duodenal en pacientes infectados por *Helicobacter pylori,* pero no es un factor de riesgo en pacientes en los que se ha erradicado esta bacteria.
7. Interrupción de la exposición:
   - La tasa de curación de las úlceras duodenales es igual con la erradicación de *Helicobacter pylori* que con el uso de antagonistas del receptor de la histamina.
   - Las tasas de recidiva a largo plazo de la úlcera eran nulas después de la erradicación de *Helicobacter pylori* con el uso de la triple terapia antimicrobiana, en comparación con una tasa de recidiva del 60-80% que suele observarse en los pacientes con úlceras duodenales tratados con antagonistas del receptor de la histamina.
8. Concordancia con otros conocimientos:
   - La prevalencia de la infección por *Helicobacter pylori* es igual en varones y mujeres. La incidencia de la úlcera duodenal, que previamente se consideraba mayor en los varones que en las mujeres, ha sido igual en los últimos años.
   - Se cree que la prevalencia de la enfermedad ulcerosa alcanzó su máximo en la última parte del siglo XIX, y la prevalencia de *Helicobacter pylori* puede haber sido mucho mayor en esa época debido a las malas condiciones de vida. Este razonamiento también se basa en observaciones actuales de que la prevalencia de *Helicobacter pylori* es mucho mayor en los países en vías de desarrollo.
9. Especificidad de la asociación:
   - La prevalencia de *Helicobacter pylori* en pacientes con úlceras duodenales es del 90-100%. Sin embargo, se ha encontrado en algunos pacientes con úlcera gástrica e incluso en personas asintomáticas.

Datos de Megraud F, Lamouliatte H. *Helicobacter pylori* and duodenal ulcer: Evidence suggesting causation. *Dig Dis Sci.* 1992;37:769–772; y DeCross AJ, Marshall BJ: The role of *Helicobacter pylori* in acid-peptic disease. *Am J Med Sci.* 1993;306:381–392.

En la actualidad, cada vez es mayor la evidencia que respalda la asociación de la infección por *H. pylori* y el desarrollo de cáncer gástrico. Uemura y

cols.[13] realizaron un estudio prospectivo con 1.526 pacientes japoneses que tenían úlceras duodenales o gástricas, hiperplasia gástrica o hiperplasia no ulcerosa. De este grupo, 1.246 personas tenían una infección *por H. pylori* y 280 no. El periodo medio de seguimiento fue de 7,8 años. Treinta y seis (2,9%) de los pacientes infectados desarrollaron un cáncer gástrico, frente a ninguno de los no infectados. Las personas que tienen anticuerpos contra *H. pylori* pueden tener un riesgo 2-3 veces mayor de cáncer gástrico que quienes carecen de ellos (fig. 14.23). El riesgo de cáncer gástrico también parece estar relacionado con el tipo de cepa de *H. pylori* causante de la infección. Cada vez hay más evidencia que apoya la idea de que el tratamiento contra *H. pylori* puede prevenir el cáncer gástrico. En el futuro, puede que el cáncer gástrico se considere una neoplasia de origen infeccioso prevenible en gran medida.



N.º en situación de riesgo						
H. pylori negativo	280	272	251	245	213	57
H. pylori positivo	1.246	1.219	1.086	907	782	258

FIG. 14.23 Análisis de Kaplan-Meier de la proporción de pacientes positivos y negativos para *Helicobacter pylori* que no presentaban cáncer gástrico. Durante el seguimiento se desarrolló un cáncer gástrico en 36 de los 1.246 pacientes infectados por *H. pylori* (2,9%), pero en ninguno de los 280 pacientes no infectados (P <0,001). (De Uemura N, Okamoto S, Yamomoto S, et al. *Helicobacter pylori* infection and the development of gastric cancer. *N Engl J Med.* 2001,345:784–789.)

Edad de inicio del consumo de alcohol y consumo de alcohol a lo largo de la vida

En 1997, Grant y Dawson[14] publicaron los datos sobre la relación de la edad en el momento en que se consumió alcohol por primera vez y la prevalencia de la dependencia y consumo de alcohol a lo largo de la vida. Estos autores analizaron los datos de 27.616 bebedores activos y previos que fueron entrevistados como parte del National Longitudinal Alcohol Epidemiologic

Survey de 1992. Las tasas de *dependencia* durante la vida disminuían desde más del 40% en las personas que comenzaron a beber a los 14 años o menos a alrededor del 10% en quienes empezaron a beber a los 20 años o más (fig. 14.24). La configuración de la curva de la figura 14.24 sugiere una relación dosis-respuesta, como la que se ha observado entre la duración más prolongada del tabaquismo y el mayor riesgo de cáncer de pulmón. Sin embargo, los datos también pueden indicar la existencia de un periodo de una susceptibilidad especialmente elevada, es decir, que el periodo de la preadolescencia y adolescencia temprana sería un periodo de mayor riesgo de desarrollar un trastorno de alcoholismo. Por tanto, las intervenciones preventivas deben dirigirse a este grupo con la esperanza de retrasar el inicio del consumo de alcohol. Sin embargo, la adopción de este planteamiento presupone que la relación entre la aparición temprana del consumo de alcohol y el consumo posterior a lo largo de la vida es de tipo causal, por lo que el retraso de la edad de inicio del consumo de alcohol reduciría el riesgo de dependencia del alcohol a lo largo de la vida. Otra explicación posible es que quienes están destinados a presentar una dependencia del alcohol a lo largo de la vida tienden a comenzar a beber antes, pero que la edad más temprana de inicio del consumo de alcohol no es necesariamente una causa de una dependencia subsiguiente. Se necesitan, por tanto, más investigaciones para explicar esta asociación enigmática que se ha observado. Más adelante se volverá sobre este ejemplo en el capítulo 16.

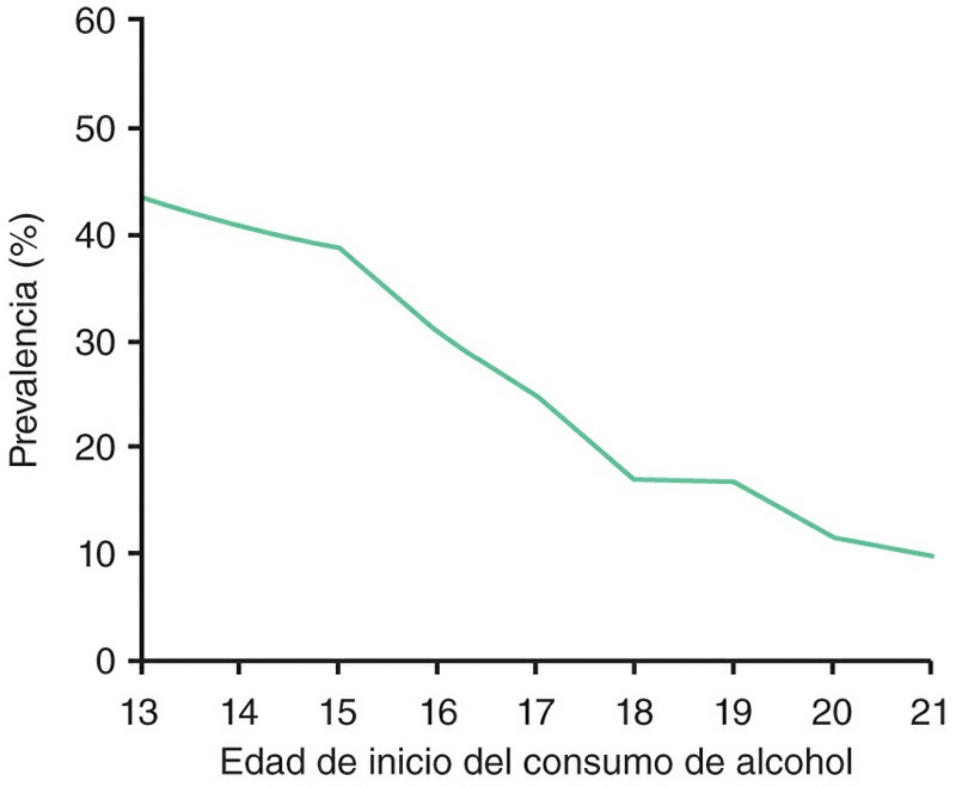


FIG. 14.24 Relación de la edad de inicio del consumo de alcohol con la prevalencia del alcoholismo durante toda la vida. (Modificada de Grant BF, Dawson DA. Age at onset of alcohol use and its association with DSM-IV alcohol

abuse and dependence: Results from the National Longitudinal Alcohol Epidemiologic Survey. *J Subst Abuse.* 1997;9:103–110.)

Modificaciones de las directrices para las inferencias de causalidad

En 1986, el Servicio de Salud Pública de EE. UU. (U.S. Public Health Service) reunió a un grupo de 19 expertos para analizar la base científica del contenido de la asistencia prenatal y para responder a esta pregunta: «¿Qué medidas aplicadas durante la asistencia prenatal han demostrado que se asocian en realidad con un mejor resultado?» El informe de este grupo de expertos se publicó en 1989 y sirvió como base para un informe exhaustivo[15]. Cuando el grupo comenzó sus deliberaciones, se hizo evidente que las cuestiones de la causalidad ocupaban el núcleo del trabajo de los expertos y que se necesitaban directrices para evaluar la relación entre las medidas prenatales y los resultados de salud. Un subcomité revisó las directrices actuales (que se acaban de enumerar en el apartado anterior) y definió un procedimiento para utilizar la evidencia, que incluía 1) la clasificación de la evidencia según la calidad de sus fuentes y 2) la evaluación de la evidencia de una relación causal mediante directrices estandarizadas[16]. Estas recomendaciones se presentan en el cuadro 14.3. Aunque estas directrices modificadas utilizan claramente los componentes originales, establecen prioridades razonables al ponderarlas. Definen, por tanto, un planteamiento para considerar la relación causal que puede ser aplicable mucho más allá de las preguntas sobre la eficacia de las medidas prenatales.

Cuadro 14.3 Proceso para usar la evidencia en la elaboración de recomendaciones sobre la eficacia de las intervenciones prenatales

Etapa I. Clasificación de la evidencia según la calidad de su fuente. ***(En cada categoría, los estudios se enumeran por orden decreciente de calidad.)***

1. Ensayos clínicos (intervenciones planificadas con asignación simultánea de tratamiento y no tratamiento):
    a. Aleatorizado, doble ciego, controlado con placebo, con suficiente potencia y analizado adecuadamente.
    b. Aleatorizado, pero no doble ciego.
    c. Ensayos no aleatorizados con un buen control de los factores de confusión y bien realizados en otros aspectos.
    d. Aleatorizados, pero con deficiencias en la ejecución o el análisis (potencia insuficiente, abandono significativo durante el seguimiento, aleatorización sospechosa, análisis con exclusiones).
    e. Ensayos no aleatorizados con deficiencias en la ejecución o el análisis.

2. Estudios de cohortes o de casos y controles:
    a. Hipótesis especificada antes del análisis, datos adecuados, factores de confusión tenidos en cuenta.
    b. Como el anterior, pero hipótesis no especificada antes del análisis.
    c. A posteriori, con problemas en los datos o el análisis.
3. Estudios de series temporales:
    a. Análisis que tienen en cuenta los factores de confusión.
    b. Análisis que no tienen en cuenta los factores de confusión.
4. Estudios de series de casos: series de casos sin grupo de comparación específico

Entre otros problemas que deben tenerse en cuenta al revisar la evidencia, hay que citar la precisión de la definición de los resultados que se están midiendo, el grado en que se ha descrito la metodología del estudio, la idoneidad del tamaño muestral y el grado en que se han descrito las características de la población estudiada y de la intervención que se está evaluando.

Un estudio puede estar bien diseñado y llevado a cabo de una manera ejemplar (validez interna), pero, si la población estudiada es inusual o está muy seleccionada, puede que los resultados no sean generalizables (validez externa).

Etapa II. Directrices para la evaluación de la evidencia de una relación causal. *(En cada categoría, los estudios se enumeran por orden decreciente de prioridad.)*

1. Criterios mayores:
    a. Relación temporal: una intervención puede considerarse como evidencia de una reducción del riesgo de enfermedad o anomalía solo si la intervención se aplicó antes del momento en el que la enfermedad o anomalía se desarrolló.
    b. Plausibilidad biológica: un mecanismo biológicamente plausible debe poder explicar por qué sería previsible que se produjese esta relación.
    c. Concordancia: los estudios individuales pocas veces son definitivos. Los resultados del estudio que se replican en diferentes poblaciones y por distintos investigadores tienen más peso que los que no se reproducen. Si los resultados de los estudios no son concordantes, se debe explicar la discordancia.
    d. Explicaciones alternativas (factor de confusión): el grado en que se han explorado las explicaciones alternativas es un criterio importante para juzgar la causalidad.
2. Otras consideraciones:
    a. Relación dosis-respuesta: si un factor es en realidad la causa

de una enfermedad, por lo general (pero no siempre) cuanto mayor sea la exposición al factor, mayor será el riesgo de la enfermedad. Puede que este tipo de relación dosis-respuesta no siempre se observe, porque muchas relaciones biológicas importantes son dicotómicas y alcanzan un nivel umbral para los efectos observados.

b. Fuerza de la asociación: la fuerza de la asociación suele medirse por el grado en que el riesgo relativo o las posibilidades se distancian de la unidad, ya sea por encima de 1 (en el caso de exposiciones causantes de enfermedad) o por debajo de 1 (en el caso de intervenciones preventivas).

c. Efectos de la interrupción: si una intervención tiene un efecto beneficioso, el beneficio debe interrumpirse cuando se elimina de una población (a menos que intervenga un efecto de arrastre).

Modificado de Gordis L, Kleinman JC, Klerman LV, et al. Criteria for evaluating evidence regarding the effectiveness of prenatal interventions. En Merkatz IR, Thompson JE, eds. New Perspectives on Prenatal Care. Nueva York: Elsevier; 1990:31-38.

El Grupo de Trabajo de los Servicios Preventivos de EE. UU. (U.S. Preventive Services Task Force, USPSTF), que es responsable del desarrollo de directrices de práctica clínica para la prevención y el cribado, utiliza un planteamiento similar que clasifica los estudios en función de su calidad y de su evidencia (tabla 14.1)[17]. Se trata de un comité independiente de expertos respaldado por el gobierno estadounidense. Sus miembros son expertos en atención primaria, prevención, medicina basada en la evidencia y métodos de investigación. También están representadas diversas áreas clínicas y expertos en medicina preventiva, salud pública y política sanitaria.

Tabla 14.1

Niveles de certeza[a] del Grupo de Trabajo de los Servicios Preventivos de EE. UU. con respecto al beneficio neto

Alto	La evidencia disponible suele incluir resultados concordantes de estudios bien diseñados y bien realizados en poblaciones representativas de atención primaria. Estos estudios evalúan los efectos del servicio de prevención sobre los resultados de salud. Por tanto, es poco probable que esta conclusión se vea fuertemente afectada por los resultados de futuros estudios.
Moderado	La evidencia disponible es suficiente para determinar los efectos del servicio preventivo sobre los resultados de salud, pero la confianza en la estimación se ve limitada por los siguientes factores: • El número, el tamaño o la calidad de los estudios individuales. • La discordancia de los resultados de los estudios individuales. • Posibilidad limitada de generalizar los resultados a la práctica habitual de atención primaria. • Falta de coherencia en la cadena de evidencia. A medida que se dispone de más información, la magnitud o la dirección del efecto

	observado podría cambiar, y este cambio puede ser lo bastante intenso como para modificar la conclusión.
Bajo	La evidencia disponible es insuficiente para evaluar los efectos sobre los resultados de salud. La evidencia es insuficiente debido a: • El número o tamaño limitado de los estudios. • Defectos importantes en el diseño o los métodos del estudio. • Discordancia de los hallazgos de estudios individuales. • Carencias en la cadena de evidencia. • Resultados no generalizables a la práctica habitual de atención primaria. • Falta de información sobre los resultados importantes de salud. Más información puede permitir una estimación de los efectos sobre los resultados de salud.

[a] El US Preventive Services Task Force *(USPSTF)* define certeza como «la probabilidad de que la evaluación por el USPSTF del beneficio neto de un servicio preventivo sea correcta». El beneficio neto se define como el beneficio menos el perjuicio del servicio preventivo tal como se aplica en una población general de atención primaria. El USPSTF asigna un nivel de certeza basado en las características de la evidencia total disponible para evaluar el beneficio neto de un servicio preventivo.

Modificada del US Preventive Services Task Force Procedure Manual. Diciembre de 2015. http://www.uspreventiveservicestaskforce.org/Home/GetFile/6/7/procedure-manual_2015/pdf. Consultado el 25 de octubre de 2017.

Para cada tema que analiza el Grupo de Trabajo, define las cuestiones que hay que abordar e identifica y recopila la evidencia relevante. Después de evaluar la calidad de cada estudio individual, se valora la solidez de toda la evidencia disponible y se realizan estimaciones del balance entre beneficio y perjuicio. Este balance se expresa como el beneficio neto (la diferencia entre beneficios y perjuicios). El Grupo de Trabajo elabora recomendaciones de intervenciones preventivas que se basan en estos análisis.

En la figura 14.25 se muestra un ejemplo genérico del plan analítico elaborado por el Grupo de Trabajo como marco para evaluar la evidencia de un programa de cribado. Las flechas rectas indican posibles vías de beneficio y las flechas curvas azules muestran los posibles efectos adversos relacionados con las diferentes etapas. La pregunta principal (pregunta 1 en la figura) suele ser si el cribado es eficaz para reducir el riesgo de un resultado adverso, como la mortalidad y, en caso afirmativo, en qué medida.

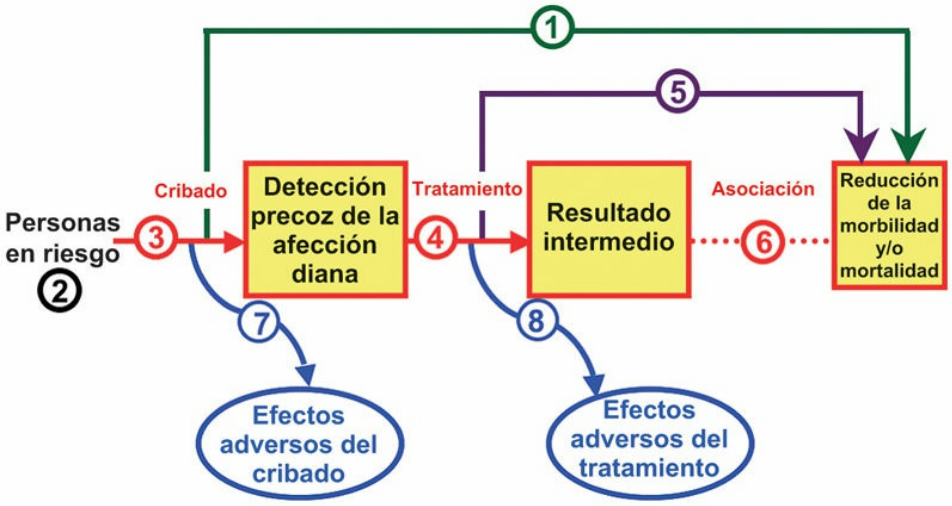


FIG. 14.25 Marco analítico genérico para los temas de cribado utilizado por el Grupo de Trabajo de los Servicios Preventivos de EE. UU. Los números hacen referencia a las cuestiones clave de la figura. (1) ¿El cribado de X reduce la morbilidad y/o mortalidad? (2) ¿Puede identificarse un grupo de alto riesgo de X basándose en criterios clínicos? (3) ¿Se dispone de pruebas de cribado precisas (es decir, sensibles y específicas)? (4) ¿Se dispone de tratamientos que modifiquen los resultados intermedios cuando la enfermedad se detecta a tiempo, o se detecta mediante cribado? (5) ¿Se dispone de tratamientos que modifiquen la morbilidad o la mortalidad cuando la enfermedad se detecta a tiempo, o se detecta mediante cribado? (6) ¿Qué solidez tiene la asociación entre los resultados intermedios y los resultados de los pacientes? (7) ¿Qué perjuicios conlleva la prueba de cribado? (8) ¿Qué perjuicios conlleva el tratamiento? (Modificada del U.S. Preventive Services Task Force Procedure Manual. Diciembre de 2015. http://www.uspreventiveservicestaskforce.org/Home/GetFile/6/7/procedure-manual_2015/pdf. Consultado el 25 de octubre de 2017.)

Las deliberaciones del Grupo de Trabajo suelen abordar los diferentes pasos o vínculos que componen esta vía general. La flecha de color púrpura de la figura (paso 5) muestra la relación del tratamiento con el resultado. Las flechas rojas (pasos 3, 4 y 6) muestran los componentes individuales de la pregunta 1. Estas evaluaciones suelen depender de una revisión de los ensayos aleatorizados relevantes para preparar una cadena de evidencia probatoria en la que basar la respuesta a la pregunta 1. La evidencia para cada vínculo se resume en la revisión de la evidencia y después se resume a través de los diferentes vínculos para proporcionar una evaluación general de la evidencia probatoria para el servicio preventivo que se está evaluando.

La certeza del beneficio neto se califica según una escala de tres puntos: alta, moderada o baja (v. tabla 14.1). Las recomendaciones del Grupo de Trabajo se basan en la consideración conjunta de la certeza y de la magnitud del beneficio neto, como se muestra en la matriz de la figura 14.26, en la que se utiliza un sistema de calificación de A, B, C, D e *Insuficiente*. El significado de la calificación de cada grado se explica en la tabla 14.2.

Certeza del beneficio neto	Magnitud del beneficio neto			
	Sustancial	Moderada	Pequeña	ula/ negativa
Alta	A	B	C	D
Moderada	B	B	C	D
Baja	Evidencia insuficiente			

FIG. 14.26 Tabla utilizada por el Grupo de Trabajo de los Servicios Preventivos de EE. UU. para evaluar la certeza del beneficio y la magnitud del beneficio neto para determinar el grado de sus recomendaciones. (Modificada del U.S. Preventive Services Task Force Procedure Manual. Diciembre de 2015. http://www.uspreventiveservicestaskforce.org/Home/GetFile/6/7/procedure-manual_2015/pdf. Consultado el 25 de octubre de 2017.)

Tabla 14.2

Significado de los grados del Grupo de Trabajo de los Servicios Preventivos de EE. UU. (USPSTF) y sugerencias para la práctica

Grado	Definiciones de los grados	Sugerencias para la práctica
A	El USPSTF recomienda el servicio. Hay una alta certeza de que el beneficio neto es sustancial.	Ofrecer/proporcionar este servicio.
B	El USPSTF recomienda el servicio. Hay una alta certeza de que el beneficio neto es moderado o hay una certeza moderada de que el beneficio neto es moderado o sustancial.	Ofrecer/proporcionar este servicio.
C	*Nota: La siguiente declaración está en revisión.* Los médicos pueden proporcionar este servicio a pacientes seleccionados dependiendo de circunstancias individuales. Sin embargo, para la mayoría de los individuos sin signos o síntomas, probablemente el beneficio de este servicio sea pequeño.	Ofrecer/proporcionar este servicio solo si hay otras consideraciones que apoyen que se ofrezca/proporcione el servicio a un paciente individual.
D	El USPSTF no recomienda el servicio. Hay una certeza moderada o alta de que el servicio no tiene ningún beneficio neto o de que los perjuicios superan a los beneficios.	Desaconsejar el uso de este servicio.
I	El USPSTF concluye que la evidencia actual es insuficiente para evaluar el balance entre los beneficios y los perjuicios de este servicio. La evidencia es nula, de mala calidad o discordante, y el balance entre los beneficios y los perjuicios no se puede determinar.	Véase la sección «Consideraciones clínicas» de las Recomendaciones del USPSTF. Si se ofrece, los pacientes deben comprender la incertidumbre sobre el balance entre los beneficios y los perjuicios.

Modificada del U.S. Preventive Services Task Force Procedure Manual. Diciembre 2015. https://www.uspreventiveservicestaskforce.org/Home/GetFile/6/7/procedure-manual_2015/pdf. Consultado el 25 de octubre de 2017.

El Grupo de Trabajo se ha ocupado del cribado de muchas enfermedades y trastornos. Algunos ejemplos ilustran la variedad de sus actividades. Ha

revisado la evidencia existente sobre el cribado de diferentes tipos de cáncer, de enfermedades cardiovasculares como la hipertensión, la AC y el aneurisma de aorta abdominal, de enfermedades infecciosas como la gonococia, la sífilis, la clamidiasis y las hepatitis B y C, y de enfermedades mentales como la demencia, la depresión y el riesgo de suicidio, así como sobre el cribado del glaucoma y la diabetes tipo 2. El Grupo de Trabajo también ha revisado la evidencia existente sobre la eficacia del asesoramiento acerca de muchas afecciones, como los consejos para prevenir el alcoholismo, el tabaquismo y las enfermedades relacionadas con este y los consejos para fomentar una dieta saludable y la actividad física. Estas cuestiones se han abordado en adultos, pero el Grupo de Trabajo también se ha ocupado de afecciones infantiles, como la prevención de la caries dental en niños de edad preescolar, el cribado de la escoliosis en adolescentes, el cribado auditivo neonatal, el cribado de la discapacidad visual en niños menores de 5 años y el cribado de la obesidad en niños y adolescentes. Estas y muchas más revisiones de la evidencia y recomendaciones del Grupo de Trabajo se pueden consultar en la página de internet de la Agencia para la Investigación y Calidad de la Atención Médica estadounidense (Agency for Health Care Research and Quality) (https://www.ahrq.gov). Las deliberaciones y recomendaciones del Grupo de Trabajo proporcionan un modelo muy útil para evaluar la solidez de la evidencia y convertir las inferencias causales en recomendaciones normativas.

Conclusión

Aunque las directrices de causalidad descritas en este capítulo suelen denominarse *criterios,* este término no parece del todo apropiado. Pese a que basar las inferencias causales en unos cimientos cuantitativos y estructurales firmes puede ser un objetivo deseable, en la actualidad no suele tenerse toda la información necesaria para hacerlo. Por tanto, la lista presentada en este capítulo debería considerarse solo como unas directrices cuya utilidad puede ser mayor cuando se combina con una evaluación razonada de todo el corpus de evidencia disponible a la hora de tomar decisiones sobre la causalidad.

En el siguiente capítulo se abordan varias cuestiones adicionales que deben tenerse en cuenta a la hora de derivar inferencias causales a partir de estudios epidemiológicos.

Bibliografía

1. Cameron WB. Informal Sociology: A Casual Introduction to Sociological Thinking. New York: Random House; 1963: 13. (This quotation was also attributed to Albert Einstein some years later.).
2. Ochsner A, DeBakey M. Primary pulmonary malignancy. *Surg Gynecol Obstet*. 1939;68:435.
3. Yerushalmy J. Infants with low birth weight born before their mothers started to smoke cigarettes. *Am J Obstet Gynecol*. 1972;112:277–284.
4. De Martel C, Forman D, Plummer M. Gastric cancer: epidemiology and risk factors. *Gastroenterol Clin North Am*. 2013;42:219–240.
5. Rothman KJ. Modern Epidemiology. Boston/Toronto: Little Brown and Company; 1986: [Chapter 2].
6. Evans AS. Causation and Disease: A Chronological Journey. New York: Plenum; 1993: 13–39.
7. US Department of Health, Education and Welfare Smoking and Health: Report of the Advisory Committee to the Surgeon General. Washington, DC: Public Health Service; 1964.
8. Schwartz J. Air pollution and daily mortality: a review and meta analysis. *Environ Res*. 1994;64:36–52.
9. Li C, Engstrom G, Hedblad B, et al. Blood pressure control and risk of stroke. A population-based cohort study. *Stroke*. 2005;36:725–730.
10. Franco M, Ordunez P, Caballero B, et al. Impact of energy intake, physical activity, and population-wide weight loss on cardiovascular disease and diabetes mortality in Cuba, 1980-2005. *Am J Epidemiol*. 2007;166:1374–1380.
11. Hill AB. The environment and disease: association or causation? *Proc R Soc Med*. 1965;58:295–300.
12. Marshall BJ, Warren JR. Unidentified curved bacilli in the stomachs of patients with gastritis and peptic ulceration. *Lancet*. 1984;1:1311–1315.
13. Uemura N, Okamoto S, Yamamoto S, et al. *Helicobacter pylori* infection and the development of gastric cancer. *N Engl J Med*. 2001;345:784–789.

14. Grant BF, Dawson DA. Age at onset of alcohol use and its association with DSM-IV alcohol abuse and dependence: results from the National Longitudinal Alcohol Epidemiologic Survey. *J Subst Abuse*. 1997;9:103–110.
15. Merkatz IR, Thompson JE, eds. New Perspectives on Prenatal Care. New York: Elsevier; 1990.
16. Gordis L, Kleinman JC, Klerman LV, et al. Criteria for evaluating evidence regarding the effectiveness of prenatal interventions. In: Merkatz IR, Thompson JE, eds. New Perspectives on Prenatal Care. New York: Elsevier; 1990:31–38.
17. US Preventive Services Task Force. *Procedure Manual.* December 2015 [fecha de última consulta: 25 de octubre de 2017. https://www.uspreventiveservicestaskforce.org/Home/GetFile manual_2015/pdf.

Preguntas de repaso del capítulo 14

1. *En un gran estudio de casos y controles de pacientes con cáncer de páncreas, el 17% de los pacientes eran diabéticos en el momento del diagnóstico, en comparación con el 4% de un grupo control bien emparejado (por edad, sexo, grupo étnico y otras características) en el que se evaluó la presencia de diabetes al mismo tiempo que se diagnosticaron los casos. Se concluyó que la diabetes desempeña un papel causal en el cáncer de páncreas. Esta conclusión:*

a. Es correcta.

b. Puede ser incorrecta porque no hay grupo control o de comparación.

c. Puede ser incorrecta debido a la imposibilidad de establecer la secuencia temporal entre la aparición de la diabetes y diagnóstico de cáncer de páncreas.

d. Puede ser incorrecta debido a la evaluación menos completa de la diabetes en los casos de cáncer de páncreas.

e. Puede ser incorrecta debido a la evaluación más completa del cáncer de páncreas en personas no diabéticas.

2. *Un investigador examinó los casos de muerte fetal en 27.000 embarazos y clasificó a las madres en función de si habían tenido relaciones sexuales en el mes previo al parto. Se observó que el 11% de las madres de fetos que fallecieron y el 2,5% de las madres de fetos que sobrevivieron habían tenido relaciones sexuales durante ese periodo. Se concluyó que las relaciones sexuales durante el mes previo al parto causaron las muertes fetales. Esta conclusión:*

a. Puede ser incorrecta porque las madres que tuvieron relaciones sexuales durante el mes previo al parto pueden diferir en otras características importantes de las que no las tuvieron.

b. Puede ser incorrecta porque no hay grupo de comparación.

c. Puede ser incorrecta debido a que se han utilizado las tasas de prevalencia en lugar de las tasas de incidencia, que son las necesarias.

d. Puede ser incorrecta debido a la imposibilidad de lograr un alto nivel de significación estadística.

e. *b* y *c*.

3. *Todos los criterios siguientes son importantes a la hora de hacer inferencias causales,* excepto:

a. Coherencia con los conocimientos existentes.

b. Relación dosis-respuesta.

c. Concordancia de la asociación en varios estudios.

d. Fuerza de la asociación.

e. Valor predictivo.

Las preguntas 4 y 5 se basan en la siguiente información.

Los factores A, B o C pueden causar de forma individual una determinada enfermedad sin los otros dos factores, pero solo cuando se siguen de la exposición al factor X. La exposición únicamente al factor X no se sigue de la aparición de la enfermedad, pero la enfermedad nunca se produce en ausencia de exposición al factor X.

4. *El factor X es:*

a. Una causa necesaria y suficiente.

b. Una causa necesaria, pero no suficiente.

c. Una causa suficiente, pero no necesaria.

d. Ni necesario ni suficiente.

e. Ninguna de las anteriores.

5. *El factor A es:*

a. Una causa necesaria y suficiente.

b. Una causa necesaria, pero no suficiente.

c. Una causa suficiente, pero no necesaria.

d. Ni necesario ni suficiente.

Ninguna de las anteriores.

CAPÍTULO 15

Más sobre inferencias causales: sesgos, confusión e interacción

Palabras clave

sesgo
sesgo de selección
sesgo de información
clasificación errónea
factor de confusión
estratificación
interacción

Objetivos de aprendizaje

- Revisar algunos de los posibles sesgos en los estudios epidemiológicos, como el sesgo de selección y el sesgo de información.
- Definir los factores de confusión y describir las posibles formas de enfrentarse a los factores de confusión en el diseño y/o análisis de un estudio observacional (no aleatorizado).
- Definir la interacción y presentar un marco para detectar si dos factores interactúan para influir en el riesgo de una enfermedad y en qué medida.

En este capítulo ampliamos la descripción de la relación de causalidad que se inició en el capítulo 14. Ahora la exposición la centraremos en tres problemas importantes a la hora de realizar inferencias causales: 1) los sesgos, 2) la confusión y 3) la interacción. Estos tres problemas son importantes para cualquier tipo de diseño de estudio, aunque si un estudio aleatorizado se hace correctamente, los sesgos y la confusión se minimizarán. En el contexto de los

diseños de estudio específicos se describen ejemplos de cada uno de estos problemas, pero debe tenerse en cuenta que pueden afectar a todos los tipos de diseños de estudio, y no se limitan a los tipos de estudio que se han seleccionado como ejemplos en este capítulo.

Sesgo

El sesgo se ha abordado en muchos de los capítulos anteriores, ya que es una consideración fundamental en cualquier tipo de diseño de estudio epidemiológico. Por tanto, aquí solo se añadirán algunos comentarios adicionales.

¿Qué se entiende por *sesgo*? El sesgo se ha definido como «cualquier error sistemático del diseño, realización o análisis de un estudio que tiene como resultado una estimación errónea del efecto de la exposición sobre el riesgo de la enfermedad[1]».

Sesgo de selección

En los estudios epidemiológicos se pueden encontrar varios tipos de sesgo. El primero es el *sesgo de selección*. Si la forma en que los casos y los controles, o las personas expuestas y no expuestas, se seleccionaron es tal que se observa una asociación evidente (aunque en realidad la exposición y la enfermedad no están asociadas), la asociación aparente se debe a un sesgo de selección.

El sesgo de selección se puede deber a la falta de respuesta de los participantes potenciales del estudio. Por ejemplo, si se está analizando la posible relación entre una exposición y el resultado de una enfermedad y la tasa de respuesta de los sujetos potenciales es mayor en las personas con la enfermedad que estuvieron expuestas que en las personas con la enfermedad que no estuvieron expuestas, podría observarse una asociación aparente, aunque en realidad no existe dicha asociación. Alternativamente, la asociación, aunque sea real, puede exagerarse al tener una mayor participación entre las personas con la enfermedad que estuvieron expuestas. Lo contrario también puede llevar a un sesgo (es decir, cuando las personas enfermas y expuestas tienen menos probabilidades de participar en el estudio).

En general, entre las personas que declinan participar en un estudio y las que sí lo hacen existen diferencias en muchas características demográficas, socioeconómicas, culturales, de estilo de vida y en factores del estado de salud[2]. Ronmark y cols. publicaron en 1999 un estudio que trató de caracterizar a las personas que no responden[3]. Mientras realizaban un estudio sobre la prevalencia del asma, la bronquitis crónica y los síntomas respiratorios, estos autores evaluaron las características de los pacientes que no responden y las razones de su falta de respuesta. En este estudio se solicitó la participación de 9.132 personas que vivían en Suecia. Los datos se obtuvieron mediante un cuestionario enviado por correo, y la tasa de respuesta fue del 85%. Se contactó por teléfono con una muestra de quienes no habían respondido y se les entrevistó con el mismo cuestionario empleado en el estudio principal. Los autores encontraron una proporción significativamente mayor de fumadores activos y de trabajadores manuales

entre quienes no respondieron que entre los participantes del estudio. Además, las tasas de prevalencia de sibilancias, tos crónica, expectoración, episodios de disnea, así como de asma y uso de fármacos para el asma, fue significativamente mayor entre los pacientes que no respondieron que entre los que sí lo hicieron. De este modo, las inferencias del estudio se vieron disminuidas por la pérdida de participación de los no respondedores, ya que las asociaciones que se encontraron en los encuestados estaban sesgadas hacia la nula.

En la mayoría de los estudios no puede obtenerse información de los pacientes que no responden, por lo que la falta de respuesta puede introducir un sesgo grave en el estudio que puede ser difícil (o quizás imposible) de evaluar. Por tanto, es importante que la falta de respuesta sea mínima. Por ejemplo, el Biobank del Reino Unido reclutó a 500.000 hombres y mujeres de 40 a 69 años durante 5 años para evaluar los factores de riesgo de las principales enfermedades de las edades media y avanzada[4]. Si bien es un número muy grande para un estudio de cohortes, solo el 5% de los adultos contactados accedió a participar en el estudio. El grado en que la muestra reclutada representa a la población objetivo puede ser motivo de preocupación. Sin embargo, dado que todos los ciudadanos del Reino Unido están incluidos en el National Health Service, las comparaciones de la cohorte con la población objetivo pueden estimarse. Además, todos los pacientes que no responden deben caracterizarse tanto como se pueda mediante el uso de toda la información disponible para determinar las formas en que se diferencian de quienes sí responden (edad estimada, sexo, localización geográfica, etc.) y para evaluar el impacto probable de su falta de respuesta en los resultados del estudio.

En los estudios de cohortes, las pérdidas de participantes durante el seguimiento también pueden dar como resultado un sesgo de selección. Sin embargo, dado que hay información de referencia sobre estos participantes perdidos en el momento de la inscripción, es posible comparar aquellos que se pierden con los que no se pierden en la observación respecto a una serie de factores sociodemográficos y de otro tipo. En Estados Unidos y otros países, también es posible vincular la información sobre los no respondedores a un registro nacional de defunciones, lo que permite una comparación de las tasas de mortalidad del resultado en estudio para los respondedores y los no respondedores. Este tipo de vínculo es particularmente útil cuando el resultado primario del estudio es la mortalidad[5–8]. Es importante tener en cuenta la distinción entre la *selección de sujetos para un estudio* y el *sesgo de selección.* Prácticamente en todos los estudios realizados en poblaciones humanas se seleccionan los sujetos de estudio a partir de una población mayor (objetivo). La naturaleza de esta selección puede afectar a la *generalización* o *validez externa* del estudio, pero no necesariamente afecta a la validez de las comparaciones realizadas en el seno del estudio o a su *validez interna.* Por otro lado, cuando se ha cometido un error sistemático en la selección de uno o más de los grupos de estudio que se compararán, puede

producirse un *sesgo de selección.* Este sesgo puede dar lugar a *odds ratios* (OR) o riesgos relativos (RR) que tal vez no sean una estimación correcta y, por tanto, dar origen a inferencias no válidas respecto a las asociaciones de la exposición y la enfermedad. El sesgo de selección es un error en la selección de un grupo o grupos de estudio dentro de dicho estudio y puede tener un gran impacto sobre la validez interna del estudio y la legitimidad de la inferencia respecto a la asociación de la exposición y el resultado. La selección de una población de estudio a partir de una población objetivo más amplia no se debe confundir con el sesgo de selección, que se debe a un error sistemático en la selección de los sujetos de uno o más de los grupos de estudio, como participantes expuestos/no expuestos o casos/controles.

Un ejemplo «clásico» interesante de sesgo de selección se demostró en 1974 con la publicación de los datos que parecían sugerir una relación entre el uso de reserpina (agente antihipertensivo de uso común) y el riesgo de cáncer de mama. Tres artículos que apoyaban esta asociación se publicaron en la misma edición de *Lancet* en septiembre de 1974[9-11]. Los tres artículos describían los resultados de tres estudios llevados a cabo en Boston, Gran Bretaña y Helsinki, respectivamente.

En uno de estos artículos, que ejemplifica el problema del sesgo de selección, Heinonen y cols.[11] describieron los resultados de un estudio de casos y controles con datos emparejados realizado en pacientes quirúrgicos de un hospital de Helsinki. Las mujeres con cáncer de mama se compararon con mujeres sin dicha neoplasia en cuanto al uso de reserpina. Las mujeres con cáncer de mama recién diagnosticado se identificaron a partir de un registro de altas hospitalarias y de los registros donde se indicaban las operaciones en el hospital. Estas pacientes eran los «casos» y cada una se emparejó por edad y año de su operación con un control, que era una mujer ingresada para cirugía programada por alguna afección benigna. El análisis se realizó sobre un total de 438 pares de casos y controles. Como se observa en la tabla 15.1, había 45 pares en los que el caso tomaba reserpina y el control no y 23 pares en los que el control usaba reserpina y el caso no. El OR resultante de los datos emparejados fue de 45/23 o 1,96.

Tabla 15.1

Resultados de un análisis de datos emparejados de un estudio de casos y controles sobre el uso de reserpina y el cáncer de mama

	Controles	
Casos de cáncer de mama	*Usaban reserpina*	*No usaban reserpina*
Usaban reserpina	8	45
No usaban reserpina	23	362

$$\textit{odds ratio}\ \text{de datos emparejados} = \frac{45}{23} = 1{,}96$$

Modificada de Heinonen OP, Shapiro S, Tuominen L, Turunen MI. Reserpine use in relation to breast cancer. *Lancet.* 1974;2:675–677.

Sin embargo, se advirtió un problema en el método utilizado para la selección de los controles. Al realizar dicha selección, los autores excluyeron a las mujeres que habían sido sometidas a las siguientes operaciones: colecistectomía, tiroidectomía por tirotoxicosis, cirugía por nefropatía y cualquier operación cardiaca, simpatectomía o injerto vascular. Se excluyeron porque, en el momento en el que se realizó el estudio, la reserpina era uno de los fármacos que solían usarse en el tratamiento de estas afecciones. Los autores temían que, si las pacientes con estas afecciones se hubiesen incluido en este estudio de casos y controles, la prevalencia del uso de reserpina en los controles habría sido artificialmente alta, de modo que, incluso si el uso de reserpina fuese mayor en los casos de cáncer de mama, el aumento podría no detectarse.

Por desgracia, al intentar solucionar esta preocupación, los autores crearon un problema diferente, porque estas exclusiones no se aplicaron a los casos.

Al excluir del grupo control a las pacientes con estas afecciones, se creó un grupo control en el que la prevalencia del uso de reserpina era artificialmente baja, debido a que se excluyó a un gran grupo de posibles usuarias de reserpina. Por tanto, aunque en realidad el uso de reserpina no era mayor en las mujeres que desarrollaron cáncer de mama, este estudio podría mostrar una diferencia en cuanto a dicho uso entre los casos y los controles debido únicamente al modo en el que se seleccionaron los controles.

Este tipo de sesgo de selección se ha denominado *sesgo de exclusión*[12]. Se produce cuando los investigadores aplican diferentes criterios de elegibilidad («inclusión») a los casos y a los controles en cuanto a qué afecciones clínicas previas permitirían la elegibilidad en el estudio y cuáles podrían ser motivo de exclusión. Horwitz y Feinstein[12] intentaron replicar el estudio sobre la reserpina en 257 mujeres con cáncer de mama y 257 controles, calculando el OR de dos maneras: primero, incluyendo a todas las mujeres y, después, tras excluir de los controles a las mujeres con antecedentes de enfermedad cardiovascular. El OR determinado al incluir a todas las mujeres fue de 1,1, pero cuando se excluyeron las mujeres con enfermedad cardiovascular se elevó a 2,5. Los resultados apoyan la idea de que la relación aparente entre el uso de reserpina y el cáncer de mama en el estudio de Helsinki se debió a un sesgo de selección secundario a los diferentes criterios para seleccionar a los controles en el estudio. El estudio realizado sobre el café y el cáncer de páncreas[13] tenía un problema similar; se comentó en el capítulo 7.

La aplicación de los mismos criterios de elegibilidad a la selección de casos y controles debería resultar idealmente en el fenómeno acuñado por Schlesselman como *sesgo de compensación*[1]. Un corolario de esta noción es que, cuando el sesgo en la selección de casos y controles es de la misma magnitud, se logra compensar el sesgo. Un ejemplo hipotético de sesgo de compensación es un estudio de casos y controles en el que tanto los casos como los controles se identifican a través de un programa de cribado. Las personas que deciden someterse a un examen de cribado no saben antes de la selección si serán casos o controles y, por lo tanto, la autoselección para el cribado es independiente del resultado del examen. En el cáncer de colon, por ejemplo, las personas que eligen someterse a una colonoscopia tienen más probabilidades de tener antecedentes familiares de cáncer de colon y de tener un historial de pólipos en el colon. Sin embargo, como el proceso de autoselección es el mismo para aquellos que posteriormente tienen cáncer de colon y aquellos que no, se logra compensar el sesgo, y, como resultado, el OR es imparcial. En otras palabras, estos factores de selección (antecedentes familiares y antecedentes de pólipos de colon) están igualmente inflados (sesgados) en ambos, casos y controles, y por lo tanto, al calcular el OR, se compensa el «factor de inflación» (sesgo). Si las personas con antecedentes familiares de cáncer de colon tienen el doble de probabilidades de ser incluidas en el estudio, entonces:

$$\begin{aligned}
&\text{Cociente de posibilidades} \\
&= \frac{\text{Probabilidad de exposición en casos} \times 2,0}{\text{Probabilidad de exposición en controles} \times 2,0} \\
&= \frac{\text{Probabilidad de exposición en casos}}{\text{Probabilidad de exposición en controles}} \times 1,0 \\
&= \text{Cociente de posibilidades real}
\end{aligned}$$

Este fenómeno explica por qué la medida de la asociación puede ser imparcial incluso si cada frecuencia de exposición está sesgada.

Sesgo de información

El *sesgo de información* puede ocurrir cuando los medios para obtener información acerca de los sujetos del estudio son imperfectos, de modo que algunos de los datos recogidos sobre la exposición y/o los resultados de la enfermedad no son correctos.

Teniendo en cuenta las inexactitudes de los métodos de adquisición de datos, en ocasiones es posible clasificar erróneamente los sujetos y, de este modo, introducir un *sesgo de clasificación errónea.* Por ejemplo, en un estudio de casos y controles, algunas personas que tienen la enfermedad (casos) pueden clasificarse erróneamente sin saberlo como controles, y algunos sin la enfermedad (controles) pueden clasificarse erróneamente como casos. Esto puede deberse, por ejemplo, a una sensibilidad y especificidad limitadas de las pruebas diagnósticas utilizadas o a la insuficiencia de la información obtenida de las historias clínicas o de otros registros. Otra posibilidad es que se puede clasificar erróneamente el estado de exposición de una persona: puede creerse que la persona estuvo expuesta cuando no lo estuvo, o viceversa. Si los datos de exposición se basan en entrevistas, por ejemplo, puede que los sujetos ignoren su exposición o pueden pensar erróneamente que no se produjo. También es posible que los participantes de la investigación puedan retener información si piensan que podría ser potencialmente embarazosa. Si la determinación de la exposición se basa en registros antiguos, los datos pueden haberse perdido, ser incompletos o inexactos.

La clasificación errónea puede ser de dos tipos: diferencial y no diferencial. En la *clasificación errónea diferencial,* la tasa de clasificación errónea es distinta en los diferentes grupos de estudio. Por ejemplo, la clasificación errónea de la exposición se puede producir de tal manera que los casos no expuestos estén mal clasificados como expuestos con más frecuencia que los controles no expuestos estén mal clasificados como expuestos. Esto sucedía en el ejemplo hipotético del sesgo de recuerdo (declaraciones maternas de infección

durante el embarazo y bebés nacidos con malformaciones congénitas) comentado al describir los estudios de casos y controles (v. cap. 7). Las mujeres que tuvieron un bebé con una malformación tendían a recordar (o comunicar) más infecciones leves que se produjeron durante sus embarazos que las madres de los bebés sin malformaciones. Los controles probablemente tenían menos probabilidades de recordar una infección leve durante el embarazo. Por tanto, hubo una tendencia a una clasificación errónea diferencial respecto a la infección prenatal, de modo que hubo un número mayor de casos no expuestos mal clasificados como expuestos que controles no expuestos. Esto dio lugar a una asociación aparente entre malformaciones e infecciones, a pesar de que no existía ninguna. Por tanto, un sesgo de clasificación errónea diferencial puede dar lugar tanto a una asociación aparente, aunque no exista en realidad, como a una aparente falta de asociación, cuando lo cierto es que sí existe.

Por el contrario, la *clasificación errónea no diferencial* se debe al grado de imprecisión que caracteriza al modo en que se obtiene la información de cualquier grupo de estudio, ya sean casos y controles o personas expuestas y no expuestas. Tal clasificación errónea no se relaciona con el estado de exposición o con el estatus de caso o control, sino que es solo un problema inherente de los *métodos de recogida de datos*. El efecto habitual de la clasificación errónea no diferencial es que el riesgo relativo (RR) o el OR tiende a diluirse y se desplaza hacia un valor de 1,0. Dicho de otro modo, es menos probable detectar una asociación, incluso si existe realmente.

Esto se puede ver de forma intuitiva. Supongamos que, en realidad, existe una fuerte asociación entre una exposición y una enfermedad, es decir, las personas sin la enfermedad tienen mucha menos exposición que aquellas con la enfermedad. Por desgracia, de forma errónea, se han incluido algunas personas enfermas en el grupo control (falsos negativos) y algunas personas no enfermas en el grupo de casos (falsos positivos). Dicho de otro modo, se han clasificado erróneamente algunos de los sujetos en lo que se refiere al diagnóstico. En esta situación, los controles no tendrán una tasa tan baja de exposición debido a que algunas personas enfermas se han incluido erróneamente en este grupo, y los casos no tendrán una tasa tan alta de exposición, porque algunas personas no enfermas se han incluido erróneamente en el grupo de casos. Como resultado, se observará una diferencia menor en cuanto a la exposición entre los casos y los controles de la que existe en realidad entre las personas enfermas y no enfermas.

Algunos de los tipos y fuentes de sesgo de información en los estudios epidemiológicos se muestran en el cuadro 15.1.

Cuadro 15.1 Algunos tipos y fuentes de sesgo de información

Sesgo al resumir las historias clínicas.

Sesgo en la entrevista.
Sesgo en las entrevistas sustitutas.
Sesgo de vigilancia.
Sesgo de recuerdo.
Sesgo de notificación.

El sesgo puede introducirse por el modo en el que la información se obtiene de las historias clínicas, de los registros laborales o de otro tipo, o por la forma en la que los entrevistadores hacen las preguntas. También puede deberse a *entrevistas sustitutas.* ¿Qué quiere decir esto? Supongamos que estamos realizando un estudio de casos y controles sobre el cáncer de páncreas. La letalidad de esta enfermedad es muy alta y el tiempo de supervivencia es muy corto (la mediana de supervivencia tratada es menor de 1 año). Cuando nos preparamos para entrevistar a los casos, nos encontramos con que muchos de ellos han fallecido y muchos de los que han sobrevivido están demasiado enfermos para ser entrevistados. Como alternativa, se puede preguntar a un miembro de la familia para obtener información sobre los antecedentes laborales del caso, su dieta y otras exposiciones y características. El entrevistado suele ser la pareja o un hijo. A la hora de obtener la información a partir de tales sustitutos surgen varios problemas. En primer lugar, puede que no tengan información precisa sobre los antecedentes del caso. Puede que la pareja no conozca las exposiciones del caso relacionadas con su trabajo, o las conozca con inexactitud. Los hijos a menudo saben aún menos que la pareja. En segundo lugar, hay evidencia de que cuando una mujer informa sobre el trabajo y el estilo de vida de su pareja después de su fallecimiento, tiende a elevar su nivel laboral y su estilo de vida. Puede atribuirle una categoría profesional más elevada de la que en realidad ocupaba. También puede que le convierta en no bebedor y/o no fumador una vez que ha fallecido.

Si se monitoriza una población durante un periodo de tiempo, es posible que la constatación de la enfermedad sea mejor en la población monitorizada que en la población general, y puede introducirse un *sesgo de vigilancia,* que da lugar a una estimación errónea del RR o del OR. Por ejemplo, hace algunos años se suscitó un interés considerable por la posible relación entre el uso de anticonceptivos orales y la tromboflebitis. Se sugirió que los médicos monitorizasen a las pacientes que habían recibido anticonceptivos orales mucho más de cerca que a otras pacientes. Como resultado, fueron más propensos a identificar los casos de tromboflebitis que se desarrollaron en las pacientes que tomaban anticonceptivos orales (y, por tanto, que estaban siendo vigiladas más de cerca) que en otras pacientes no tratadas con anticonceptivos orales, que por tanto no estaban tan monitorizadas. Como consecuencia, simplemente por la mejor constatación de tromboflebitis en las mujeres que recibían anticonceptivos orales, se puede observar una asociación aparente entre la tromboflebitis y el uso de anticonceptivos orales, aunque no exista ninguna asociación verdadera.

En el capítulo 7 se describió el *sesgo de recuerdo* en los estudios de casos y controles. Este sesgo mejora el recuerdo en los casos en comparación con los controles. Por tanto, un fragmento de información determinado, como una exposición potencialmente relevante, puede ser recordado por un caso, pero olvidado por un control (como vimos en el recuerdo de infecciones en las madres de bebés con malformaciones congénitas). Un tipo relacionado de sesgo es el *sesgo de notificación,* en el que un sujeto puede ser reacio a notificar una exposición que conoce, debido a sus actitudes, creencias y percepciones. Este es un problema común en los estudios sobre VIH/SIDA, donde los factores de riesgo incluyen prácticas sexuales y uso de sustancias, que pueden estar sujetos a un sesgo de notificación significativo. Se han desarrollado enfoques metodológicos para superar este sesgo y evitar que los participantes traten de «causar una buena impresión». La autoentrevista con audio asistida por ordenador (ACASI, *Audio Computer-Assisted Self-Interviewing*) se desarrolló para permitir que los participantes del estudio respondiesen preguntas sensibles a un ordenador en lugar de que se las haga verbalmente un entrevistador, con algunos resultados importantes[14,15]. Si esta falta de notificación es más frecuente en los casos o en los controles, puede dar lugar a un sesgo. A continuación, se presenta un ejemplo.

El término *sesgo de deseo* fue acuñado por Wynder y cols.[16] para indicar el sesgo introducido por los sujetos que han desarrollado una enfermedad y que en el intento de responder a la pregunta «¿Por qué yo?» pretenden mostrar, a menudo sin intención, que la enfermedad no es su culpa. Por tanto, pueden negar ciertas exposiciones asociadas al estilo de vida (como fumar o beber); si están pensando en presentar una demanda, pueden exagerar las exposiciones relacionadas con el lugar de trabajo. El sesgo de deseo puede considerarse un tipo de sesgo de notificación.

Hay que recordar que *el sesgo se debe a un error en el diseño o la realización de un estudio.* Por tanto, se debe intentar reducir o eliminar el sesgo o, por lo menos, identificarlo y tenerlo en cuenta al interpretar los resultados de un estudio. Sin embargo, puede que no siempre se disponga de los datos necesarios para documentar y evaluar el tipo y el grado de sesgo.

A continuación, se presenta otro ejemplo. La relación entre el aborto inducido y el riesgo de cáncer de mama ha sido un tema que ha suscitado gran interés en los últimos años. Aunque, en general, no se ha descrito ninguna asociación entre el aborto *espontáneo* y el riesgo de cáncer de mama, los datos han sido contradictorios en cuanto a la posible relación entre aborto *inducido* y cáncer de mama. Se ha sugerido que el sesgo de notificación podría haber intervenido en los estudios de casos y controles donde se indicó una asociación positiva: los controles sanos pueden haber sido más reticentes que las mujeres con cáncer de mama a notificar que habían tenido un aborto inducido.

Un estudio sobre el aborto inducido y el riesgo de cáncer de mama permitió a los investigadores evaluar el alcance y el posible papel de este sesgo de notificación, que es un tipo de sesgo de información. Rookus y van

Leeuwen[17] publicaron un estudio de casos y controles realizado en los Países Bajos en el que la estimación del RR ajustado global fue de 1,9 para el aborto inducido y el cáncer de mama en mujeres que habían tenido hijos. (No se observó ninguna asociación en mujeres nulíparas.) Después, compararon los resultados en dos regiones del país (la región del sudeste, que tiene una mayoría de población católica romana, que es más conservadora, y la región occidental, incluida Ámsterdam, que tiene actitudes más liberales hacia el aborto). Esta diferencia de actitudes se refleja en el hecho de que las tasas de abortos inducidos en el sudeste han sido siempre más bajas que en el oeste. Como se observa en la tabla 15.2, los autores observaron que la asociación entre el aborto inducido y el cáncer de mama era mucho más fuerte en el sudeste conservador (RR ajustado estimado = 14,6) que en el oeste más liberal (RR ajustado estimado = 1,3), lo que sugiere que la conclusión general de una asociación entre el cáncer de mama y el aborto inducido en este estudio era atribuible en gran parte a una infranotificación de abortos por parte de los controles en el sudeste. Además, dado que este estudio fue parte de un estudio poblacional de casos y controles sobre el uso de anticonceptivos orales y riesgo de cáncer de mama, fue posible buscar datos que apoyaran la posibilidad de que este sesgo de infranotificación explicase las diferencias regionales. En el análisis de las respuestas sobre el uso de anticonceptivos orales, cuando las respuestas de las mujeres se compararon con las prescripciones de sus médicos, se observó que los controles de la región del sudeste habían notificado una duración menor del uso de anticonceptivos orales con una diferencia de más de 6 meses respecto a los controles de la región occidental.

Tabla 15.2

Riesgos relativos[a] (RR) e intervalos de confianza (IC) del 95% del desarrollo de cáncer de mama en mujeres de 20-45 años en relación con abortos inducidos previos notificados por mujeres que habían tenido hijos en todas las regiones y en las regiones occidental y sudeste de los Países Bajos

	RR no ajustado	RR ajustado[b]	IC del 95%
Todas las regiones	1,8	1,9	1,1-3,2
Región occidental	1,2	1,3	0,7-2,6
Región sudeste	12,3	14,6	1,8-120

[a] Riesgos relativos estimados utilizando métodos de regresión logística condicional para datos emparejados.

[b] Ajustado por aborto espontáneo o provocado, edad del primer embarazo a término, número de embarazos a término, semanas de lactancia materna, antecedentes familiares de cáncer de mama y uso de anticonceptivos inyectables.

Modificada de Rookus MA, Van Leeuwen FE. Induced abortion and risk for

breast cancer: Reporting (recall) bias in a Dutch case-control study. *J Natl Cancer Inst.* 1996;88:1759–1764.

La ubicación de la entrevista también puede influir en la veracidad de la información proporcionada por los sujetos del estudio. D'Avanzo y cols. encontraron diferentes proporciones de consumo de alcohol comunicado al entrevistar a pacientes en el hospital y luego nuevamente en casa: en las entrevistas hospitalarias el 62% de los preguntados afirmó ser bebedor, pero en las entrevistas domiciliarias posteriores dicha cifra ascendió al 72%[18].

Confusión

Un problema que se plantea en muchos estudios epidemiológicos es que la observación de una asociación verdadera suscita la tentación de derivar una inferencia causal cuando, en realidad, puede que la relación no sea causal. Esto nos lleva al tema de la *confusión*, uno de los problemas principales de los estudios epidemiológicos observacionales.

¿Qué es la *confusión*? En un estudio sobre si la exposición A es una causa de la enfermedad B, se dice que un tercer factor, el factor X, es un factor de confusión si se cumple lo siguiente:

1. El factor X es un factor de riesgo conocido para la enfermedad B.
2. El factor X se asocia con la exposición A, pero no es un resultado de la exposición A.

Recordemos el ejemplo que se expuso en el capítulo 7 sobre la relación entre el café y el cáncer de páncreas. El tabaquismo era un factor de confusión, porque, a pesar de que se estaba estudiando una posible relación entre el consumo de café (exposición A) y el cáncer de páncreas (enfermedad B), las siguientes afirmaciones sobre el tabaquismo (factor X) son ciertas:

1. El tabaquismo es un factor de riesgo conocido para el cáncer de páncreas.
2. El tabaquismo se asocia con el consumo de café, pero no es secundario a beber café.

Por tanto, si se observa una asociación entre el consumo de café y el cáncer de páncreas, puede ser 1) que el café realmente cause el cáncer del páncreas o 2) que la asociación observada entre el consumo de café y el cáncer de páncreas se deba a que el tabaquismo es un factor de confusión (es decir, se observa la asociación entre el consumo de café y el cáncer de páncreas porque el tabaquismo es un factor de riesgo para el cáncer de páncreas y además se asocia con el consumo de café; fig. 15.1).

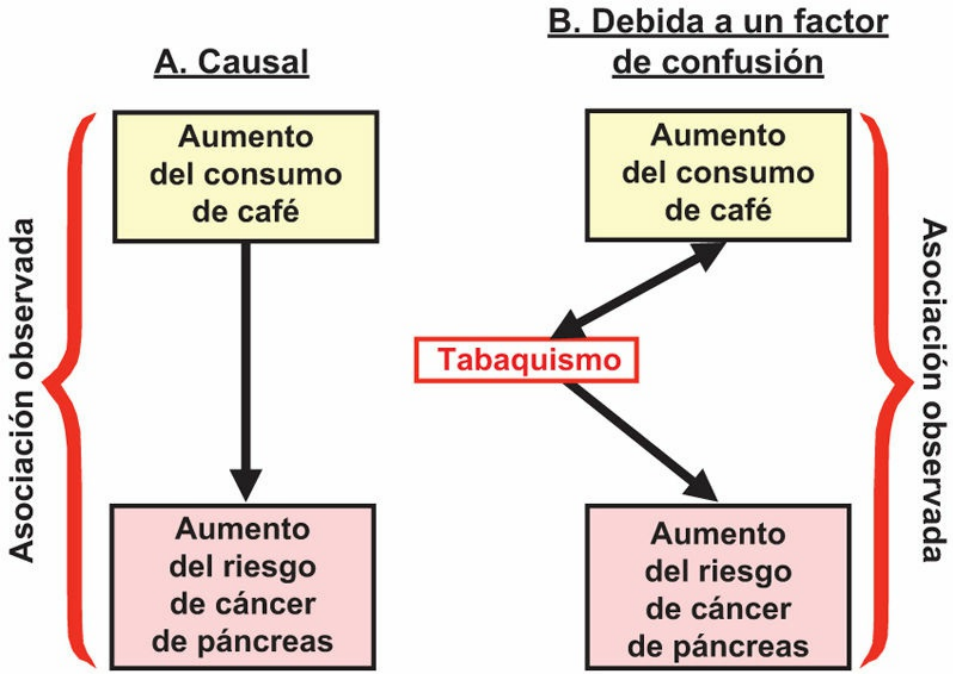


FIG. 15.1 Asociación entre el aumento del consumo de café y el aumento del riesgo de cáncer de páncreas.

Cuando se observa una asociación, hay que preguntarse si es causal (v. fig. 15.1A) o si se debe a la confusión por un tercer factor que sea a la vez un factor de riesgo para la enfermedad y que se asocie con la supuesta exposición en cuestión (v. fig. 15.1B).

A continuación, se presenta un ejemplo hipotético. En la tabla 15.3 se muestran los datos de un estudio de casos y controles sin emparejamiento sobre una exposición y una enfermedad en el que se estudiaron 100 casos y 100 controles.

Tabla 15.3

Ejemplo hipotético de confusión en un estudio de casos y controles sin emparejamiento: I. Números de casos y controles expuestos y no expuestos

Expuesto	Casos	Controles
Sí	30	18
No	70	82
Total	100	100

$$Odds\ ratio = \frac{30 \times 82}{70 \times 18} = 1{,}95$$

El OR de datos no emparejados que se calculó fue de 1,95. A continuación nos planteamos esta pregunta: ¿Esta asociación de la exposición con la enfermedad es causal o podría deberse a diferencias en la distribución de las edades de los casos y los controles? Dicho de otro modo, ¿la edad es un factor de confusión respecto a la relación observada? La primera pregunta que debe responderse al abordar esta cuestión es si la edad se relaciona con el hecho de ser un caso o un control. La respuesta a esta pregunta procede del análisis de

la tabla 15.4.

Tabla 15.4

Ejemplo hipotético de confusión en un estudio de casos y controles sin emparejamiento: II. Distribución de los casos y controles por edad

Edad (años)	Casos	Controles
<40	50	80
≥40	50	20
Total	100	100

Se observa que el 80% de los controles son menores de 40 años, en comparación con solo el 50% de los casos. Por tanto, la edad avanzada se asocia con el hecho de ser un caso (tener la enfermedad) y la edad más joven se asocia con el hecho de ser un control (no tener la enfermedad).

La siguiente pregunta es si la edad está relacionada con el estado de exposición.

En la tabla 15.5 se analiza la relación entre la edad y la exposición de los 200 sujetos estudiados, con independencia de su estatus de casos y controles. Se observa que 130 personas eran menores de 40 años (50 + 80 de la fila superior de la tabla 15.4), y, de ellas, 13 (10%) estuvieron expuestas. Entre los 70 sujetos mayores de 40 años, 35 (50%) estuvieron expuestos. Por tanto, la edad está claramente relacionada con la exposición. Ahora se sabe que la edad se relaciona con el hecho de ser un caso (los casos eran mayores que los controles); también se sabe que la exposición se relaciona con la edad avanzada.

Tabla 15.5

Ejemplo hipotético de confusión en un estudio de casos y controles sin emparejamiento: III. Relación de la exposición con la edad

Edad (años)	Total	Expuestos	No expuestos	% expuestos
<40	130	13	117	10
≥40	70	35	35	50

Como se muestra en la figura 15.2, la siguiente pregunta es: ¿La asociación entre la exposición y la enfermedad es de tipo causal (v. fig. 15.2A), o puede que se esté observando una asociación entre la exposición y la enfermedad solo porque hay una diferencia de edad entre los casos y los controles y la mayor edad también se relaciona con la exposición (v. fig. 15.2B)? Dicho de otro modo, ¿la exposición causa la enfermedad o la asociación observada entre la exposición y la enfermedad se debe a la confusión por un tercer factor (en este caso, la edad)?

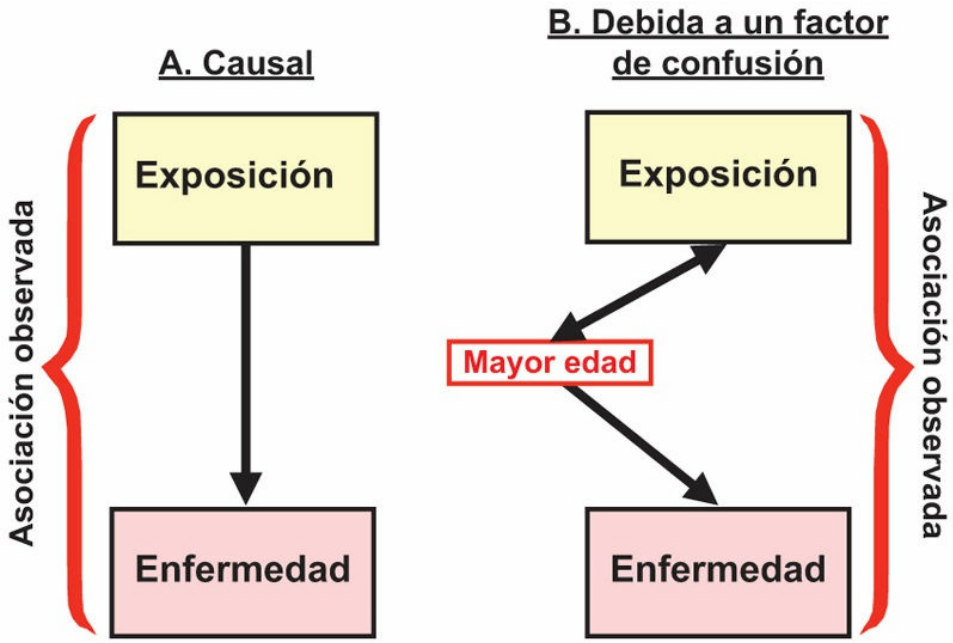


FIG. 15.2 Representación esquemática del problema de los posibles factores de confusión.

¿Cómo se puede aclarar este tema? Un método se presenta en la tabla 15.6. Se puede llevar a cabo un análisis estratificado distribuyendo los sujetos en dos grupos de edad: menores de 40 años y mayores de 40 años. Dentro de cada estrato, se elabora una tabla de 2 × 2 y se calcula el OR para cada uno. Cuando se calcula este cociente por separado para los pacientes más jóvenes y los mayores, se obtiene un valor de 1,0 en cada estrato. Por tanto, la única razón por la que originalmente se obtuvo un OR de 1,95 en la tabla 15.3 se debía a que había una diferencia en la distribución por edad entre los casos y los controles. En este ejemplo la edad es, por tanto, un factor de confusión.

Tabla 15.6

Ejemplo hipotético de confusión en un estudio de casos y controles sin emparejamiento: IV. Cálculos del *odds ratio* después de la estratificación por edad

Edad (años)	Expuestos	Casos	Controles	*Odds ratio*
<40	Sí	5	8	$\frac{5\times72}{45\times8}=\frac{360}{360}=1{,}0$
	No	45	72	
	Totales	**50**	**80**	
40	Sí	25	10	$\frac{25\times10}{25\times10}=\frac{250}{250}=1{,}0$
	No	25	10	
	Totales	**50**	**20**	

¿Cómo se puede solucionar el problema de la confusión? Como se observa

en el cuadro 15.2, este problema puede solucionarse bien mediante el diseño y la realización de un estudio o bien mediante el análisis de los datos. Al *diseñar y realizar un estudio de casos y controles,* se pueden emparejar los casos con los controles, como se comentó en el capítulo 7 (mediante un emparejamiento de grupos o individual), según el posible factor de confusión sospechado. En este ejemplo, podría realizarse un emparejamiento por edad para eliminar cualquier diferencia de edad entre los casos y los controles. Si después de realizar este emparejamiento se observa una asociación entre la exposición y la enfermedad, quedaría claro que no se puede atribuir la asociación observada a diferencias de edad entre los casos y los controles.

Cuadro 15.2 Estrategias para manejar la confusión

A la hora de diseñar y llevar a cabo el estudio:

1. Emparejamiento individual.
2. Emparejamiento de grupo.

Al analizar los datos:

1. Estratificación.
2. Ajuste.

Como alternativa, se puede abordar el problema de la confusión en el *análisis de datos* de dos posibles maneras: estratificación o ajuste. A continuación, se describe brevemente la estratificación, que se acaba de demostrar en el ejemplo hipotético (v. tabla 15.6). Supongamos que estamos interesados en la relación entre el tabaquismo y el cáncer de pulmón. Queremos saber si el mayor riesgo observado de cáncer de pulmón en los fumadores podría deberse a la confusión por la contaminación del aire y/o la urbanización. Puede que estemos observando una relación entre el tabaquismo y el cáncer de pulmón no porque fumar cause dicho cáncer, sino debido a que la contaminación del aire causa cáncer de pulmón y el tabaquismo es más frecuente en las zonas contaminadas (como las zonas urbanas). Quizás fumar sea más común en las ciudades que en las áreas rurales.

¿Cómo se puede responder a esta pregunta? Un método sería estratificar los datos por grado de urbanización (rural, pueblo o gran urbe). A continuación, se calculan las tasas de cáncer de pulmón en los fumadores y no fumadores en cada estrato de urbanización (tabla 15.7).

Tabla 15.7

Ejemplo de estratificación: tasas de cáncer de pulmón en función del tabaquismo y del grado de urbanización

	TASAS DE CÁNCER	
Grado de urbanización	**No fumadores**	**Fumadores**

Ninguno		
Ligero		
Pueblo		
Ciudad		
Totales		

Si la relación entre el cáncer de pulmón y el tabaquismo se debe en realidad a este último, y no al efecto de confusión de la contaminación y/o la urbanización, *en cada estrato de urbanización* la incidencia de cáncer de pulmón debe ser mayor en los fumadores que en los no fumadores. Esto dejaría claro que la asociación observada entre el tabaquismo y el cáncer de pulmón no podría deberse al grado de urbanización.

Es posible que no solo se quiera dicotomizar los grupos en fumadores y no fumadores, sino incluir también en el análisis el número de cigarrillos fumados al día (que por supuesto está sujeto a cierto grado de error, al ser estimado por los participantes del estudio).

En la tabla 15.8 se ha subdividido el consumo de cigarrillos en varias categorías según la respuesta de los participantes acerca de la cantidad fumada al día. De nuevo, se puede calcular la incidencia en cada celda de la tabla. Si la asociación observada entre el tabaquismo y el cáncer de pulmón no se debe a la confusión por la urbanización y/o la contaminación, sería previsible observar un patrón de dosis-respuesta *en cada estrato de urbanización.*

Tabla 15.8

Ejemplo de una mayor estratificación: tasas de cáncer de pulmón en función el nivel de tabaquismo y del grado de urbanización

	TASAS DE CÁNCER			
		FUMADORES		
Grado de urbanización	**No fumadores**	**<1 paquete/día**	**1 paquete/día**	**>1 paquete/día**
Ninguno				
Ligero				
Pueblo				
Ciudad				
Totales				

En la figura 15.3 se muestran las tasas de mortalidad por cáncer de pulmón reales ajustadas por edad por 100.000 personas-años según la clasificación del hábitat urbano/rural y la categoría de tabaquismo de uno de los estudios epidemiológicos clásicos sobre tabaquismo y mortalidad por cáncer de pulmón. Para cada nivel de urbanización, las tasas de mortalidad por cáncer de pulmón en los fumadores se indican con barras azules y las tasas de mortalidad en no fumadores, con barras de color verde claro. Se observa que, en todos los niveles (o estratos) de urbanización, la mortalidad por cáncer de

pulmón es mayor en los fumadores que en los no fumadores. Por tanto, la asociación observada entre el tabaquismo y el cáncer de pulmón no se puede atribuir al grado de urbanización. Al analizar cada estrato por separado, se mantiene la urbanización constante, y aún así se observa una mortalidad mucho más alta por cáncer de pulmón en los fumadores que en los no fumadores.

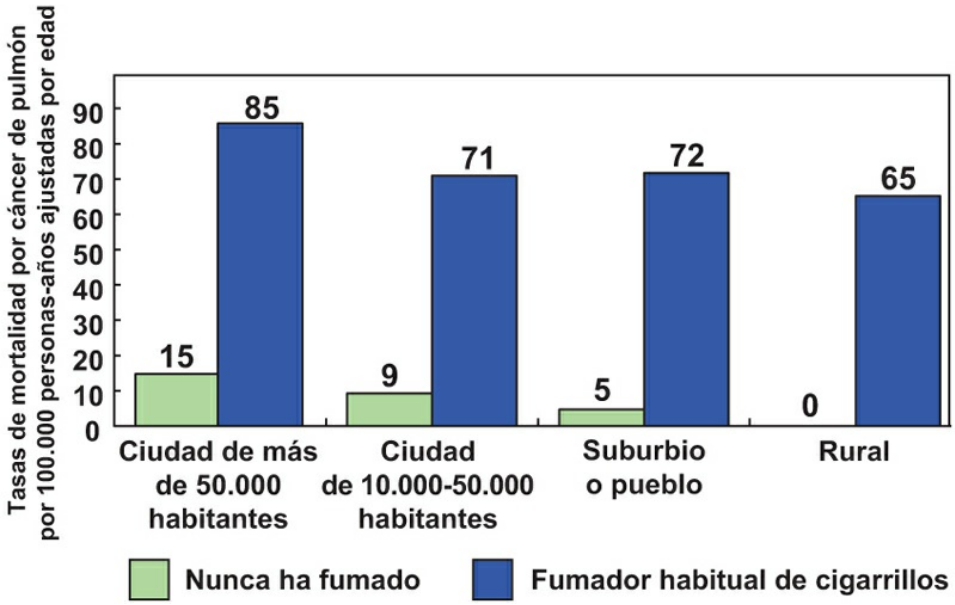


FIG. 15.3 Tasas de mortalidad por cáncer de pulmón por 100.000 personas-años ajustadas por edad según la clasificación del hábitat urbano/rural y la categoría de tabaquismo. (Modificada de Hammond EC, Horn D. Smoking and death rates: Report on 44 months of follow-up of 187,783 men: II. Death rates by cause. *JAMA.* 1958;166:1294–1308. Copyright 1958, American Medical Association.)

Al mismo tiempo, es interesante analizar los datos de los no fumadores (barras verdes). Si se traza una línea que una los extremos superiores de estas barras, se observa que, cuanto mayor es el nivel de urbanización, mayor es la incidencia de cáncer de pulmón en los no fumadores (fig. 15.4). Por tanto, existe una relación dosis-respuesta entre el cáncer de pulmón y la urbanización en los no fumadores. Sin embargo, como ya se ha visto, esta relación no puede explicar la asociación entre el cáncer de pulmón y el tabaquismo, ya que esta última relación se mantiene dentro de cada nivel de urbanización.

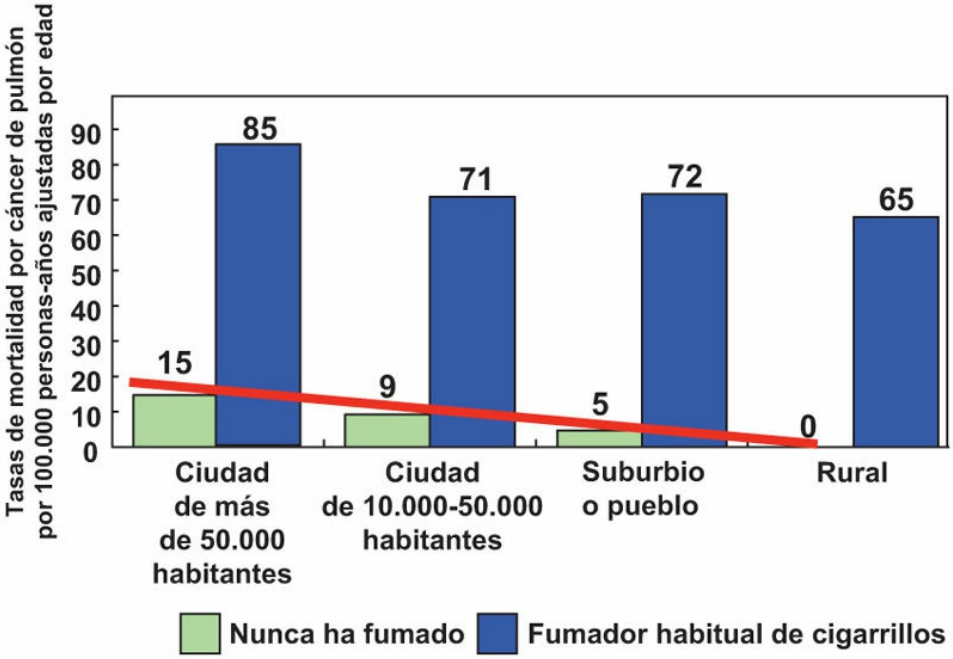


FIG. 15.4 Relación entre el grado de urbanización y las tasas de mortalidad por cáncer de pulmón en no fumadores. La *línea inclinada* conecta las tasas de mortalidad por cáncer de pulmón por 100.000 personas-años ajustadas por edad según la clasificación del hábitat urbano/rural en no fumadores. (Modificada de Hammond EC, Horn D. Smoking and death rates: Report on 44 months of follow-up of 187,783 men: II. Death rates by cause. *JAMA.* 1958;166:1294–1308. Copyright 1958, American Medical Association.)

En la figura 15.5 se muestra la relación entre el tabaquismo, el consumo de alcohol y el cáncer de esófago. Se muestran cuatro estratos (niveles) de la cantidad fumada. Dentro de cada estrato de tabaquismo, se representa el riesgo de cáncer de esófago en relación con la cantidad de alcohol consumido.

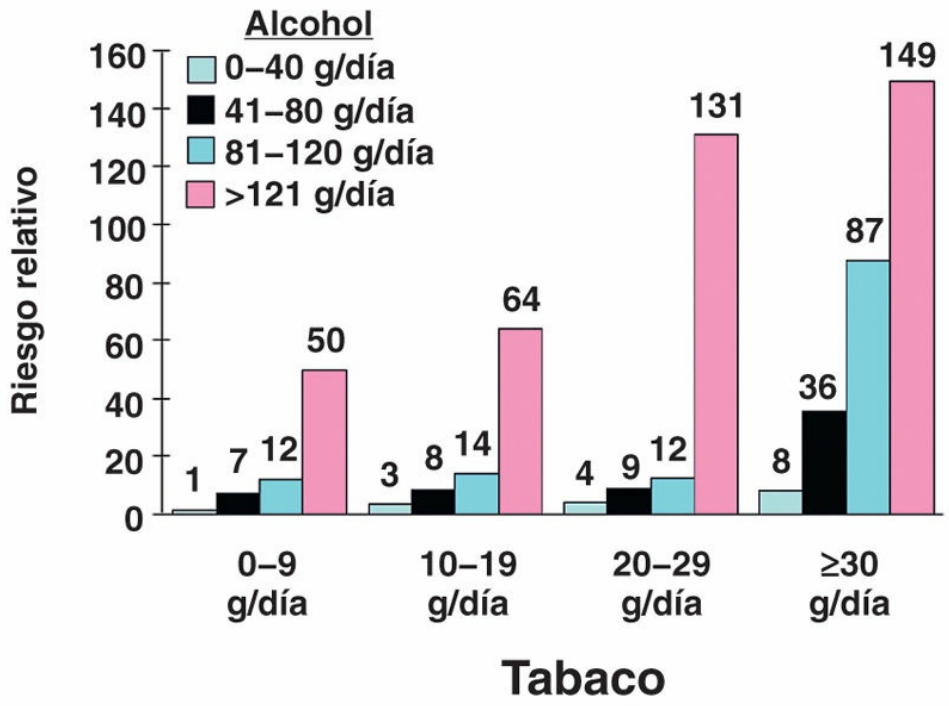


FIG. 15.5 Riesgo relativo de desarrollar cáncer de esófago en relación con el tabaquismo y el consumo de alcohol. (Modificada de Tuyns AJ, Pequignot G, Jensen OM. Esophageal cancer in Ille-et-Vilaine in relation to levels of alcohol and tobacco consumption: Risks are multiplying. *Bull Cancer.* 1977;64:45–60.)

¿Qué es lo que se observa? Cuanto más comunica que fuma un individuo, mayor es la tasa de cáncer de esófago. Sin embargo, dentro de cada estrato de tabaquismo, hay una relación dosis-respuesta entre el cáncer de esófago y la cantidad de alcohol consumido. Por tanto, no se pueden atribuir al tabaquismo los efectos del consumo de alcohol sobre el cáncer de esófago. Tanto el tabaquismo como el consumo de alcohol tienen efectos independientes sobre el riesgo de cáncer de esófago.

Es importante señalar que, en esta presentación de los datos, no es posible comparar fumadores con no fumadores o bebedores con no bebedores, porque los autores han juntado el grupo que fuma 0-9 gramos de tabaco al día, y también han juntado los no bebedores con bebedores mínimos. Por tanto, no se dispone de tasas para las personas realmente *no expuestas* al alcohol o al tabaco. Hubiera sido preferible mantener los datos de las personas no expuestas por separado, de manera que se hubiesen podido calcular los RR basándose en la tasas de las personas no expuestas. Mezclar los no bebedores con los bebedores moderados y los no fumadores con los fumadores moderados dificulta analizar adecuadamente los datos.

Para finalizar, dos últimas consideraciones sobre la confusión. En primer lugar, cuando se identifica un factor de confusión, se suele considerar que es un problema y se desea encontrar la manera de solucionar la cuestión de la confusión. Sin embargo, a veces el hallazgo de una relación con un factor de confusión también puede ser esclarecedor. Aunque la asociación aparente entre la exposición A (el factor que es el principal foco de interés) y la enfermedad B se deba en realidad a algún tercer factor de confusión X, de

modo que la exposición A no tenga una relación causal con la enfermedad B, el cribado de la exposición A puede ser útil, porque permite identificar a las personas que tienen un alto riesgo de contraer la enfermedad y aplicarles las intervenciones preventivas y terapéuticas adecuadas. Por tanto, una relación de confusión también puede ser una guía útil para el cribado de la población, incluso cuando no se identifique el agente etiológico específico involucrado.

En segundo lugar, la confusión no es un error del estudio, sino más bien un fenómeno auténtico que se identifica en el estudio y que debe conocerse. El sesgo es el resultado de un error metodológico del estudio, pero la confusión es un hallazgo válido que describe la naturaleza de la relación entre varios factores y el riesgo de la enfermedad. Sin embargo, *el hecho de no tener en cuenta la confusión al interpretar los resultados de un estudio* sí es un error metodológico del estudio y puede sesgar sus conclusiones.

Interacción

Hasta el momento, se ha asumido generalmente la presencia de una sola exposición (factor de riesgo) en la etiología de una enfermedad. Aunque este enfoque es útil para los propósitos de la exposición, en la vida real pocas veces nos enfrentamos a causas individuales. En los ejemplos anteriores de la relación entre el cáncer de pulmón y el tabaquismo y la urbanización, y la relación entre el cáncer de esófago y el consumo de alcohol y el tabaquismo, ya hemos visto que hay más de un factor implicado en la etiología de la enfermedad. En esta sección nos plantearemos la siguiente pregunta: ¿Cómo interactúan múltiples factores a la hora de causar una enfermedad?

¿Qué se entiende por *interacción*? MacMahon[19] definió la interacción de la siguiente manera: «Cuando la incidencia de la enfermedad en presencia de dos o más factores de riesgo difiere de la incidencia que sería previsible por sus efectos individuales». El efecto puede ser mayor de lo esperado (interacción positiva, sinergismo) o menor de lo esperado (interacción negativa, antagonismo). El problema consiste en determinar qué sería lo *esperado* de los efectos individuales de las exposiciones.

En la figura 15.6 se muestra un algoritmo para analizar la posibilidad de interacción.

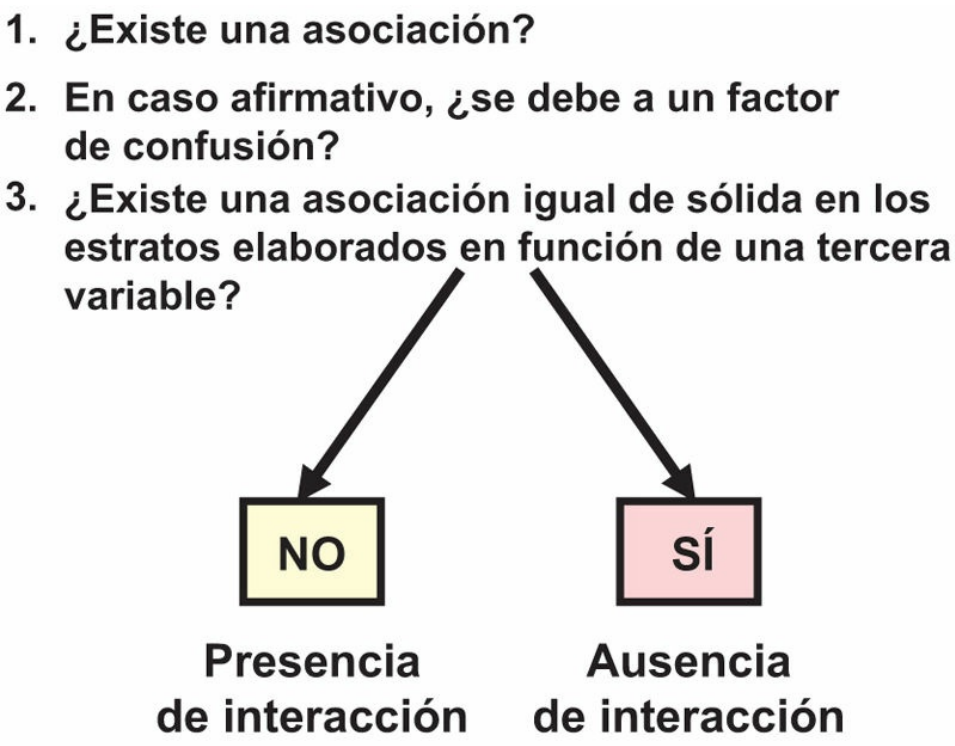


FIG. 15.6 Preguntas que deben plantearse respecto a la naturaleza de la relación entre la exposición y el resultado.

Al examinar los datos, la primera pregunta es si se ha observado una

asociación entre una exposición y una enfermedad. Si es así, ¿se debe a un fenómeno de confusión? Si decidimos que *no* se debe a confusión, es decir, que la asociación es causal, entonces hay que preguntarnos si la asociación es igual de sólida en cada uno de los estratos que se crean basándose en una tercera variable. Por ejemplo, ¿es la asociación entre el tabaquismo y el cáncer de pulmón igual de sólida en los estratos creados en función del grado de urbanización? Si la asociación es igual de sólida en todos los estratos, no hay interacción. Pero si la asociación tiene una solidez diferente en los distintos estratos creados en función de la edad, por ejemplo, si la asociación es más fuerte en las personas mayores que en las más jóvenes, se ha observado una interacción entre la edad y la exposición a la hora de producir la enfermedad. Si no hubiera interacción, sería de esperar que la asociación tuviese la misma solidez en todos los estratos.

A continuación, se analizará más en detalle la interacción. En la tabla 15.9 se muestra la incidencia en las personas expuestas a un factor de riesgo (A o B), a ambos o a ninguno de ellos en un ejemplo hipotético.

Tabla 15.9

Tasas de incidencia para los grupos expuestos a ningún factor de riesgo o a uno o dos factores de riesgo (datos hipotéticos)

		Factor A: −	Factor A: +
Factor B	−	3,0	9,0
	+	15,0	

En las personas con ninguna exposición, la incidencia es de 3,0. En las personas expuestas al factor A y no al factor B, la incidencia es de 9,0. En las personas expuestas al factor B y no al factor A, la incidencia es de 15,0. Estos son los efectos individuales de cada uno de los factores considerados por separado.

¿Cuál sería la incidencia esperada en las personas que están expuestas a los dos factores (A y B) (la celda inferior derecha de la tabla 15.9) si esas personas experimentasen el riesgo resultante de las contribuciones independientes de ambos factores? La respuesta depende del tipo de modelo que se proponga.

Supongamos que, cuando se producen dos exposiciones, el efecto de una exposición se *suma* al efecto de la segunda exposición, es decir, el modelo es *aditivo.* En tal caso, ¿qué sería previsible observar en la celda inferior derecha de la tabla? Como ejemplo, se utilizarán las personas que no tienen ninguna exposición, cuyo riesgo en ausencia de ambas exposiciones es de 3,0. ¿De qué manera la exposición al factor A afecta a su riesgo? Suma un riesgo de 6,0 a 3,0 para producir un riesgo de 9,0. Si el factor A suma un riesgo de 6,0 al riesgo que existe sin el factor A, debería tener el mismo efecto tanto en las personas expuestas como en las no expuestas al factor B. Debido a que el factor A suma 6,0 a 3,0, también sería de esperar que añadiese 6,0 a la tasa de 15,0 que presentan las personas expuestas al factor B cuando estas se exponen simultáneamente al factor A. Por tanto, sería de esperar que los efectos de la exposición a ambos factores diesen lugar a una incidencia de 21,0.

Esto también puede plantearse del siguiente modo: si el factor B suma 12,0 a la incidencia de 3,0 de las personas sin ninguna exposición, es de esperar que sume 12,0 a cualquier grupo, incluido el grupo expuesto solo al factor A, cuya incidencia es 9,0. Por tanto, el efecto esperado de la exposición conjunta a A y B sería 9,0 más 12,0, es decir, 21,0. (Hay que recordar que 3,0 es el riesgo de base que está presente en ausencia de A y B. Cuando se calcula el efecto combinado de los factores A y B, no se puede sumar simplemente 9,0 y 15,0, pues hay que asegurarse de que no se tiene en cuenta el riesgo de base [3,0] dos veces.) La parte izquierda de la tabla 15.10 muestra la tabla completa de los datos parciales presentados en la tabla 15.9.

Tabla 15.10

Tasas de incidencia y riesgos atribuibles para los grupos expuestos a ningún factor de riesgo o a uno o dos factores de riesgo (datos hipotéticos en un modelo aditivo: I)

		Tasas de incidencia		Riesgos atribuibles	
		Factor A		Factor A	
		−	+	−	+
Factor B	−	3,0	9,0	0	6
	+	15,0	21,0	12	

Hay que recordar que, al comentar las diferencias en cuanto a riesgos, nos referimos a los *riesgos atribuibles.* Esto se muestra en la parte derecha de la

tabla 15.10. Si se analizan las personas que no tienen ninguna exposición, presentan un riesgo de base, pero el riesgo atribuible (el que puede atribuirse a la exposición al factor A o B) es 0. Como ya se ha indicado, la exposición solo al factor A suma 6, y la exposición solo al factor B suma 12. ¿Cuál será el riesgo atribuible a ambas exposiciones? La respuesta es 18, es decir, 18 más que el riesgo de base. El modelo aditivo se resume en la tabla 15.11.

Tabla 15.11

Tasas de incidencia y riesgos atribuibles para los grupos expuestos a ningún factor de riesgo o a uno o dos factores de riesgo (datos hipotéticos en un modelo aditivo: II)

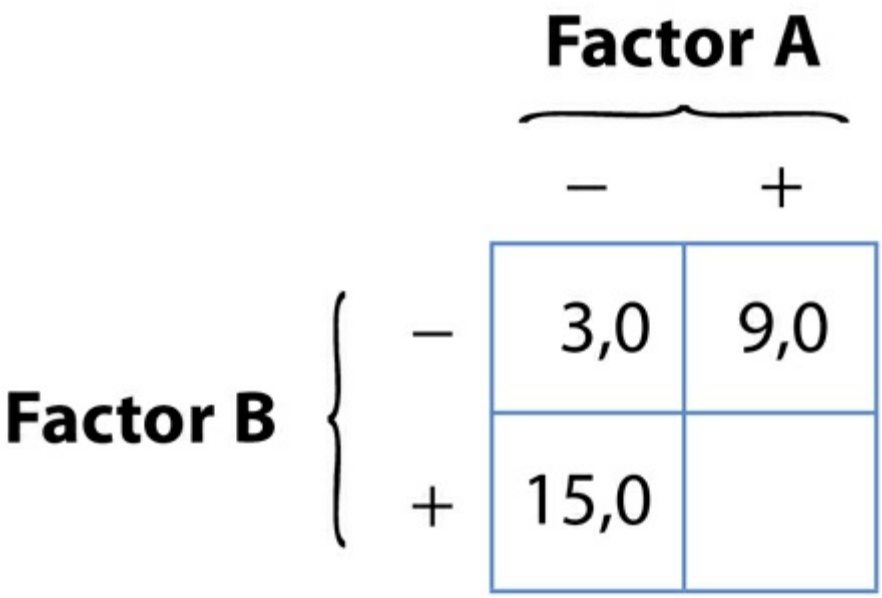
		Tasas de incidencia		Riesgos atribuibles	
		Factor A		Factor A	
		−	+	−	+
Factor B	−	3,0	9,0	0	6
	+	15,0	21,0	12	18

¿Qué sucede si un modelo aditivo no describe correctamente el efecto de la exposición a dos factores independientes? Tal vez una segunda exposición no *se sume* al efecto de la primera exposición, sino que *multiplique* el efecto de la primera. Si tener una exposición determinada duplica el riesgo de una persona, sería de esperar que duplicase ese riesgo, con independencia de si esa persona tuviese o no otra exposición. Por ejemplo, si el alcohol tiene el efecto de duplicar el riesgo de una persona para un cierto tipo de cáncer, se podría esperar que duplicase el riesgo para los fumadores y los no fumadores. El modelo apropiado para los efectos de dos factores independientes podría ser, por tanto, *multiplicativo* en lugar de aditivo.

Volvamos a los datos originales sobre el riesgo resultante de ninguna exposición o de la exposición al factor A o B. Estos datos se muestran de nuevo en la tabla 15.12.

Tabla 15.12

Tasas de incidencia para los grupos expuestos a ningún factor de riesgo o a uno o dos factores de riesgo (datos hipotéticos)



Puede observarse que la exposición al factor A triplica el riesgo, en comparación con lo que se observa cuando el factor A está ausente (9,0 frente a 3,0). ¿Qué sería previsible encontrar en la celda inferior derecha de la tabla cuando ambas exposiciones estuviesen presentes? Dado que, en ausencia del factor B, el factor A ha triplicado el riesgo de 3,0, también sería de esperar que triplicase el riesgo de 15,0 observado cuando la exposición al factor B está presente. En tal caso, el efecto de la exposición a ambos factores sería 45,0. También en este caso se puede calcular esto de una manera diferente. El factor B multiplica el riesgo por 5 (15,0 frente a 3,0) cuando el factor A está ausente. Por tanto, sería de esperar que tuviese el mismo efecto cuando el factor A está presente. Debido a que el riesgo cuando el factor A está presente es 9,0, sería previsible que la presencia del factor B diese lugar a un riesgo de 45,0 (9,0 × 5; tabla 15.13).

Tabla 15.13

Tasas de incidencia y riesgos relativos para los grupos expuestos a ningún factor de riesgo o a uno o dos factores de riesgo (datos hipotéticos en un modelo multiplicativo: I)

		Tasas de incidencia		Riesgos relativos	
		Factor A		Factor A	
		−	+	−	+
Factor B	−	3,0	9,0	1	3
	+	15,0	45,0	5	

La parte izquierda de la tabla 15.13 muestra los datos de incidencia completos. Esta descripción del modelo multiplicativo corresponde a un *modelo de RR*. Esto se muestra en la parte derecha de la tabla. ¿Qué valor sería previsible observar en la celda en blanco?

Si ahora se asigna un valor de 1 al riesgo de base (3,0), frente al que comparar los demás valores de la tabla, la exposición al factor A triplica el riesgo y se obtiene un RR de 3 para el factor A en ausencia del factor B. El factor B multiplica el riesgo por 5, lo que produce un RR de 5 para la exposición al factor B en ausencia del factor A. Cuando actúan ambos factores (A y B), sería previsible observar un RR de 15 (45,0/3,0), como se ve en la parte izquierda, o 3 × 5, como se ve en la parte derecha de la tabla 15.14.

Tabla 15.14

Tasas de incidencia y riesgos relativos para los grupos expuestos a ningún factor de riesgo o a uno o dos factores de riesgo (datos hipotéticos en un modelo multiplicativo: II)

		Tasas de incidencia		Riesgos relativos	
		Factor A		Factor A	
		−	+	−	+
Factor B	−	3,0	9,0	1	3
	+	15,0	45,0	5	15

Hasta aquí se han considerado dos modelos, aditivo y multiplicativo, pero siguen pendientes estas cuestiones: ¿Qué sería previsible observar como resultado de los efectos independientes de dos factores de riesgo? ¿Sería de esperar un modelo aditivo o un modelo multiplicativo?

Puede que las respuestas no sean evidentes. Si intervienen dos factores y la incidencia es 21,0, el resultado concuerda con un modelo aditivo. Si la incidencia es 45,0, el resultado concuerda con un modelo multiplicativo. Sin embargo, si la incidencia resultante de los dos factores es 60,0, por ejemplo, se supera claramente incluso el valor para un modelo multiplicativo, y existe una interacción (es decir, un efecto mayor del que sería de esperar de los efectos independientes de ambos factores por separado).

Sin embargo, si la incidencia es 30,0, es menos de lo esperado según un modelo multiplicativo, pero aún más de lo esperado según un modelo aditivo. Entonces surge la pregunta: «¿Es esto más de lo que cabría esperar de los efectos independientes de los dos factores?» Es difícil saber la respuesta sin más información sobre la biología de la enfermedad, los mecanismos implicados en su patogenia, y el modo en el que dichos factores actúan a nivel celular y molecular. La mayoría de los expertos aceptan que cualquier efecto mayor que el aditivo es una evidencia de una interacción positiva, también denominada *sinergismo.* Sin embargo, esta opinión suele basarse en consideraciones estadísticas, mientras que la validez del modelo debería basarse idealmente en el conocimiento biológico. El modelo puede ser diferente entre las distintas enfermedades y exposiciones.

A continuación, se presentarán algunos ejemplos. En un estudio de cohortes sobre el tabaquismo y el cáncer de pulmón, Hammond y cols.[20] estudiaron el riesgo de cáncer de pulmón respecto al tabaquismo en 17.800 trabajadores del amianto en Estados Unidos y en 73.763 varones que no estuvieron expuestos al amianto. En la tabla 15.15 se muestran los datos de mortalidad por cáncer de pulmón en relación con la exposición. Si la relación entre el tabaquismo y la exposición al amianto fuese aditiva, sería de esperar

que el riesgo en los expuestos tanto al tabaco como al asbesto (la celda inferior derecha) fuese 58,4 + 122,6 − 11,3, es decir, 169,7. (Recuerde que el riesgo de base 11,3 se resta para no contabilizarlo dos veces.) Es evidente que el valor observado de 601,6 es mucho mayor que el valor aditivo esperado. De hecho, los datos de la tabla 15.15 se aproximan mucho a un modelo multiplicativo y sugieren con fuerza la existencia de sinergismo entre la exposición al asbesto y el tabaquismo.

Tabla 15.15

Fallecimientos por cáncer de pulmón (por 100.000) entre los individuos con y sin exposición al tabaquismo y el amianto

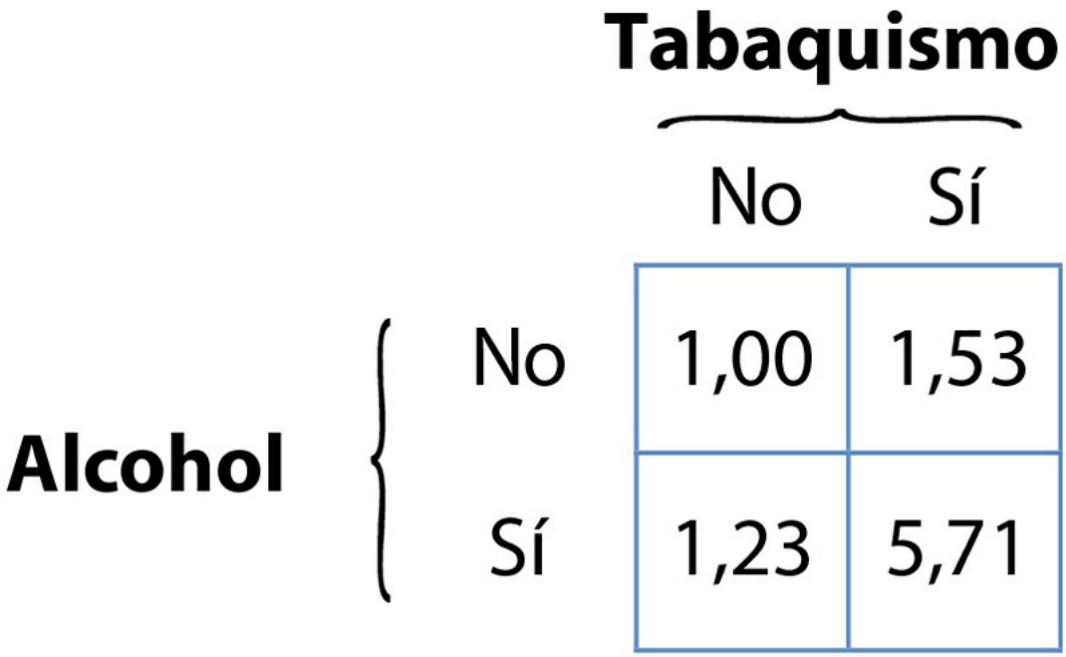
	EXPOSICIÓN AL AMIANTO	
Tabaquismo	**No**	**Sí**
No	11,3	58,4
Sí	122,6	601,6

Modificada de Hammond EC, Selikoff IJ, Seidman H. Asbestos exposure, cigarette smoking and death rates. *Ann NY Acad Sci.* 1979;330:473–490.

Un segundo ejemplo se observa en la tabla 15.16, que muestra el RR de cáncer oral en función de la presencia o ausencia de dos exposiciones: el tabaquismo y el consumo de alcohol. El riesgo se fija en 1,00 para las personas con ninguna exposición. ¿Hay evidencia de una interacción? ¿Cuál sería el riesgo previsible si el efecto fuese multiplicativo? Sería de esperar un valor de 1,53 × 1,23, es decir, 1,88. Es evidente que el efecto observado de 5,71 es mayor que un efecto multiplicativo e indica la existencia de interacción.

Tabla 15.16

Riesgos relativos[a] de cáncer oral según la presencia o ausencia de dos exposiciones: tabaquismo y consumo de alcohol



[a] Los riesgos se expresan respecto a un riesgo de 1,00 para las personas que no fumaban ni bebían alcohol.

De Rothman K, Keller A. The effect of joint exposure to alcohol and tobacco on risk of cancer of the mouth and pharynx. *J Chronic Dis.* 1972;25:711–716.

Ahora se analizarán unos datos más detallados de estas relaciones usando datos de dosis para el consumo de alcohol y el tabaquismo (tabla 15.17).

Tabla 15.17

Riesgo relativo[a] de cáncer oral según el nivel de exposición al alcohol y al tabaco: I

Consumo de alcohol (g/día)	EQUIVALENTES DE CIGARRILLOS/DÍA			
	0	<20	20–39	≥40
0	1,00	1,52	1,43	2,43
<11	0,40	1,67	3,18	3,25
11-42	1,60	4,36	4,46	8,21
>42	2,33	4,13	9,59	15,50

[a] Los riesgos se expresan en relación con un riesgo de 1,00 para las personas que no fumaban ni bebían alcohol.

De Rothman K, Keller A. The effect of joint exposure to alcohol and tobacco on risk of cancer of the mouth and pharynx. *J Chronic Dis.* 1972;25:711–716.

De nuevo, el riesgo en las personas que no beben y no fuman se ha fijado en 1,0. En las personas con el nivel más alto tanto de consumo de alcohol como de tabaquismo, el riesgo es 15,50. ¿Esto indica una interacción? Los

datos parecen apoyar esto. Los valores más altos en fumadores que son no bebedores y en bebedores que son no fumadores son 2,43 y 2,33, respectivamente; el valor de 15,5 supera claramente el producto resultante de 5,66 que sería de esperar si hubiese un efecto multiplicativo.

Sin embargo, hay que mencionar un problema con estos datos. Se debe tener en cuenta que cada categoría de tabaquismo o de consumo de alcohol tiene unos límites superior e inferior, excepto las categorías más altas, que no tienen límites superiores. Por tanto, el alto riesgo de 15,50 podría deberse a la presencia de alguno de los valores atípicos extremos, o bien fumadores muy empedernidos o bien personas con un consumo muy elevado de alcohol.

¿Hay alguna manera de evitar este problema y seguir utilizando los datos que aparecen aquí? Se podría ignorar la columna de la derecha y la fila inferior y utilizar solo la tabla resultante de 3 × 3 (tabla 15.18). Ahora todas las categorías tienen ambos límites (superior e inferior). Si el modelo fuese multiplicativo, sería previsible observar un valor de 1,43 × 1,60, es decir, 2,29, en lugar del observado en realidad de 4,46. Por tanto, todavía se aprecia una evidencia de interacción, pero mucho más débil de lo que se observaba en la tabla completa, con sus categorías de alta exposición indefinidas. Esto sugiere que el problema de la falta de límites superiores de las categorías contribuía al elevado valor de 15,50 observado en la tabla de 4 × 4.

Tabla 15.18

Riesgo relativo[a] de cáncer oral según el nivel de exposición al alcohol y al tabaco: II

-->

Consumo de alcohol (g/día)	EQUIVALENTES DE CIGARRILLOS/DÍA			
	0	<20	20–39	≥40
0	1,00	1,52	1,43	2,43
<11	0,40	1,67	3,18	3,25
11-42	1,60	4,36	4,46	8,21
>42	2,33	4,13	9,59	15,50

[a] Los riesgos se expresan en relación con un riesgo de 1,00 para las personas que no fumaban ni bebían alcohol.

De Rothman K, Keller A. The effect of joint exposure to alcohol and tobacco on risk of cancer of the mouth and pharynx. *J Chronic Dis.* 1972;25:711–716.

Como ya se ha comentado, la decisión de si un modelo aditivo o multiplicativo es más relevante en una situación concreta debería depender de la biología de la enfermedad. En la tabla 15.19 se muestran datos respecto a los riesgos de cáncer secundarios a la radiación y el tabaquismo en dos poblaciones diferentes: los trabajadores del uranio (izquierda) y los

supervivientes del bombardeo atómico de Japón de 1945 (derecha). En la tabla 15.19 se muestran los niveles bajos y altos de tabaquismo y los niveles bajos y altos de radiación para ambos grupos.

Tabla 15.19

Riesgos relativos de cáncer de pulmón según el tabaquismo y la exposición a la radiación en dos poblaciones

Nivel de radiación	TRABAJADORES DEL URANIO (NIVEL DE TABAQUISMO)		SUPERVIVIENTES DE LA BOMBA ATÓMICA (NIVEL DE TABAQUISMO)	
	Bajo	Alto	Bajo	Alto
Bajo	1,0	7,7	1,0	9,7
Alto	18,2	146,8	6,2	14,2

De Blot WJ, Akiba S, Kato H. Ionizing radiation and lung cancer: A review including preliminary results from a case-control study among A-bomb survivors. En Prentice RL, Thompson DJ, eds. *Atomic Bomb Survivor Data: Utilization and Analysis.* Philadelphia: Society for Industrial and Applied Mathematics; 1984:235-248.

¿Qué tipo de modelo sugiere la tabla 15.19? Es evidente que apunta hacia una relación multiplicativa; 146,8 se aproxima al producto de 7,7 × 18,2 (140,1). Las columnas de la derecha sugieren un modelo aditivo; 14,2 está próximo a la suma de 9,7 + 6,2 − 1,0 (14,9). Por tanto, aunque los datos tienen en cuenta la radiación y el tabaquismo en dos poblaciones, en un contexto, las exposiciones se relacionan de modo aditivo y, en el otro, lo hacen de una manera multiplicativa. No se sabe si esto se debe a las diferencias en cuanto a la exposición a la radiación en las minas de uranio en comparación con la de las bombas atómicas. Esta hipótesis no es irracional; se sabe que hubo incluso una diferencia en cuanto a la radiación emitida por las bombas atómicas en Hiroshima y Nagasaki, y que las curvas de dosis-respuesta para el cáncer fueron diferentes en ambas ciudades. Además, la ubicación de los supervivientes en cada ciudad puede haber variado. En cualquier caso, el hecho de que dos exposiciones que son aparentemente la misma (o, al menos, similares) puedan tener interrelaciones diferentes en distintos contextos es una observación intrigante que requiere una mayor evaluación.

Por último, un ejemplo muy llamativo de interacción es el de la relación de la aflatoxina y la infección crónica por hepatitis B con el riesgo de cáncer de hígado (tabla 15.20). En este estudio chino, la infección por hepatitis B sola multiplica el riesgo de cáncer de hígado por 7,3; la exposición a la aflatoxina sola multiplica el riesgo por 3,4. Sin embargo, cuando ambas exposiciones estaban presentes, el RR se elevó a 59,4, muy por encima de lo que podría esperarse tanto en un modelo aditivo como en un modelo multiplicativo[a] . Esta observación de sinergia tiene un gran interés clínico y de salud pública,

pero también sugiere direcciones importantes para futuras investigaciones de laboratorio sobre la etiología y la patogenia del cáncer de hígado.

Tabla 15.20

Riesgos[a] de cáncer de hígado en personas expuestas a la aflatoxina o con infección crónica por hepatitis B: un ejemplo de interacción

	Aflatoxina- negativo	Aflatoxina- positivo
HBsAg negativo	1,0	3,4
HBsAg positivo	7,3	59,4

HBsAg, antígeno de superficie de la hepatitis B.

[a] Ajustado para el tabaquismo.

Modificada de Qian GS, Ross RK, Yu MC, et al. A follow-up study of urinary markers of aflatoxin exposure and liver cancer risk in Shanghai, People's Republic of China. *Cancer Epidemiol Biomarkers Prev.* 1994;3:3–10.

El hallazgo de interacción aditiva o sinergismo también puede tener implicaciones políticas prácticas respecto a cuestiones tales como quién es el responsable de una enfermedad y quién debería indemnizar a las víctimas. Por ejemplo, al principio de este capítulo se comentó la relación del tabaquismo y la exposición al asbesto con el desarrollo de cáncer, una relación que es claramente muy interactiva o sinérgica. Las demandas contra los fabricantes de amianto se remontan al menos a la década de 1970 y los tribunales fijaron indemnizaciones muy cuantiosas. En 1998, en un momento en el que las acciones legales contra las compañías tabacaleras eran cada vez más frecuentes, una asociación de algunas de las víctimas de la exposición al amianto se alió con los fabricantes de amianto para exigir al Congreso de EE. UU. que destinase una gran cantidad de dinero de cualquier acuerdo económico nacional sobre el tabaco para compensar a las personas cuyo cáncer estuviese causado por la exposición combinada al asbesto y el tabaco, exigencia que justificaron por la relación sinérgica de estas exposiciones. Los que se opusieron a esta demanda alegaron que quienes la habían interpuesto estaban eximiendo a los fabricantes de amianto de cumplir con su obligación (cambiándola a los fabricantes de tabaco) y que solo lo hacían porque creían que podría ser más fácil obtener una compensación significativamente más elevada de las empresas tabacaleras que de los fabricantes de amianto. De este modo, estaban dispuestos a aliarse con los fabricantes de amianto que previamente habían sido encontrados responsables de su enfermedad. El fundamento de esta estrategia era el sinergismo bien documentado entre el asbesto y el tabaquismo a la hora de causar cáncer.

Conclusión

En este capítulo se han revisado los conceptos de sesgo, confusión e interacción en relación con la derivación de inferencias causales. Los sesgos reflejan deficiencias del diseño o la realización de un estudio y afectan claramente a la validez de los hallazgos. Los sesgos deben, por tanto, evaluarse y, si es posible, eliminarse. La confusión y la interacción, por otro lado, describen la realidad de las interrelaciones entre ciertas exposiciones y un determinado resultado de la enfermedad. La confusión y la interacción caracterizan casi todas las situaciones en las que se aborda la etiología, porque la mayoría de las preguntas sobre causalidad implican las relaciones de varias exposiciones y de múltiples factores posiblemente etiológicos. Estas relaciones son fundamentales a la hora de investigar los papeles de los factores genéticos y ambientales en la etiología de la enfermedad y de asignar la responsabilidad de los resultados de salud adversos causados por exposiciones ambientales. La evaluación de las contribuciones relativas de los factores genéticos y ambientales se describe en el siguiente capítulo.

Bibliografía

1. Schlesselman JJ. Case-Control Studies: Design, Conduct, and Analysis. New York: Oxford University Press; 1982.
2. Johson TJ. Handbook of Health Survey Methods. Hoboken, NJ: John Wiley & Sons; 2015.
3. Ronmark E, Lundqvist A, Lundback B, et al. Non-responders to a postal questionnaire on respiratory symptoms and diseases. *Eur J Epidemiol*. 1999;15:293–299.
4. Collins R. What make the UK Biobank special? *Lancet*. 2012;379:1173–1174.
5. Pinsky PF, Prorok PC, Yu K, et al. Extended mortality results for prostate cancer screening in the PLCO trial with median follow-up of 15 years. *Cancer*. 2017;123:592–599.
6. Loprinzi PD, Davis RE. Socioecological risk predictors of physical activity and associated mortality. *Am J Health Promot*. 2018;32:106–111.
7. Dobson AJ, Hockey R, Brown WJ, et al. Cohort profile update: Australian Longitudinal Study on Women's Health. *Int J Epidemiol*. 2015;44: 1547a–1547f.
8. Koton S, Schneider AL, Rosamond WD, et al. Stroke incidence and mortality trends in US communities, 1987-2011. *JAMA*. 2014;312:259–268.
9. Boston Collaborative Drug Surveillance Program Reserpine and breast cancer. *Lancet*. 1974;2:669–671.
10. Armstrong B, Stevens B, Doll R. Retrospective study of the association between use of Rauwolfia derivatives and breast cancer in English women. *Lancet*. 1974;2:672–675.
11. Heinonen OP, Shapiro S, Tuominen L, et al. Reserpine use in relation to breast cancer. *Lancet*. 1974;2:675–677.
12. Horwitz RI, Feinstein AR. Exclusion bias and the false relationship of reserpine and breast cancer. *Arch Intern Med*. 1985;145:1873–1875.
13. MacMahon B, Yens S, Trichopoulos D, et al. Coffee and cancer of the pancreas. *N Engl J Med*. 1981;304:630–633.
14. Metzger DS, Koblin B, Turner C, et al. Randomized controlled trial of audio computer-assisted self-interviewing: utility and acceptability in longitudinal studies. *Am J Epid*. 2000;152:99–107.

15. Macalino GE, Celentano DD, Latkin C, et al. Risk behaviors by audio computer-assisted self-interviews among HIV-seropositive and HIV-seronegative injection drug users. *AIDS Educ Prev*. 2002;14:367–378.
16. Wynder EL, Higgins IT, Harris RE. The wish bias. *J Clin Epidemiol*. 1991;43:619–621.
17. Rookus MA, van Leeuwen FE. Induced abortion and risk for breast cancer: reporting (recall) bias in a Dutch case-control study. *J Natl Cancer Inst*. 1996;88:1759–1764.
18. D'Avanzo B, LaVecchia C, Katsouyanni K, et al. Reliability of information on cigarette smoking and beverage consumption provided by hospital controls. *Epidemiology*. 1996;7:312–315.
19. MacMahon B. Concepts of multiple factors. In: Lee DH, Kotin P, eds. Multiple Factors in the Causation of Environmentally Induced Disease. New York: Academic Press; 1972.
20. Hammond EC, Selikoff IJ, Seidman H. Asbestos exposure, cigarette smoking and death rates. *Ann NY Acad Sci*. 1979;330:473–490.

Preguntas de repaso del capítulo 15

1. *¿Cuál de las siguientes respuestas es una estrategia para manejar la confusión?*

a. Emparejamiento individual.

b. Estratificación.

c. Emparejamiento de grupo.

d. Ajuste.

e. Todas las anteriores.

2. *¿Cuál de los siguientes enfoques puede ocuparse de la confusión en la etapa de diseño del estudio?*

a. Estratificación.

b. Ajuste.

c. Restricción.

d. Regresión.

3. *Se ha sugerido que los médicos pueden explorar a las mujeres que usan anticonceptivos orales con más frecuencia o más a fondo que a las mujeres*

que no los usan. Si es así, y si se observa una asociación entre la flebitis y el uso de anticonceptivos orales, la asociación puede deberse a:

a. Sesgo de selección.

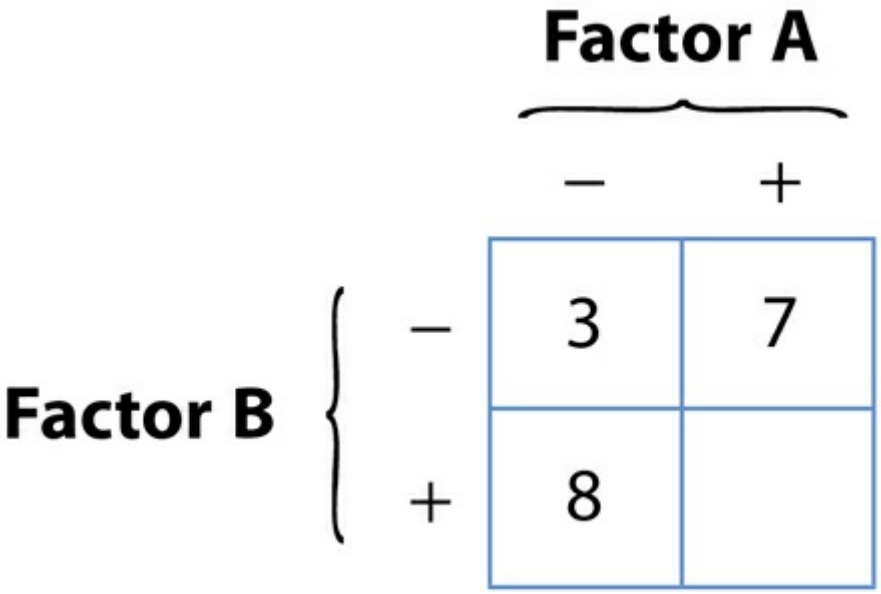
b. Sesgo del entrevistador.

c. Sesgo de vigilancia.

d. Sesgo de ausencia de respuesta.

e. Sesgo de recuerdo.

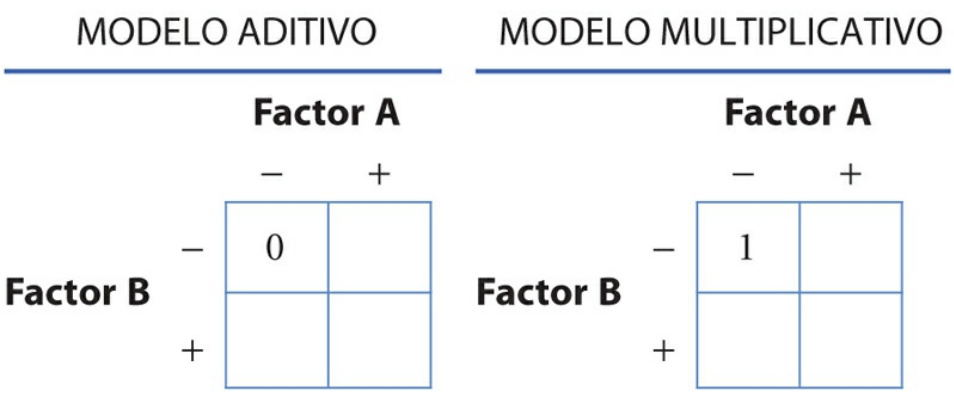
Las preguntas 4 a 7 se basan en la información que se indica a continuación:



4. *Rellene la celda vacía de la tabla usando el modelo aditivo de interacción:* _____

5. *Rellene la celda vacía de la tabla usando el modelo multiplicativo de interacción:* _____

Convierta los números de la tabla siguiente en riesgos atribuibles para el modelo aditivo (abajo, a la izquierda) y en riesgos relativos para el modelo multiplicativo (abajo, a la derecha).



6. *Rellene la celda inferior derecha de la tabla en la parte inferior de la columna izquierda con el riesgo atribuible de tener ambos factores A y B (modelo aditivo):* _____

7. *Rellene la celda inferior derecha de la tabla en la parte inferior de la columna izquierda con el riesgo relativo de presentar ambos factores A y B (modelo multiplicativo):* _____

La pregunta 8 se basa en la información que se indica a continuación:

En un estudio de casos y controles sobre la relación de la exposición a la radiación y el cáncer de tiroides, se estudiaron 50 casos ingresados por cáncer de tiroides y 100 «controles» ingresados durante el mismo periodo para el tratamiento de hernias. Solo se entrevistó a los casos, y se observó que 20 de ellos habían sido expuestos a radioterapia en el pasado, según las entrevistas y las historias clínicas. Los controles no fueron entrevistados, pero una revisión de sus historias clínicas cuando fueron ingresados para cirugía de hernia reveló que solo dos controles habían sido expuestos a radioterapia en el pasado.

8. *Basándose en la descripción indicada, ¿qué fuente de sesgo es menos probable que se presente en este estudio?*

a. Sesgo de recuerdo.

b. Sesgo debido a que los controles no son representativos de la población no enferma.

c. Sesgo debido al uso de diferentes métodos de evaluación de la exposición en los casos y los controles.

d. Sesgo debido a la pérdida de sujetos del grupo control con el tiempo.

e. Sesgo de selección para la exposición a la radioterapia en el pasado.

9. *En 1990 se realizó un estudio de casos y controles para investigar la asociación positiva entre el uso de edulcorantes artificiales y el cáncer de vejiga. Los controles fueron seleccionados de una muestra hospitalaria de los pacientes diagnosticados de trastornos relacionados con la obesidad. Dichos trastornos relacionados con la obesidad se han asociado positivamente con el uso de edulcorantes artificiales. ¿Cómo afectaría el uso de estos pacientes como controles a la hora de estimar la asociación entre el uso de edulcorantes artificiales y el cáncer de vejiga?*

a. La estimación de la asociación reflejaría con precisión la verdadera asociación, con independencia de la asociación entre el uso de edulcorantes artificiales y los trastornos relacionados con la obesidad.

b. La estimación de la asociación tendería a subestimar la verdadera asociación.

c. Se necesita más información sobre la fuerza de la asociación entre el uso de edulcorantes artificiales y los trastornos relacionados con la obesidad antes de poder realizar algún juicio.

d. La estimación de la asociación tendería a sobrestimar la verdadera asociación.

e. Se necesita más información sobre la fuerza de la asociación entre el uso de edulcorantes artificiales y el cáncer de vejiga antes de poder realizar algún juicio.

---

[a] Una fórmula simple para calcular el cociente de posibilidades (OR, *odds ratio*) o el riesgo relativo (RR) aditivo conjunto esperado es [RR u OR para el efecto aislado del factor A] + [RR u OR para el efecto aislado del factor B] − 1,0. En el ejemplo, el RR aditivo conjunto esperado es 7,3 + 3,4 − 1,0 = 9,7, que es mucho más bajo que el RR observado de 59,4, lo que define así una fuerte interacción aditiva. (Este enfoque es particularmente útil en los estudios de casos y controles, en los que no se dispone de diferencias absolutas en las tasas.) La fórmula para el RR u OR multiplicativo conjunto esperado se basa en la multiplicación de los RR u OR independientes; en el ejemplo, 7,3 × 3,4 = 24,8, que también es mucho más bajo que el RR observado de 59,4. Por lo tanto, se puede concluir que existen interacciones tanto aditivas como multiplicativas.

CAPÍTULO 16

Identificación de los factores genéticos y ambientales en la causalidad de la enfermedad

Palabras clave

epidemiología genética
enfermedades mendelianas
enfermedades complejas
estudios de gemelos
estudios de adopción
estudios en inmigrantes
mapeo genético
análisis de vinculación
estudios de asociación de genoma completo
interacción de factores genéticos y ambientales
medicina de precisión

Para crear otro Wolfgang Amadeus Mozart, necesitaríamos no solo su genoma, sino también el útero de su madre, las lecciones de música de su padre, los amigos de sus padres y los suyos, el estado de la música en el siglo XVIII en Austria, la amistad de Haydn, y así sucesivamente, en círculos cada vez más amplios. Sin el conjunto de los genes de Mozart, el resto no sería suficiente; en cualquier caso, solo hubo un Wolfgang Amadeus Mozart. Pero no tenemos derecho a la presunción inversa: que su genoma, desarrollado en otro mundo y en otro momento, habría dado lugar al mismo genio de la música. Si una cepa particular de trigo produce diferentes cosechas en diferentes condiciones de clima, suelo y cultivo, ¿cómo se puede suponer que un genoma mucho más complejo como el de un ser humano podría producir su cosecha deseada de óperas, sinfonías y música de cámara bajo diferentes circunstancias de crianza?[1]

—Leon Eisenberg, MD, DSc, psiquiatra infantil, psiquiatra social y educador médico (1922-2009)

Objetivos de aprendizaje

- Examinar cómo los diseños de estudios

epidemiológicos pueden aclarar los papeles de los factores genéticos y ambientales en el riesgo de enfermedad y sus posibles interacciones.

- Mostrar cómo se utilizan los marcadores genéticos para mapear los genes que controlan el riesgo de diferentes enfermedades, incluidas enfermedades complejas.
- Probar la interacción entre genes y factores de riesgo ambientales.
- Comentar cómo la combinación de innovadores métodos epidemiológicos y de biología molecular, incluido el mapeo del genoma humano, puede ayudar a definir los papeles etiológicos de los factores de riesgo genéticos y ambientales, y cómo puede permitir el desarrollo de tratamientos individualizados para las personas con enfermedades graves.

En los capítulos anteriores se han descrito los diseños de estudios para identificar las causas de la enfermedad centrando la exposición principalmente en el posible papel etiológico de los factores ambientales. Sin embargo, para prevenir la enfermedad, también hay que tener en cuenta el papel que desempeñan los factores genéticos del paciente y los factores ambientales. Los seres humanos presentan diferencias evidentes entre sí en cuanto a características físicas, personalidad y otros factores. Estos fenotipos pueden ser cualitativos o cuantitativos. Un fenotipo cualitativo es un rasgo que se puede clasificar en dos (o más) grupos mutuamente excluyentes (p. ej., afectados o no afectados; no afectados, afectados leve o gravemente). Por ejemplo, un individuo puede ser clasificado dicotómicamente como hipertenso o no hipertenso. El fenotipo cuantitativo subyacente consiste en la medición continua de la presión arterial de la persona. La epidemiología genética pretende comprender la contribución de la genética del huésped en enfermedades complejas y fenotipos cuantitativos, e identificar su importancia relativa. Al final de este capítulo se incluye un glosario de términos genéticos.

Genética tradicional

La genética médica tradicional se ha centrado principalmente en los rasgos de un solo gen que siguen los patrones de transmisión descritos por Gregor Mendel, un monje austríaco del siglo XIX. Las enfermedades mendelianas son generalmente raras en la población y pueden clasificarse por su transmisión como autosómica dominante, autosómica recesiva, dominante ligada a X o recesiva ligada a X. Algunas enfermedades mendelianas, como, por ejemplo, la fibrosis quística (la enfermedad autosómica recesiva más común en poblaciones de ascendencia del norte de Europa, con una prevalencia de 1/2.500 nacidos vivos entre los blancos no hispanos en los Estados Unidos) y la enfermedad de células falciformes (la hemoglobinopatía más común entre poblaciones de ascendencia de África occidental con una prevalencia de 1/360 nacidos vivos entre los afroamericanos) se deben casi siempre a mutaciones en un solo gen (respectivamente *CFTR* en el cromosoma 7 y *beta globina* en el cromosoma 11). Aunque la fibrosis quística fue una enfermedad letal de la infancia durante la mayor parte de la historia humana debido a sus graves efectos nutricionales y pulmonares, en la actualidad los pacientes sobreviven hasta la edad adulta con la atención médica adecuada. En el gen *CFTR* se han identificado cientos de mutaciones diferentes, pero una eliminación de tres pares de bases (Δ508) es, con mucho, la más común. La detección de portadores seguida por un asesoramiento genético apropiado puede reducir la carga de esta enfermedad en algunas poblaciones[1], mientras que la detección selectiva en los recién nacidos ayuda en la identificación temprana y el tratamiento de las personas afectadas en una etapa temprana de la vida. La enfermedad de células falciformes es una enfermedad debilitante con una tasa de mortalidad general elevada, pero no es uniformemente letal[2]. La detección selectiva universal de la enfermedad de células falciformes se introdujo por primera vez en los Estados Unidos en 1975[3], lo que permitió el diagnóstico temprano, el inicio de la profilaxis con penicilina y la atención integral de los niños afectados, lo que ha reducido significativamente la mortalidad infantil por esta enfermedad[4].

Si bien las enfermedades mendelianas no son lo suficientemente comunes como para considerarse una carga importante de salud pública en todas las poblaciones, cuando todas las enfermedades mendelianas se consideran en conjunto, representan un importante problema de salud pública, especialmente para el sistema de atención médica pediátrica. Además, los programas de cribado de recién nacidos se han utilizado durante más de 50 años para identificar lactantes con errores congénitos del metabolismo (muchos de los cuales reflejan enfermedades mendelianas), y estos programas funcionan en la actualidad en la mayoría de los países desarrollados y en muchos países en vías de desarrollo. Por lo tanto, aunque cada una de las 30 a 60 enfermedades comúnmente incluidas en los programas actuales de

detección selectiva de recién nacidos en los Estados Unidos sea bastante rara, la sensibilidad y especificidad de las pruebas de cribado son lo suficientemente altas y los costos son lo suficientemente bajos, como para que estas se consideren una herramienta valiosa para la salud pública[5]. Por ejemplo, la fenilcetonuria se caracteriza por una deficiencia genéticamente determinada de fenilalanina hidroxilasa, y un niño homocigoto para esta mutación no puede metabolizar el aminoácido esencial fenilalanina. Los niveles excesivos de fenilalanina resultantes conducen a un retraso mental grave. Si bien no podemos prevenir la aparición de este genotipo, mediante el cribado de recién nacidos podemos reducir o eliminar la fenilalanina en la dieta de los niños afectados y, por lo tanto, prevenir el retraso mental grave. Por lo tanto, desde los puntos de vista tanto de la medicina clínica como de la salud pública, los programas de cribado de recién nacidos pueden identificar a los que están en riesgo lo antes posible y maximizar los efectos de la intervención disponible.

Enfermedades complejas

La mayoría de las enfermedades humanas, sin embargo, están controladas por una combinación de factores genéticos *y* ambientales que actúan en conjunto. El espectro del control genético sobre las enfermedades varía de estrictamente genético a estrictamente ambiental, y algunas enfermedades se ubican en el rango medio donde los genes y los factores ambientales influyen en el riesgo. Estas incluyen algunos cánceres y malformaciones congénitas, donde existe evidencia sólida y convincente de agregación familiar de riesgo (el sello distintivo del control genético), pero también factores de riesgo ambientales reconocidos (p. ej., exposición a virus en el útero, exposición a carcinógenos para cánceres, etc.). A diferencia de las enfermedades tradicionales de un solo gen, una enfermedad compleja probablemente refleja los efectos de uno o más genes (que pueden interactuar entre sí) y el medio ambiente, y con frecuencia hay un cierto grado de heterogeneidad etiológica en la que múltiples genes pueden conducir a la enfermedad.

Es evidente que no todas las personas expuestas a un factor de riesgo ambiental desarrollan la enfermedad obligatoriamente. Aunque el riesgo relativo de desarrollo de una enfermedad sea muy alto tras la exposición a un factor específico, el concepto de riesgo atribuible transmite el mensaje de que no toda la incidencia de una enfermedad se debe solo a la exposición específica en cuestión. Por ejemplo, se ha documentado con claridad la relación entre el tabaquismo y el cáncer de pulmón. Sin embargo, no todos los fumadores desarrollan un cáncer de pulmón, mientras que algunos no fumadores sí lo desarrollan. Puede que se necesite otro factor ambiental además del tabaquismo o bien los individuos difieren en su susceptibilidad genética, o ambos contribuyen al riesgo.

Lancelot Hogben describió sucintamente la interacción entre factores genéticos y ambientales hace muchos años:

Nuestros genes no pueden fabricar ladrillos sin paja. Las diferencias individuales que muestran los varones y las mujeres se deben en parte al hecho de que reciben genes diferentes de sus progenitores y en parte al hecho de que los mismos genes habitan en casas diferentes[6].

Cómo los diseños de estudios epidemiológicos pueden aclarar los papeles de los factores genéticos y ambientales en el riesgo de enfermedad

En este capítulo analizamos cómo algunos diseños de estudios

epidemiológicos comunes pueden extenderse para identificar cuándo contribuyen los factores genéticos a la causa de la enfermedad. Claramente, si los factores genéticos influyen en el riesgo de la enfermedad, los familiares de los casos (individuos con la enfermedad) deben tener un riesgo mayor que los familiares de los controles (individuos sin la enfermedad). Por lo tanto, cuando se realiza un estudio de casos y controles convencional, es posible evaluar la agregación familiar simplemente preguntando sobre el estado de la enfermedad en familiares de ambos, casos y controles. La información fiable se puede obtener de los familiares de primer grado (padres, hermanos e hijos), y la mayoría de las personas tienen información sobre los familiares de segundo grado (hermanastros, tíos, abuelos/nietos, etc.). Sin embargo, más allá de los miembros cercanos de la familia se vuelve más difícil obtener información fiable a través de una entrevista directa solo con la persona índice (es decir, el caso o el control), y es probable que dicha información esté sujeta a sesgo de selección y de información.

Dichos datos informados sobre la prevalencia de la enfermedad en parientes se pueden resumir como una variable binaria simple (es decir, antecedentes familiares positivos o negativos) o se pueden estratificar por tipo de parientes (es decir, número de parientes varones afectados). De cualquier manera, la información de los antecedentes familiares se trata como un factor de riesgo, y se puede calcular y evaluar una relación de riesgo para determinar su importancia estadística. Si se dispone de información basal válida sobre el riesgo de la enfermedad en la población por edad, sexo y cohorte de nacimiento, esta relación de riesgo puede considerarse como una proporción de incidencia estandarizada para toda la familia. Hay que tener en cuenta, sin embargo, que las familias comparten más que los genes. La exposición común a factores ambientales (p. ej., el tabaquismo pasivo, la dieta) puede explicar, al menos en parte, el nivel de agregación familiar.

Estudios con gemelos

Los estudios con gemelos han sido muy útiles para identificar la contribución relativa de los factores genéticos y ambientales a la causalidad de las enfermedades humanas. Hay dos tipos de gemelos: monocigóticos o MC (idénticos) y dicigóticos o DC (fraternales). Los gemelos MC proceden del mismo óvulo fecundado y comparten el 100% de su material genético. Sin embargo, los gemelos DC son genéticamente como otros hermanos y, por tanto, comparten un promedio del 50% de su material genético.

Si nos fijamos en la aparición de una enfermedad en gemelos MC (que tienen el mismo material genético), se pueden observar varias posibilidades. Ambos gemelos (A y B) pueden tener o no tener la enfermedad, es decir, los miembros del par de gemelos pueden ser *concordantes* para la enfermedad. También es posible que el gemelo A tenga la enfermedad y el gemelo B no (o viceversa), es decir, el par de gemelos es *discordante* para la enfermedad.

¿Qué información aporta sobre el papel de los factores genéticos el hecho

de que los gemelos MC sean concordantes para una enfermedad? ¿Podría ser la enfermedad genética? Sí, ya que los gemelos tienen el mismo material genético. ¿Podría ser ambiental? Sí, ya que se sabe que los progenitores a menudo crían a los gemelos MC de forma similar, y obviamente compartieron las mismas exposiciones en el útero, por lo que están expuestos a muchos de los mismos factores ambientales. Sin embargo, incluso si la exposición compartida a los factores ambientales fuera igual para los gemelos MC y DC, el tipo de placentación puede influir en la concordancia de los gemelos. Las placentas de los gemelos MC pueden ser monocoriónicas o dicoriónicas, mientras que en los gemelos DC siempre son dicoriónicas. Por lo tanto, la exposición en el útero a un factor ambiental en el mismo grado en gemelos MC y DC puede resultar en una mayor concordancia en el resultado para los primeros que para los segundos, ya que las anastomosis de vasos sanguíneos placentarios en las placentas monocoriónicas que caracterizan a muchos gemelos MC pueden llevar a una mayor participación de la exposición ambiental que en las placentas dicoriónicas. Es interesante, por ejemplo, que en los gemelos MC, no así en los DC, cuyas madres habían estado igualmente expuestas a la talidomida durante el embarazo, la concordancia de defectos congénitos fue casi universal[7]. Por lo tanto, observar concordancia en gemelos MC no prueba que una enfermedad sea genética o ambiental en origen, pero observar discordancia en gemelos MC indica algún papel para los factores de riesgo ambientales. Debido a que los gemelos MC discordantes comparten el mismo material genético, la enfermedad tendría que ser, al menos en parte, de origen ambiental.

En los gemelos DC intervienen tanto factores ambientales como genéticos. Si una enfermedad es genética, sería de esperar que hubiese menos concordancia en los gemelos DC que en los MC, porque comparten menos alelos en promedio. Una suposición clave de los estudios de gemelos es que los gemelos MC y DC comparten factores ambientales en el mismo grado, lo que puede no cumplirse si los gemelos MC son tratados de manera más similar que los gemelos DC por sus progenitores y otros miembros de la familia. Otras suposiciones clave son la ausencia de emparejamiento selectivo y endogamia entre los progenitores de los gemelos[8].

¿Cómo se calculan las tasas de concordancia y de discordancia en gemelos? En la figura 16.1 se muestra una tabulación cruzada de los gemelos 1 y 2. Los números de cada celda son los números de *pares de gemelos*, de modo que hay a pares (en los que ambos gemelos 1 y 2 tienen la enfermedad), d pares (en los que ni el gemelo 1 ni el 2 tienen la enfermedad), b pares (en los que el gemelo 1 no tiene la enfermedad, pero el 2 sí) y c pares (en los que el gemelo 1 tiene la enfermedad, pero el 2 no).

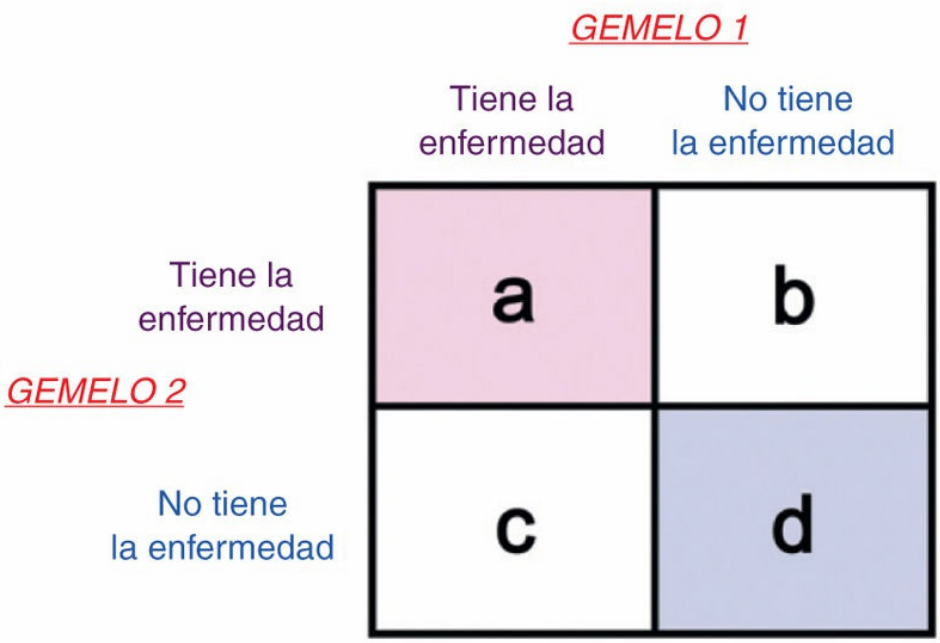


FIG. 16.1 Concordancia en gemelos para una variable dicotómica, como la leucemia.

Si se quiere calcular la tasa de concordancia en los gemelos, la mayoría de ellos se incluirán en la categoría *d*, es decir, ninguno tendrá la enfermedad. Se suelen evaluar, por tanto, las otras tres celdas, correspondientes a los pares de gemelos en los que al menos uno de ellos tiene la enfermedad. Se puede calcular la tasa de concordancia en los pares de gemelos en los que al menos uno de ellos tiene la enfermedad de la siguiente manera:

$$\text{Tasa de concordancia} = \frac{a}{a+b+c}$$

También se puede calcular la tasa de discordancia en todos los pares de gemelos en los que al menos uno tiene la enfermedad del siguiente modo:

$$\text{Tasa de discordancia} = \frac{b+c}{a+b+c}$$

En la tabla 16.1 se muestran los datos de concordancia para la leucemia en pares de gemelos monocigóticos y dicigóticos. Se observa que el porcentaje de pares concordantes es notablemente alto para la leucemia congénita, lo que sugiere con fuerza un componente genético significativo en la causalidad cuando la enfermedad se produce cerca del momento del nacimiento. Sin embargo, el número de gemelos con leucemia perinatal/congénita es pequeño (especialmente para los gemelos DC) y habría confusión entre las causas

verdaderamente genéticas y los factores intrauterinos, que son compartidos por todos los gemelos.

Tabla 16.1

Distribución por edades en los artículos clínicos publicados sobre la leucemia en gemelos, 1928-1974

	PARES MONOCIGÓTICOS		PARES DICIGÓTICOS	
	Concordante	Discordante	Concordante	Discordante
Perinatal-congénita	14	1	1	1
2-7 años	6	13	3	5
7-12 años	1	8	–	1
≥12 años	5	14	0	3
Total	**26**	**36**	**4**	**10**

De Keith L, Brown ER, Ames B, et al. Leukemia in twins: Antenatal and postnatal factors. *Acta Genet Med Gemellol.* 1976;25:336–341.

A continuación, se presentan varios ejemplos de cómo se utilizan los datos de concordancia. En la tabla 16.2 se recogen las tasas de concordancia para el alcoholismo en gemelos monocigóticos y dicigóticos descritas en varios estudios[9-12]. En casi todos los estudios publicados se han indicado mayores tasas de concordancia en gemelos monocigóticos que en gemelos dicigóticos; únicamente los resultados de un estudio realizado con un número relativamente pequeño de gemelos no concordaron con los resultados de los otros estudios. Por tanto, estos datos publicados en la literatura sugieren con fuerza un componente genético en la etiología del alcoholismo.

Tabla 16.2

Concordancia para el alcoholismo en pares de gemelos monocigóticos (MC) y dicigóticos (DC) identificados por un miembro alcohólico

	CONCORDANCIA			
Autor (año)	Número de pares de gemelos	MC (%)	DC (%)	Cociente de concordancia MC:DC
Kaij (1960)	174	71	32	2,2
Hrubec y cols. (1981)	15.924	26	13	2,0
Murray y cols. (1983)	56	21	25	0,8
Pickens y cols. (1991)	86 (V)	59	36	1,6
	44 (M)	25	5	5,0

Modificada de Lumeng L, Crabb DW. Genetic aspects and risk factors in alcoholism and alcoholic liver disease. *Gastroenterology*. 1994;107:572–578.

Se debe señalar que la cigosidad a menudo se basa en diferente información de estudios de gemelos, y al examinar datos como los que se muestran en las tablas 16.1 y 16.2 hay que preguntarse cómo se clasificaron los gemelos como MC o DC. (Debe recordarse la advertencia previa: al observar las diferencias entre grupos o cambios a lo largo del tiempo, la primera pregunta que debe plantearse es si son reales. Únicamente cuando se está convencido de que una diferencia o cambio es real y no espurio se debería proceder a interpretar los hallazgos.) La mejor manera de clasificar la cigosidad es mediante la comparación de marcadores genéticos entre gemelos (que deben mostrar consistentemente una identidad del 100% para los gemelos MC y del 50% para los gemelos DC); sin embargo, los estudios de ADN pueden no estar siempre disponibles. Los datos de los cuestionarios realizados a los gemelos o a sus progenitores son generalmente lo suficientemente exactos si se incluyen preguntas sobre la frecuencia con la que los progenitores, maestros u otros miembros de la familia confundieron a un gemelo con el otro y cuánto era el parecido físico entre los gemelos en su infancia. Esta clasificación de la cigosidad es casi tan precisa como las pruebas que usan marcadores genéticos.

Hasta ahora se ha comentado la concordancia para una variable discreta, como la leucemia o la esquizofrenia, que puede estar presente o ausente. Sin embargo, a menudo nos interesa determinar la concordancia para un fenotipo continuo, como la presión arterial. En este caso, se podrían representar en una gráfica los datos del gemelo 1 frente a los del gemelo 2 para todos los pares de gemelos y calcular el coeficiente de correlación intraclase *(r)*, que mide la correlación de valores idénticos, como se observa en la figura 16.2. El coeficiente de correlación varía de −1 a +1. Un coeficiente de correlación de +1 indica una correlación totalmente positiva, 0 indica que no hay correlación y −1 indica una correlación totalmente inversa. Si se representan en una gráfica los datos de las parejas de gemelos MC y DC, como se muestra en la figura 16.3, sería de esperar encontrar una correlación más sólida para los gemelos MC en comparación con los gemelos DC si el fenotipo cuantitativo estuviese bajo control genético. El uso de fenotipos cuantitativos dentro de un diseño gemelar clásico ofrece la oportunidad de estimar un componente genético puramente aditivo (debido a los alelos transmitidos en genes no observados), un componente residual (que refleja factores aleatorios no compartidos entre gemelos) y un componente ambiental compartido o un componente genético (que representa la interacción entre alelos en un solo gen)[8]. Es importante darse cuenta de que estos dos últimos componentes están confundidos entre sí en el diseño clásico de gemelos, donde solo se dispone de datos sobre los gemelos. La extensión del estudio para incluir a otros miembros de la familia de los gemelos crea más contrastes, lo que permite un modelado adicional utilizando herramientas estadísticas más

elaboradas.

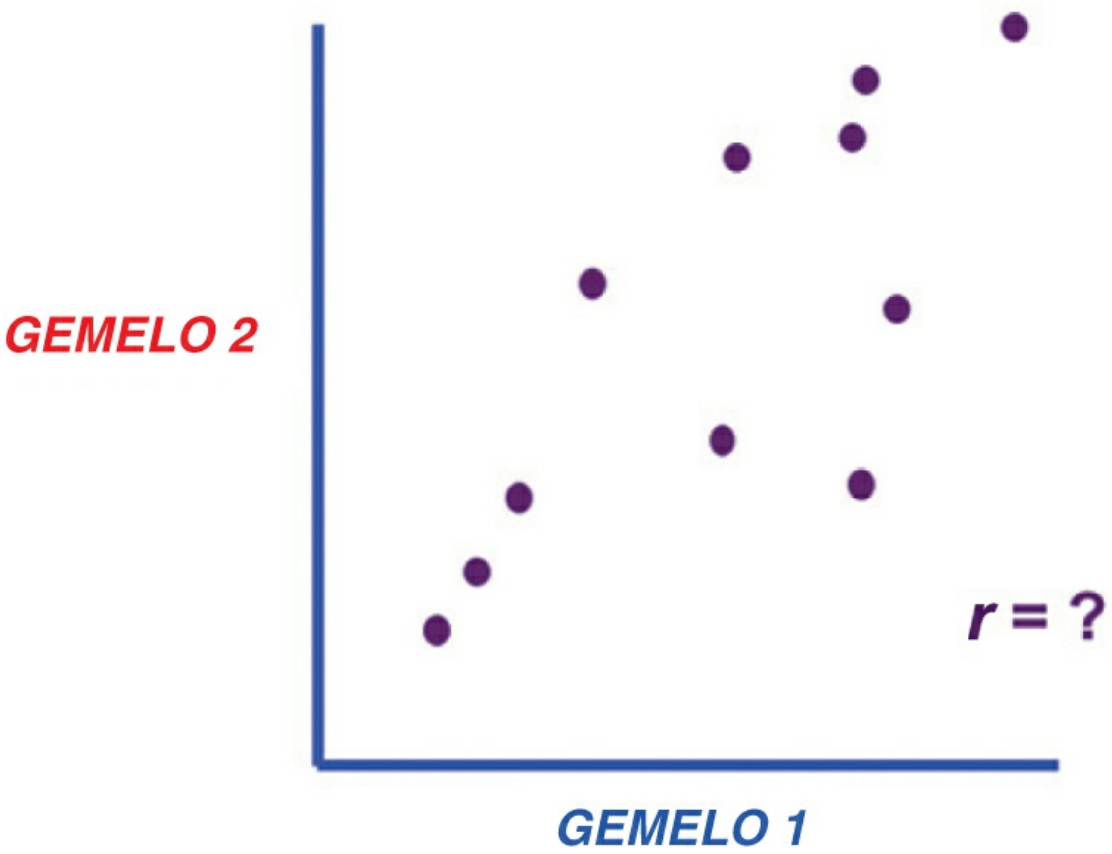


FIG. 16.2 Diagrama de dispersión que ilustra la correlación en gemelos para una variable continua, como la presión arterial sistólica. Esto se resume típicamente por el coeficiente de correlación r^2.

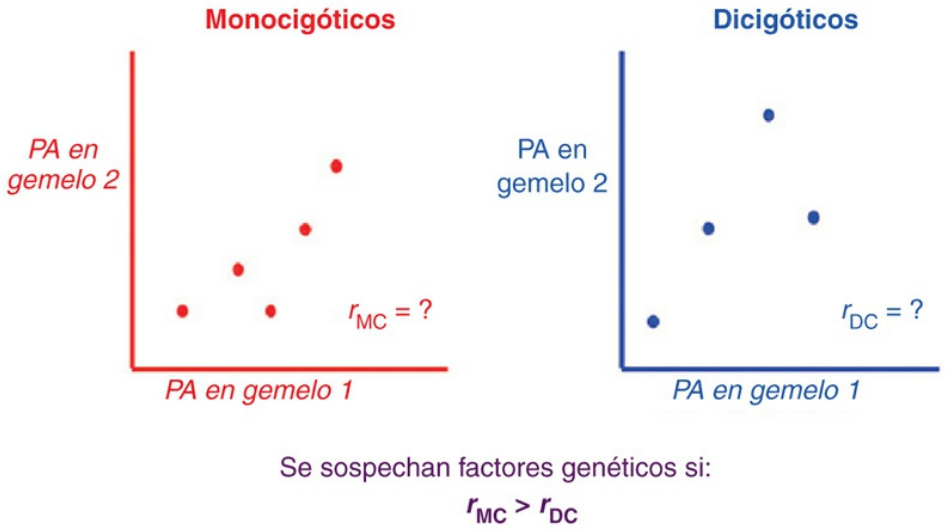


FIG. 16.3 Diagramas de dispersión que ilustran la correlación para variables

continuas, como la presión arterial *(PA)*, en gemelos monocigóticos y dicigóticos para explorar el papel etiológico de los factores genéticos en el control de una variable continua. Los gemelos monocigóticos deben tener una mayor correlación que los gemelos dicigóticos.

En la tabla 16.3 se muestran los coeficientes de correlación para la presión arterial sistólica entre familiares. El coeficiente más alto se observa en los gemelos MC; los valores para los gemelos DC y los hermanos convencionales son parecidos. También resulta interesante comprobar que prácticamente no existe ninguna correlación entre cónyuges. Una correlación fuerte entre cónyuges (que por lo general carecen de relación biológica) sugiere la implicación de factores ambientales. Sin embargo, podría sugerirse como alternativa que para el matrimonio se busca a personas similares a uno mismo («emparejamiento selectivo»). Así, los individuos con sobrepeso, por ejemplo, pueden buscar a otros individuos con sobrepeso para casarse. En tal situación se obtendría una correlación conyugal alta para el peso y el índice de masa corporal, y esto sucedería incluso para afecciones que no están determinadas por factores ambientales.

Tabla 16.3

Correlación de la presión arterial sistólica entre familiares

Familiares comparados	Coeficientes de correlación
Gemelos monocigóticos	0,55
Gemelos dicigóticos	0,25
Hermanos	0,18
Progenitores e hijos	0,34
Cónyuges	0,07

Modificada de Feinleib M, Garrison MS, Borhani N, et al. Studies of hypertension in twins. En Paul O (ed.): *Epidemiology and Control of Hypertension*. Nueva York: Grune & Stratton; 1975:3-20.

Lichtenstein y cols. publicaron un estudio a gran escala con gemelos en 2000[13] cuya finalidad fue estimar la contribución relativa de los factores ambientales y genéticos en la etiología del cáncer. Se utilizaron datos de 44.788 pares de gemelos inscritos en los registros de gemelos sueco, danés y finlandés para evaluar los riesgos de cáncer en 28 localizaciones anatómicas en los gemelos. Los gemelos de las personas con cáncer de estómago, colorrectal, de pulmón, de mama y de próstata tenían un riesgo mayor de desarrollar el mismo tipo de cáncer. Los resultados de este y de otros estudios sobre el cáncer ponen de relieve la necesidad de considerar los efectos de los factores tanto genéticos como ambientales (y sus interacciones) a la hora de investigar la etiología de los diferentes tipos de cáncer.

Estudios de adopción

Ya se ha comentado que uno de los problemas a la hora de interpretar los hallazgos de los estudios con gemelos es que incluso los gemelos MC que comparten la misma dotación genética también comparten gran parte del mismo entorno. Por tanto, en estos estudios es difícil averiguar cuál es la contribución relativa de los factores genéticos y ambientales a la causa de la enfermedad. Una estrategia común es comparar diferentes grupos de niños adoptados, sus padres adoptivos (o miembros de la familia) y sus padres biológicos. Por ejemplo, si estamos interesados en saber si el origen de la esquizofrenia es principalmente genético o ambiental, podemos realizar un estudio con niños adoptados. Hay tres contrastes básicos:

1. Hijos de padres biológicos normales criados por padres adoptivos esquizofrénicos.
2. Hijos de padres biológicos normales criados por padres adoptivos normales.
3. Hijos de padres biológicos esquizofrénicos criados por padres adoptivos normales.

Podemos examinar la descendencia de padres biológicos normales que son adoptados y criados por padres esquizofrénicos. Si la enfermedad es de origen puramente genético, ¿qué riesgo de esquizofrenia esperaríamos en estos niños adoptados? Debería aproximarse al observado en el resto de la población porque el entorno familiar no tendría ningún efecto en el aumento del riesgo. Si la enfermedad es en gran medida ambiental, esperaríamos que la crianza en un entorno con padres adoptivos esquizofrénicos aumentara el riesgo de esquizofrenia en estos niños adoptados. Como grupo de control, también podríamos examinar la descendencia de padres biológicos normales criados por padres adoptivos normales, y esperaríamos que tuvieran la misma tasa de esquizofrenia que la observada en la población general. Obviamente, para las enfermedades con baja prevalencia, será difícil encontrar padres adoptivos afectados, por lo que la tercera opción se utiliza con mayor frecuencia cuando se evalúa el mayor riesgo de enfermedad en niños de padres biológicos esquizofrénicos que fueron adoptados por padres normales.

Al interpretar los datos de los estudios de adopción, deben tenerse en cuenta ciertos factores. El primero es la edad a la que se realizó la adopción. Por ejemplo, si la adopción se produjo al final de la infancia, puede que parte del ambiente del niño haya sido el de los progenitores biológicos. Lo ideal sería estudiar a niños que son adoptados al nacer. Otra fuente de complicaciones es que, después de la adopción, algunos niños mantienen relaciones con sus progenitores biológicos, como visitas y otras exposiciones al ambiente de estos, por lo que la separación entre el ambiente de los progenitores biológicos y el de los progenitores adoptivos no es completa.

Se han realizado muchos estudios de adopción detallados en los países escandinavos, que tienen excelentes registros de enfermedades, registros de

adopción y sistemas de vinculación de registros. A título de ejemplo, en la tabla 16.4 se muestran los datos de un estudio sobre la esquizofrenia realizado por Kety e Ingraham en el que se estudiaron las tasas de esquizofrenia en familiares biológicos y adoptivos de niños adoptados[14]. Utilizando datos de registros de adopción y registros psiquiátricos hospitalarios nacionales, los autores identificaron a 34 personas adoptadas que más tarde se convirtieron en esquizofrénicos y también a 34 niños adoptados sin enfermedad mental grave. Después, analizaron las tasas de esquizofrenia en los familiares biológicos y adoptivos de los niños adoptados esquizofrénicos y en los adoptados del grupo control. La tasa de esquizofrenia en los familiares biológicos de los adoptados esquizofrénicos fue del 5,0%, frente al 0,4% en los familiares biológicos de niños adoptados sin enfermedad mental grave. Los hallazgos sugieren con fuerza que hay un componente genético en la etiología de la esquizofrenia.

Tabla 16.4

Esquizofrenia en familiares biológicos y adoptivos de personas adoptadas que desarrollaron esquizofrenia (estudio nacional de adoptados de Dinamarca)

	FAMILIARES BIOLÓGICOS			FAMILIARES ADOPTIVOS		
		ESQUIZOFRÉNICOS			ESQUIZOFRÉNICOS	
	N.° total	N.°	%	N.° total	N.°	%
Adoptados que desarrollaron esquizofrenia ($N = 34$)	275	14	5,0	111	0	0
Adoptados del grupo control (sin enfermedad mental grave) ($N = 34$)	253	1	0,4	124	0	0

De Kety SS, Ingraham LJ. Genetic transmission and improved diagnosis of schizophrenia from pedigrees of adoptees. *J Psychiatr Res.* 1992;26:247–255.

En la tabla 16.5 se muestran los coeficientes de correlación para la agregación paternofilial de la presión arterial, comparando los niños biológicos con los niños adoptados. Se observa con claridad que las correlaciones son mucho más débiles (cercanas a 0) entre progenitores y niños adoptados que las correlaciones entre progenitores e hijos biológicos. Los resultados sugieren con fuerza la existencia de un componente genético en la presión arterial, un rasgo cuantitativo.

Tabla 16.5

Coeficientes de correlación para la agregación progenitores-hijos de la presión arterial

	ENTRE PROGENITORES E	
	Hijo biológico	**Hijo adoptado**
Sistólica	0,32 (*P* <0,001)	0,09 (NS)
Diastólica	0,37 (*P* <0,001)	0,10 (NS)

NS, no significativo.

Modificada de Biron P, Mongeau JG, Bertrand D. Familial aggregation of blood pressure in 558 adopted children. *Can Med Assoc J.* 1976;115:773–774.

Tendencias temporales en la incidencia de la enfermedad

Si se observan tendencias temporales en el riesgo de la enfermedad, o bien un aumento o bien una reducción en un periodo de tiempo relativamente corto, y si existe la convicción de que la tendencia es real, la observación implica la existencia de factores ambientales en la etiología de la enfermedad. Es evidente que las características genéticas de las poblaciones humanas no suelen cambiar en periodos relativamente cortos. Por tanto, la variación de la mortalidad por arteriopatía coronaria (AC) en los varones de 1979 a 2004 que se observa en la figura 16.4 puede deberse sobre todo a los cambios de exposición a factores ambientales. (La mortalidad también puede haber disminuido debido a las mejoras en la atención médica de los pacientes con AC.)

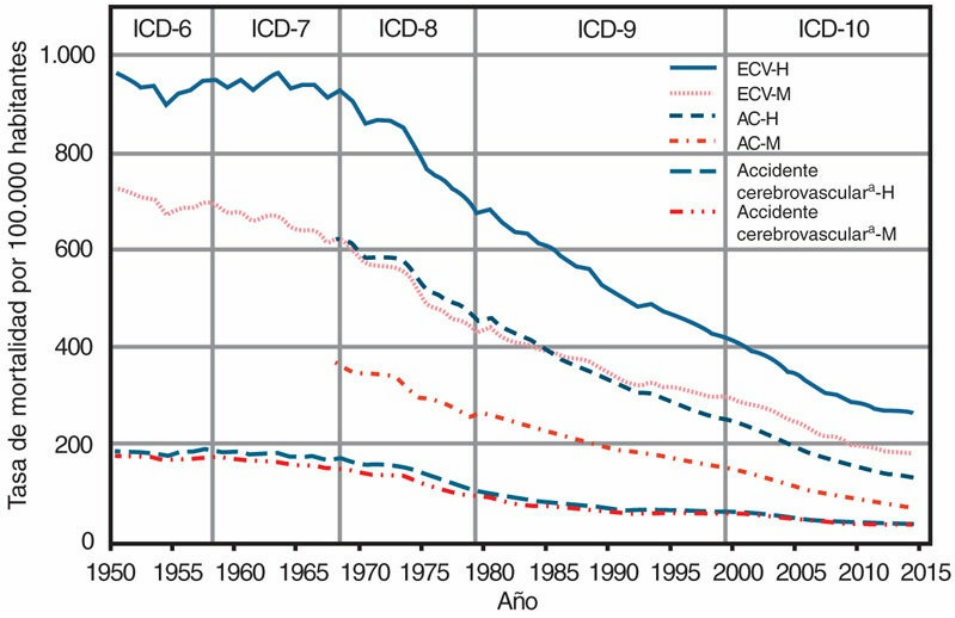


FIG. 16.4 Tasas de mortalidad ajustadas por edad y por sexo para la enfermedad cardiovascular *(ECV),* 1950-2014. *ECV,* enfermedad cardiovascular; *H,* hombre*; ICD,* International Classification of Diseases; *M,* mujer. (De Mensah GA, Wei GS, Sorlie PD, et al. Decline in cardiovascular mortality: possible causes and implications. *Circ Res.* 2017;120:366–380.)

La figura 16.5 muestra las tasas de mortalidad ajustadas por edad para el cáncer de estómago en varones de varios países. La tasa más alta se observa en Japón; la de Estados Unidos es bastante baja. ¿Son estas diferencias reales? ¿Podrían deberse a diferencias en la calidad de la atención médica o en el

acceso a la asistencia médica en los diferentes países, o bien a las diferencias internacionales en el modo de cumplimentación de los certificados de defunción? Los resultados de otros estudios sugieren que estas diferencias son reales y probablemente debidas a las diferencias en la prevalencia de un factor ambiental que se sabe que es una probable causa necesaria de cáncer de estómago, la infección por *Helicobacter pylori.* (Además, como la mortalidad por cáncer de estómago es alta, estas diferencias entre países probablemente reflejen diferencias en la incidencia de este cáncer.)

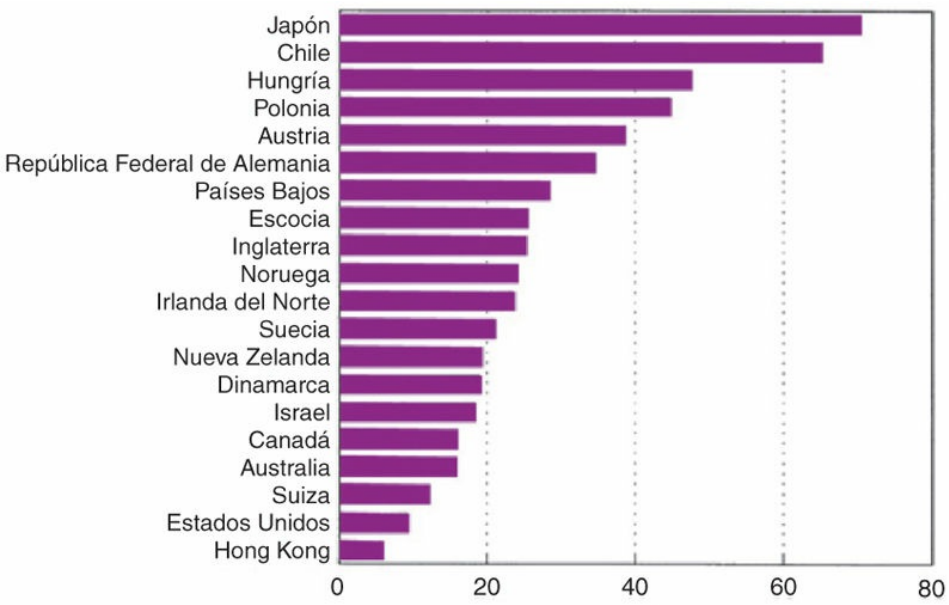


FIG. 16.5 Tasas de mortalidad por 100.000 habitantes ajustadas por edad para el cáncer de estómago en 20 países, varones, 1976-1977. (Datos de Page HS, Asire AJ: *Cancer Rates and Risks.* 3.ª ed. Washington, DC: US Government Printing Office; 1985, NIH Publication N.° 85–691.)

La figura 16.6 muestra datos comparables para el cáncer de mama en las mujeres. En ella se observa que una de las tasas más bajas del mundo está en Japón. ¿Se deben las diferencias entre los países a factores ambientales o genéticos? Nuevamente, las influencias ambientales son probablemente más importantes en la etiología del cáncer de mama, aunque se sabe que algunos genes causan esta enfermedad (p. ej., *BRCA1* y *BRCA2*), pero estos genes reconocidos no pueden explicar la mayoría de las variaciones en el riesgo[15,16].

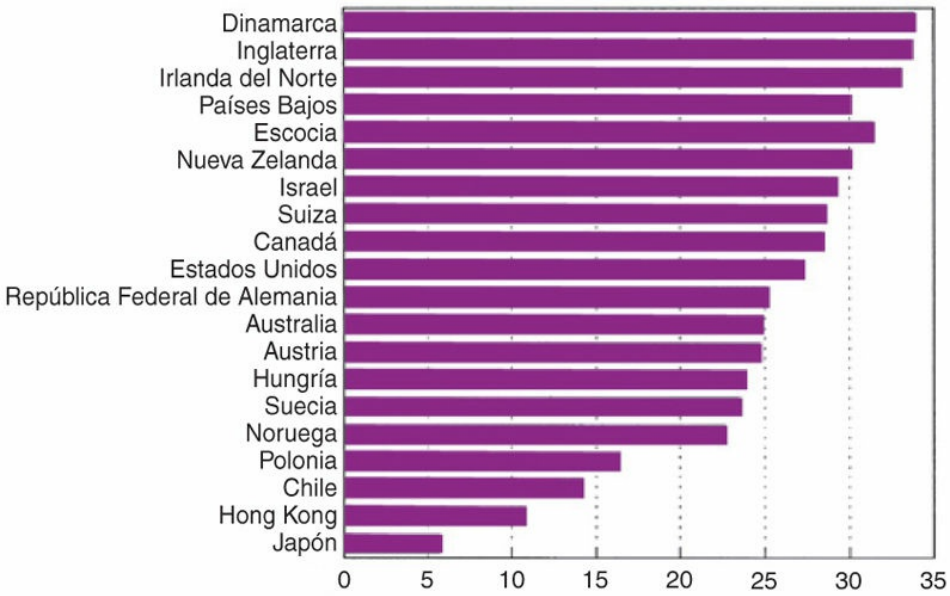


FIG. 16.6 Tasas de mortalidad por 100.000 habitantes ajustadas por edad para el cáncer de mama en 20 países, mujeres, 1976-1977. (Datos de Page HS, Asire AJ: *Cancer Rates and Risks.* 3.ª ed. Washington, DC: US Government Printing Office; 1985, NIH Publication N.° 85–691.)

¿Cómo se pueden desglosar las contribuciones relativas de los factores genéticos y ambientales a las diferencias internacionales en cuanto al riesgo de enfermedad? Es posible hacerlo mediante el estudio en inmigrantes de forma análoga a la que acaba de describirse para el estudio de adoptados.

Estudios en inmigrantes

Supongamos que una persona japonesa que vive en Japón, un país con un alto riesgo de cáncer de estómago, se traslada a Estados Unidos, un país con un bajo riesgo de dicho cáncer. ¿Cómo sería de esperar que se modificase su riesgo de cáncer de estómago? Si la enfermedad es principalmente de origen genético, sería de esperar que el riesgo elevado de cáncer de estómago se mantuviese incluso cuando las personas se trasladan de una zona de alto riesgo a una de bajo riesgo. Sin embargo, si la enfermedad es principalmente de origen ambiental, podría esperarse que, con el tiempo, el riesgo para este grupo de inmigrantes se modificase hasta el menor riesgo del país de adopción.

La tabla 16.6 muestra las tasas de mortalidad estandarizada (TME) para el cáncer de estómago en los varones japoneses que viven en Japón, los varones japoneses que emigraron a EE. UU. («Issei») y los hijos de los inmigrantes japoneses («Nisei») nacidos en EE. UU., en comparación con las TME de los varones de raza blanca estadounidenses. Se observa que las TME se desplazan progresivamente hacia las TME más bajas de los varones blancos de EE. UU. Estos datos sugieren con fuerza la implicación de un componente ambiental significativo. Por ejemplo, como se mencionó anteriormente, es bien sabido que la prevalencia de una causa importante de cáncer de estómago, la infección por *Helicobacter pylori,* varía entre los países[17].

Tabla 16.6

Tasas de mortalidad estandarizada para el cáncer de estómago en varones japoneses, Issei, Nisei y varones estadounidenses de raza blanca

Grupo	Tasa de mortalidad estandarizada
Varones japoneses	100
Issei[a]	72
Nisei[a]	38
Varones estadounidenses de raza blanca	17

[a] Issei y Nisei son inmigrantes japoneses de primera y segunda generación, respectivamente.

De Haenszel W, Kurihara M: Studies of Japanese migrants: I. Mortality from cancer and other disease among Japanese in the United States. *J Natl Cancer Inst.* 1968;40:43–68.

Sin embargo, se debe tener en cuenta que, cuando las personas emigran a otro país, ellas y sus familias no se desprenden inmediatamente del ambiente de su país de origen. Se conservan muchos aspectos de su cultura original, como ciertas preferencias alimentarias. Por tanto, el microambiente del inmigrante, en especial las características ambientales relacionadas con el estilo de vida, suele ser una combinación de las del país de origen y de las del país de adopción. Otra consideración importante es la edad a la que la persona emigra; al interpretar los resultados de los estudios de inmigrantes, se debe saber cuánto tiempo ha pasado la persona en cada país.

Veamos otro ejemplo. Se ha demostrado que el riesgo de esclerosis múltiple se relaciona con la latitud: cuanto mayor es la distancia desde el ecuador, mayor es el riesgo[18]. Esta observación es muy interesante y ha suscitado muchas investigaciones. Sin embargo, sigue habiendo dudas sobre el grado en que la relación con la latitud se debe a factores ambientales, así como sobre el modo en el que se puede determinar qué factores ambientales podrían estar implicados.

Los estudios de personas que han emigrado de zonas de alto riesgo a otras de bajo riesgo son idóneos para abordar algunos de estos aspectos. Un país que se presta muy bien para un estudio de este tipo es Israel, que, debido a su latitud, tiene un bajo riesgo para la esclerosis múltiple. Israel tuvo oleadas sucesivas de inmigración durante el siglo XX. Algunos de los inmigrantes procedían de zonas de alto riesgo, como las latitudes relativamente septentrionales de Estados Unidos, Canadá y el norte de Europa, mientras que otros provenían de latitudes de bajo riesgo más cercanas al ecuador, como zonas del norte de África y de la península arábiga.

La tabla 16.7 muestra los datos de incidencia de la esclerosis múltiple en inmigrantes europeos, africanos y asiáticos en Israel. Esta enfermedad no es común, por lo que los tamaños de las muestras son pequeños (solo 67 casos entre 668.880 inmigrantes). Al observar las tasas en los inmigrantes africanos y asiáticos que se trasladaron de un área de bajo riesgo a otra, se observa que

su riesgo se mantuvo bajo. Después, si se analizan los datos de los inmigrantes europeos que se trasladaron de una zona de alto riesgo (Europa) a otra de bajo riesgo (Israel), se observa que los europeos que emigraron antes de los 15 años (fila superior) tenían una tasa baja, similar a la de los inmigrantes africanos y asiáticos. Sin embargo, los europeos que emigraron después de los 15 años de edad tendían a mantener la tasa elevada de su país de origen. Estos hallazgos sugieren que el riesgo de esclerosis múltiple se determina en la infancia y que el factor fundamental es si la infancia se pasó en una zona de alto riesgo o de bajo riesgo. Una persona que pasó su infancia en una zona de riesgo bajo mantiene un riesgo bajo y una que la pasó en una zona de riesgo alto conserva un riesgo alto, aunque después haya emigrado a un área de riesgo bajo. Esto ha sugerido que algún evento en la infancia puede ser relevante en la etiología de la esclerosis múltiple.

Tabla 16.7

Incidencia de esclerosis múltiple (EM) por 100.000 entre inmigrantes europeos, africanos y asiáticos a Israel por edad en el momento de la inmigración

	INCIDENCIA DE EM EN INMIGRANTES	
Edad en el momento de la inmigración	**Europeos**	**Africanos y asiáticos**
<15 años	0,76	0,65
15-29 años	3,54	0,40
30-34 años	1,35	0,26

Modificada de Alter M, Leibowitz U, Speer J. Risk of multiple sclerosis related to age at immigration to Israel. *Arch Neurol.* 1966;15:234–237.

Los estudios con inmigrantes no están desprovistos de problemas. En primer lugar, los inmigrantes no son representativos de las poblaciones de sus países de origen. Por tanto, hay que determinar qué selección de factores llevaron a ciertas personas a emigrar (factores de selección). Por ejemplo, las personas que están gravemente enfermas o con discapacidad no suelen emigrar. Otros factores, como las características socioeconómicas y culturales, también se relacionan con qué personas tienen propensión a emigrar y cuáles no. Por tanto, dado este problema de la selección, hay que preguntarse si es legítimo comparar las tasas de cáncer de estómago en inmigrantes japoneses de primera y segunda generación con las tasas de japoneses nativos. En segundo lugar, hay que determinar cuál era la edad en el momento de la emigración. ¿Cuántos años habían vivido los inmigrantes en su país de origen y cuántos en su país de adopción? En tercer lugar, hay que recordar que los inmigrantes no se desprenden por completo del ambiente de su país de origen después de emigrar.

Estos y otros factores deben tenerse en cuenta al interpretar los resultados de los estudios en inmigrantes. Existe un paralelismo obvio con los estudios de adopción y, como se ve en la tabla 16.8, muchos de los problemas que surgen en la interpretación de los resultados son similares para los dos tipos

de estudios.

Tabla 16.8

Problemas en la interpretación de los resultados de los estudios de adopción y en inmigrantes

Estudios de adopción	Estudios en inmigrantes
• Los adoptados están muy seleccionados • Edad variable en el momento de la adopción • Los adoptados pueden conservar varios grados de contacto con su o sus progenitores biológicos	• Los inmigrantes están muy seleccionados • Edad variable en el momento de la emigración • Los inmigrantes pueden conservar muchos elementos de su ambiente original, sobre todo los relacionados con la cultura y el estilo de vida

Cómo se usan los marcadores genéticos para mapear los genes que controlan el riesgo de enfermedades, incluidas las enfermedades complejas

Los marcadores genéticos son variantes en la secuencia de ADN que pueden tipificarse directamente. Los marcadores se transmiten de los progenitores a la descendencia claramente, siguiendo patrones mendelianos regulares, y las ubicaciones cromosómicas de los marcadores genéticos son generalmente conocidas. Si el marcador es polimórfico en la población, entonces muchas personas portarán uno u otro alelo en el marcador, y habrá un número razonable de heterocigotos y homocigotos para los marcadores bialélicos. Mientras que los grupos sanguíneos (como el tipo de sangre ABO) pueden considerarse como marcadores genéticos, la forma más común de marcadores genéticos son los polimorfismos de nucleótido único (SNP, *single nucleotide polymorphisms*). La mayoría de los SNP no tienen ningún efecto fisiológico directo, pero una minoría se localiza en regiones de codificación o reguladoras de un gen y puede afectar directamente a la susceptibilidad de una persona a la enfermedad o a la respuesta a los tratamientos médicos.

El propósito final de los estudios de mapeo es identificar genes asociados con la susceptibilidad a una enfermedad para mejorar nuestra comprensión de la patogenia de la enfermedad y facilitar el desarrollo de estrategias preventivas apropiadas. La búsqueda de genes de susceptibilidad a la enfermedad utiliza dos enfoques principales:

1. Uso de estudios familiares para identificar la *vinculación* o cosegregación entre un determinado marcador y un posible gen de la enfermedad[19]. La herencia conjunta de marcadores genéticos y la enfermedad representa una evidencia convincente de que tal gen existe al localizar el gen potencialmente causal en una región cromosómica específica. Sin embargo, si hay varios genes que pueden

causar la misma enfermedad, mostrarían evidencia de *heterogeneidad de vinculación,* donde algunas familias presentan evidencia de vinculación a un marcador dado, mientras que otras no.

2. Búsqueda de una *asociación* entre un alelo marcador y una enfermedad utilizando muestras de personas no relacionadas. Si bien incluso una sólida evidencia de asociación sigue siendo una evidencia algo menos convincente de causalidad directa, y tal evidencia está sujeta a posibles confusiones debido a la heterogeneidad dentro de la muestra, los métodos para probar la asociación estadística entre enfermedad y marcadores genéticos se basan en los diseños de estudios epidemiológicos convencionales y son fáciles de implementar.

Análisis de vinculación en estudios familiares

Como se mencionó anteriormente, cuando una persona sufre cierta enfermedad, es útil examinar a sus parientes de primer grado en busca de evidencia de una prevalencia de la enfermedad mayor a la esperada. El riesgo excesivo en los familiares de primer grado sugiere (aunque no prueba) la existencia de algún componente genético. También es posible examinar familias de alto riesgo, como el pedigrí que se muestra en la figura 16.7, que muestra una familia múltiple con retinoblastoma en cuatro generaciones sucesivas. El árbol genealógico no solo brinda una imagen visual de la naturaleza familiar de esta enfermedad, sino que también se puede usar para mapear el gen de la enfermedad mediante la prueba de *cosegregación* con un marcador genético (o muchos marcadores, como en una exploración de genoma completo). Claramente, esta familia es probable que refleje un gen autosómico dominante que controla el riesgo. Observe cómo este pedigrí en particular también demuestra cómo, incluso en familias de alto riesgo, la enfermedad puede saltarse generaciones y ser transmitida por individuos *(hombres en la tercera generación que tienen una descendencia femenina con la enfermedad)* que no están afectados ellos mismos (generalmente se denota como «penetrancia incompleta» del gen putativo de la enfermedad).

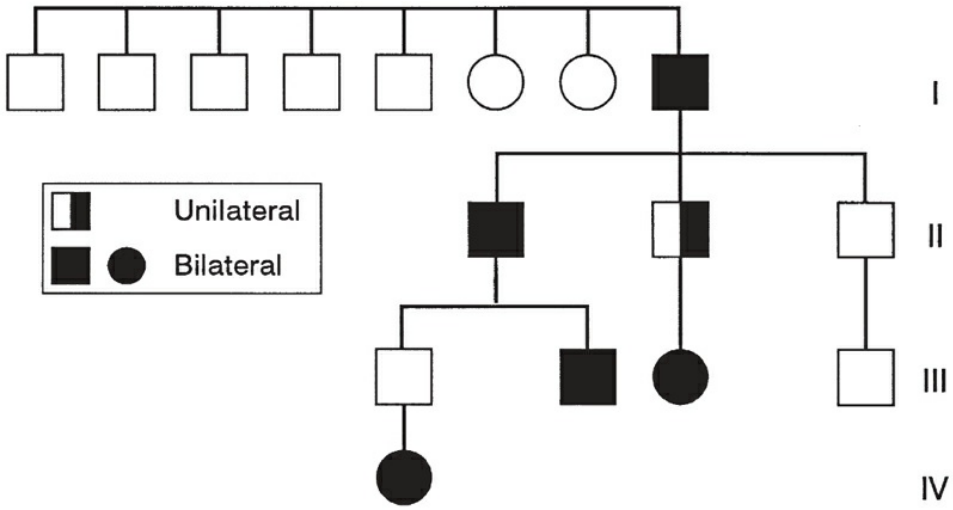


FIG. 16.7 Árbol genealógico publicado con casos de retinoblastoma en cuatro generaciones sucesivas. *Cuadrados*, varones; *círculos*, mujeres. (De Migdal C: Retinoblastoma occurring in four successive generations. *Br J Ophthalmol.* 1976;60:151–152.)

Suponiendo que el gen de la enfermedad tiene una mutación rara, autosómica dominante que causa el retinoblastoma, se puede probar la vinculación entre un marcador genético y el gen de la enfermedad hipotética.

Los *análisis de vinculación* buscan determinar si los alelos de dos loci (el locus de la enfermedad hipotética y algún marcador genético) se segregan juntos en una familia y se pasan como una unidad de padres a hijos. Los genes físicamente próximos entre sí en el mismo cromosoma (es decir, los dos genes vinculados) tenderán a transmitirse juntos; es decir, cosegregados. El vínculo genético solo puede identificarse a través de estudios familiares, y generalmente requiere familias multigeneracionales con muchos miembros afectados (denominadas familias múltiples). Se puede obtener cierta información sobre la vinculación de familias nucleares (dos padres y varios hermanos completos) si dos o más de los descendientes están afectados. Sin embargo, incluso cuando se demuestra la vinculación en algunas familias, es posible que no muestre evidencia de vinculación en otras familias si dos genes diferentes controlan el riesgo; es decir, si existe heterogeneidad de ligamiento. Además, los casos de enfermedad en las familias múltiples necesarios para el análisis de ligamiento no son totalmente representativos de todos los casos de enfermedad. El ligamiento a menudo aclara los mecanismos biológicos subyacentes a la transmisión y la patogenia de la enfermedad. Por ejemplo, se ha caracterizado el gen causal de la poliquistosis renal, una enfermedad autosómica dominante. Como se observa en la familia que se muestra en la figura 16.8, el alelo 1 muestra evidencia de vinculación con la enfermedad y se ve en el padre y dos de sus descendientes, todos los cuales fueron afectados. En el caso de la fibrosis quística (fig. 16.9), una enfermedad autosómica recesiva, la variante causal debe heredarse tanto del padre como de la madre. Por lo tanto, la enfermedad no se observa en ninguno de los padres, sino solo en el niño que posee ambos alelos.

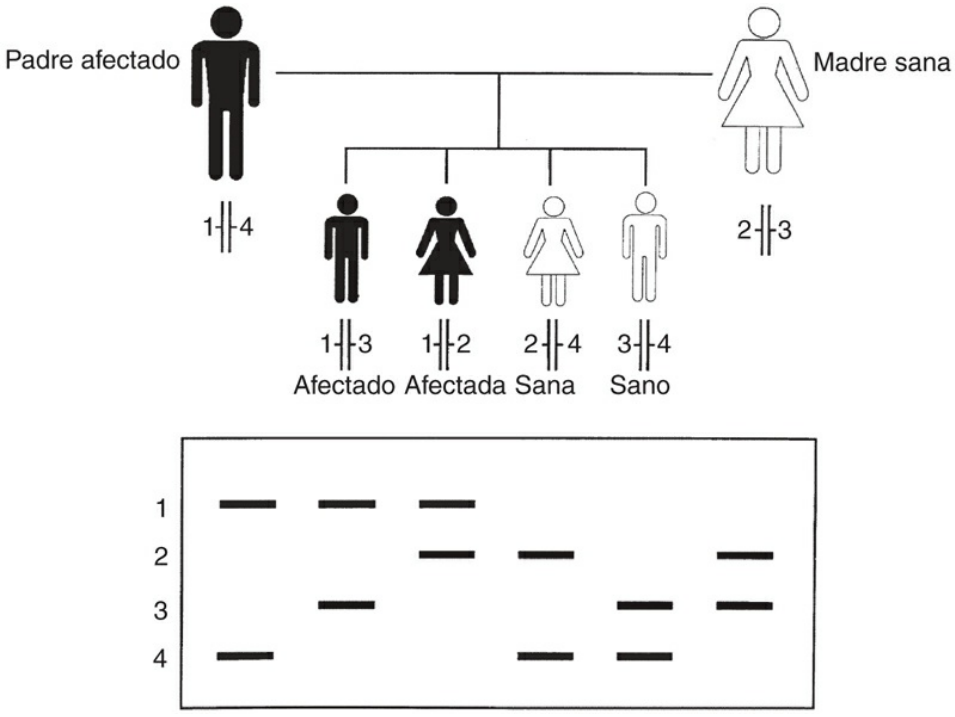


FIG. 16.8 Análisis de ADN de los trastornos autosómicos dominantes. Ejemplo:

poliquistosis renal. (De Taylor HA, Schroer RJ, Phelan MC, et al. *Counseling Aids for Geneticists.* 2.ª ed. Greenwood, SC: Greenwood Genetic Center; 1989.)

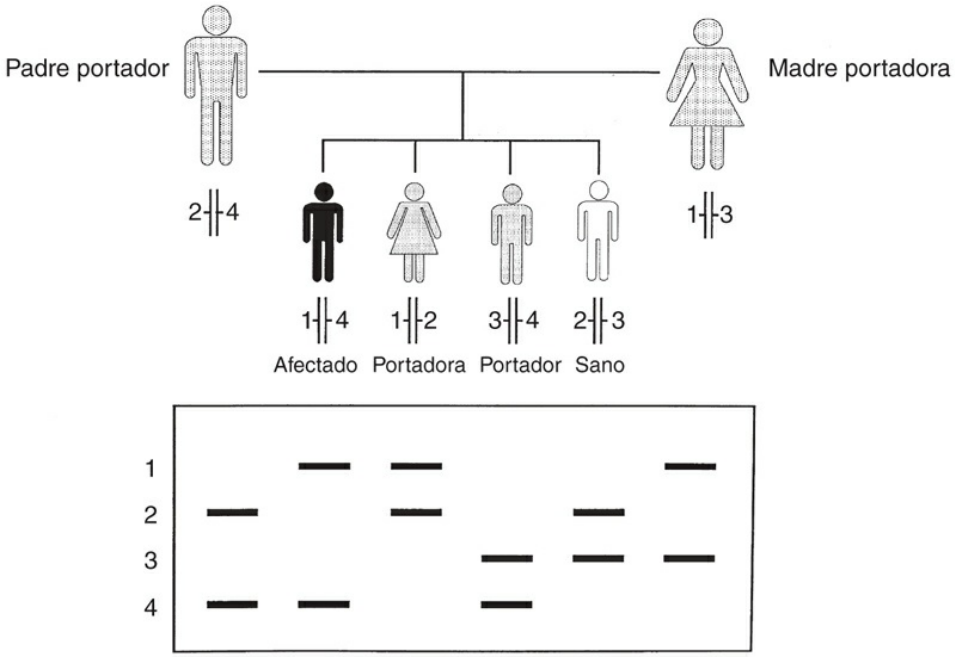


FIG. 16.9 Análisis de ADN de los trastornos autosómicos recesivos. Ejemplo: fibrosis quística. (De Taylor HA, Schroer RJ, Phelan MC, et al. *Counseling Aids for Geneticists.* 2.ª ed. Greenwood, SC: Greenwood Genetic Center; 1989.)

Estudios de asociación

Similar al enfoque de las pruebas para la agregación familiar de la enfermedad, también podemos probar una asociación entre un SNP y una enfermedad (o un fenotipo continuo). Esta prueba de asociación es la misma que la utilizada en la epidemiología tradicional, pero el alelo o genotipo marcador genético se convierte en la exposición de interés. Podemos probar la asociación con un solo gen, con un conjunto de genes que se sabe que están potencialmente asociados (genes candidatos) o de manera independiente en todo el genoma (genoma completo). La posibilidad de estudiar todo el genoma para identificar asociaciones genéticas, en forma de estudios de asociación del genoma completo (GWAS, *genome-wide association studies*), ha cambiado el enfoque general para estudiar las asociaciones entre marcadores genéticos (la mayoría de las veces, SNP) y la enfermedad. A través de los GWAS, que utilizan de 1 a 5 millones de SNP, estamos buscando regiones del genoma asociadas estrechamente con una enfermedad (y, por lo tanto, es probable que alberguen el gen causal), pero los SNP individuales que proporcionan evidencia de asociación por lo general no son directamente causales y la mayoría de las veces se localizan en regiones no codificantes del genoma. Sin embargo, estos SNP significativos pueden estar correlacionados

con algunos alelos de alto riesgo no observados en un gen causal desconocido, y aún pueden ser útiles en el mapeo de genes cercanos que controlan el riesgo. Los GWAS se han hecho posibles en los últimos 10 años aproximadamente, en gran parte debido a los avances tecnológicos en el genotipado de SNP. Se han utilizado ampliamente para diversas enfermedades comunes de las que se conoce que varios genes y algunos factores de riesgo ambientales desempeñan algún papel en la causa de la enfermedad[20,21]. La premisa subyacente en los GWAS es que los SNP comunes dispersos por todo el genoma pueden «etiquetar» los haplotipos con precisión suficiente como para revelar correlaciones entre los SNP observados y los genes causales no observados cuando se analizan bajo el diseño tradicional de estudio de casos y controles o de cohortes. La prueba estadística (generalmente una prueba de chi-cuadrado o un modelo de regresión logística, temas no cubiertos en este texto) es simple, y las diferencias significativas de frecuencia de alelos o genotipos entre grupos de casos y controles deben revelar diferencias importantes que reflejen una asociación directa o indirecta entre un marcador y algún gen causal desconocido. Debido a que la cantidad de SNP utilizada en los GWAS es grande (originalmente al menos 100.000 SNP, pero ahora típicamente de 1 a 2,5 millones de SNP por persona en un estudio), el valor crítico convencional del error alfa (valor *P*) del 5% para cada SNP ya no es apropiado porque en cada estudio se realizan muchas pruebas. Para minimizar la tasa de resultados falsos positivos se establece un nivel de significación de todo el genoma en $P = 5 * 10^{-7} – 5 * 10^{-8}$.

Muchos de los problemas en la metodología e interpretación de los resultados tratados en capítulos anteriores se aplican a las asociaciones entre enfermedades y marcadores genéticos. Esto incluye sesgos (selección, información) y factores de confusión. Un factor de confusión importante para los estudios de asociación genética es la ascendencia. Es un hecho conocido que, a pesar de las similitudes genéticas entre individuos (> 99%), existen diferentes frecuencias alélicas entre individuos que se originan en diferentes partes del mundo debido a la variación natural entre subpoblaciones humanas o, posiblemente, a presiones ambientales o selectivas presentes a través de la historia de diferentes subpoblaciones. Por lo tanto, la ascendencia puede asociarse tanto con la exposición (frecuencia alélica o genotípica) como con el resultado de la enfermedad, y así confundir los resultados del estudio. Existen métodos para abordar este importante factor de confusión, entre los que se encuentran la estratificación simple por ascendencia autoinformada o la estimación del porcentaje de ascendencia de cada persona en función de marcadores genéticos, incluido el del modelo estadístico. Seleccionar los controles apropiados siempre es importante. Cuando se nos presenta una lista de asociaciones con una enfermedad y un marcador genético (grupo sanguíneo o SNP), deberíamos preguntarnos cómo fueron las conclusiones con respecto a tales asociaciones y qué grupos de comparación se utilizaron. Por lo tanto, los aspectos metodológicos expuestos en capítulos anteriores en

el contexto de diferentes tipos de diseños de estudios epidemiológicos son altamente relevantes cuando se investiga en qué forma los factores genéticos se relacionan con estas enfermedades.

A pesar de sus limitaciones, mucho se ha logrado a través de los GWAS. Desde el primer descubrimiento basado en GWAS de la asociación del gen del factor H del complemento (CFH) y la degeneración macular relacionada con la edad[22], se han identificado más de 2.000 asociaciones estrechas entre genes y enfermedades complejas[23]. Al combinar múltiples estudios, en la actualidad resulta claro que múltiples genes influyen en el riesgo de la mayoría de las enfermedades complejas que representan importantes cargas para la salud pública[20]. Estos factores de riesgo genéticos replicados deberían ser útiles para mejorar la predicción del riesgo (a un nivel amplio), la clasificación diagnóstica y, finalmente, el desarrollo de fármacos (tanto para la eficacia potencial como para minimizar los efectos adversos). Este progreso ha generado entusiasmo por la medicina de precisión donde la información genética puede guiar la medicina clínica (ver más adelante). Recientemente, Khoury y Evans[24] abogaron por un énfasis similar en el desarrollo de la «prevención de precisión», caracterizada por estrategias que desarrollan el papel y la importancia de la genética en la salud pública.

Interacción entre los factores genéticos y los factores de riesgo ambientales

La cuestión de la susceptibilidad genética a factores ambientales y la posibilidad de interacción entre ellos también deben tenerse en cuenta. En el capítulo 14 se analizó el estudio de Grant y Dawson que describía una asociación entre la menor edad de inicio del consumo de alcohol y la prevalencia del consumo de alcohol a lo largo de la vida (fig. 16.10). Como se observa en la figura 16.11, cuando los sujetos se dividieron en aquellos con antecedentes familiares positivos de alcoholismo y aquellos con antecedentes familiares negativos, la relación global persistía, aunque la prevalencia era más elevada entre aquellos con antecedentes familiares positivos[25]. Esta observación sugiere que, aunque la relación observada entre el riesgo de consumo de alcohol a lo largo de la vida y la edad de inicio del consumo de alcohol puede reflejar las influencias ambientales, el efecto de los antecedentes familiares sugiere una interacción con factores genéticos o cierta influencia de la crianza de los hijos respecto a los antecedentes familiares de consumo de alcohol.



FIG. 16.10 Prevalencia de la dependencia del alcohol durante la vida en función de la edad de inicio del consumo. (Modificada de Grant BF, Dawson DA. Age at onset of alcohol use and its association with DSM-IV alcohol abuse and dependence: Results from the National Longitudinal Alcohol Epidemiologic

Survey. *J Substance Abuse.* 1997;9:103–110.)

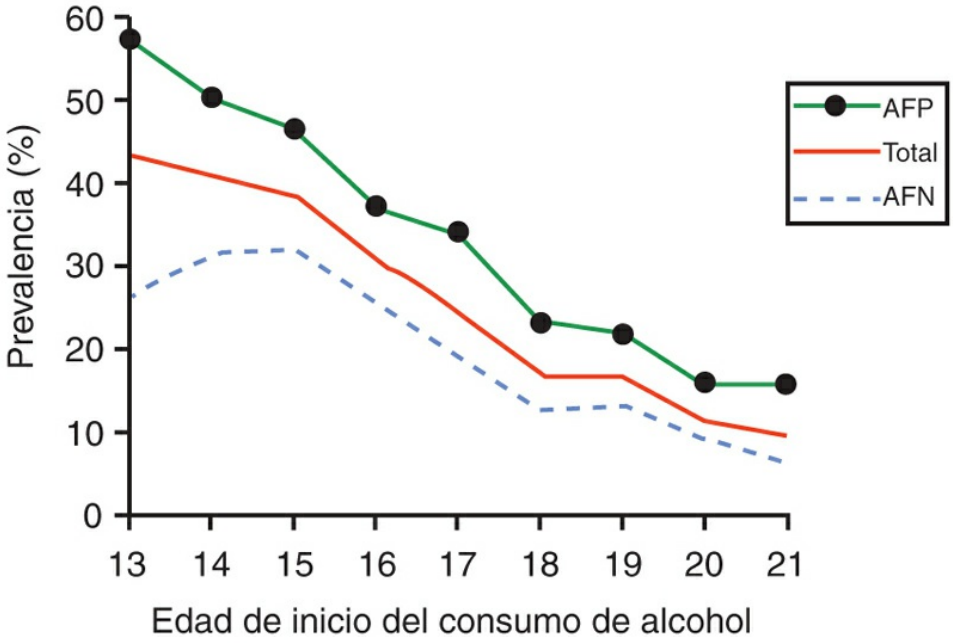


FIG. 16.11 Prevalencia de la dependencia del alcohol durante la vida en función de la edad de inicio del consumo y los antecedentes familiares de alcoholismo. *AFN,* antecedentes familiares negativos; *AFP,* antecedentes familiares positivos. (Modificada de Grant BF. The impact of a family history of alcoholism on the relationship between age at onset of alcohol use and DSM-IV alcohol dependence: Results from the National Longitudinal Alcohol Epidemiologic Survey. *Alcohol Health Res World.* 1998;22:144–147.)

La relación entre el tabaquismo y la mutación del factor V Leiden es un ejemplo de posible interacción gen-ambiente. Fumar es un factor de riesgo conocido de infarto de miocardio (IM), y el factor V de Leiden es una anomalía hereditaria común que afecta la coagulación de la sangre y aumenta el riesgo de trombosis venosa. Rosendaal y cols. realizaron un estudio de casos y controles de base poblacional de 472 mujeres jóvenes, incluidas 84 que habían sufrido un IM antes de los 45 años[26]. La mutación del factor V de Leiden fue más prevalente en el grupo de casos con IM, con un aumento de 2,4 en el cociente de posibilidades en comparación con aquellas sin esta mutación. Sin embargo, al considerar tanto el hábito de fumar como el estado de portadora, parece existir una interacción estadística entre el tabaquismo y el factor V de Leiden, como se ilustra en la tabla 16.9. En particular, las mujeres no portadoras que fumaban mostraron un cociente de posibilidades indicativo de una probabilidad 9 veces superior de sufrir un IM prematuro, mientras que las mujeres portadoras que fumaban tenían 32 veces más probabilidades de sufrir un MI prematuro en comparación con aquellas mujeres que no fumaban y no eran portadoras. Debido a que los efectos combinados estimados del genotipo y el hábito de fumar excedieron en gran medida lo que se esperaría con un modelo multiplicativo o aditivo, estos

hallazgos sugieren una interacción estadística (o una desviación de una simple combinación de los efectos marginales del genotipo y las exposiciones ambientales). Si bien este tipo de análisis es intrigante, a menudo es imposible concluir de manera definitiva si dicha interacción estadística representa verdaderas interacciones biológicas (p. ej., relaciones sinérgicas o antagónicas entre los efectos de los factores de riesgo genéticos y ambientales)[27]. Sin embargo, existe verosimilitud biológica para cada una de estas exposiciones, y su interacción también es biológicamente convincente. La detección de la interacción gen-ambiente siempre requiere muestras de mayor tamaño, y confirmar su existencia en las poblaciones puede resultar difícil porque las frecuencias alélicas y las tasas de exposición pueden diferir entre las poblaciones, según lo analizado por Aschard[28].

Tabla 16.9

Asociación entre el tabaquismo y el factor V de Leiden con el riesgo de infarto de miocardio en mujeres jóvenes

Odds ratio (IC del 95%)	GENOTIPO FACTOR V	
Fumadora actual	**Salvaje**	**Leiden**
No	1,0 (ref)	1,1 (0,1; 8,5)
Sí	9,0 (5,1; 15,7)	32,0 (7,7; 133)

El riesgo de infarto de miocardio es elevado en las mujeres jóvenes que fuman, pero es mayor en aquellas con el genotipo del factor V Leiden. Esta interacción estadística entre el genotipo y el entorno supera lo que sería esperable por la genética o el entorno solo (OR = 32).

IC, intervalo de confianza; *OR*, cociente de posibilidades.

De Austin MA, Schwartz SM. Cardiovascular disease. En: Costa LG, Eaton DL, eds. *Gene-Environment Interactions: Fundamentals of Ecogenetics*. Hoboken, NJ: John Wiley & Sons; 2006. Modificado de Rosendaal FR, Siscovick DS, Schwartz SM, et al. Factor V Leiden (resistance to activated protein C) increases the risk of myocardial infarction in young women. *Blood.* 1997; 89:2817–2821.

En 1995, Brennan y cols. publicaron un estudio sobre el tabaquismo y el cáncer epidermoide de cabeza y cuello[29]. Estos autores observaron que, en los pacientes con cáncer invasivo de cabeza y cuello, el tabaquismo se asociaba con un aumento marcado de las mutaciones del gen p^{53}, que codifica una proteína supresora de tumores. Tales mutaciones pueden contribuir tanto a la aparición como al crecimiento de los cánceres. Los investigadores estudiaron muestras de tumores de 127 pacientes con cáncer de cabeza y cuello y encontraron mutaciones de p^{53} en el 42% (54 de 127) de los pacientes. Quienes habían fumado al menos 1 paquete diario durante al menos 20 años tenían más del doble de probabilidad de tener mutaciones en p^{53} que los pacientes que no eran fumadores. Quienes fumaban y bebían más de 28 gramos de alcohol de alta graduación diarios tenían una probabilidad 3,5 veces mayor

de tener mutaciones de p^{53} que los pacientes que no fumaban ni bebían. Como se ve en la figura 16.12, se encontraron mutaciones de p^{53} en un 58% de los pacientes que fumaban y bebían, en un 33% de los pacientes que fumaban pero no bebían y en un 17% de los pacientes que no fumaban ni bebían. Además, el tipo de mutación encontrada en pacientes que no fumaban ni bebían parecía probable que fuese endógena en lugar de haber sido causada por mutágenos ambientales (es decir, exógena). Los resultados sugieren que el tabaquismo puede tender a inactivar el gen supresor tumoral p^{53} y proporcionar así una base molecular de la relación bien conocida entre el tabaquismo y el cáncer de cabeza y cuello.

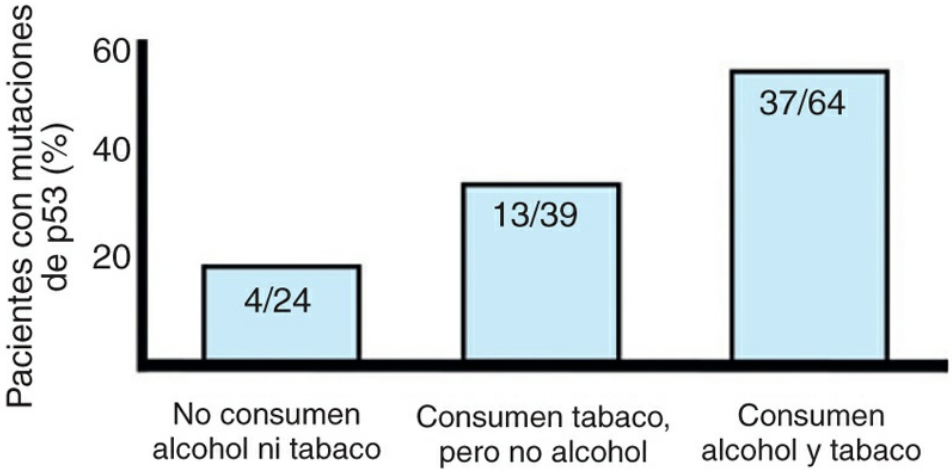


FIG. 16.12 Asociación de las mutaciones del gen p53 con el tabaquismo y el consumo de alcohol en 129 pacientes con carcinoma epidermoide de cabeza y cuello. (De Brennan JA, Boyle JO, Koch WM, et al. Association between cigarette smoking and mutation of the p53 gene in squamous cell carcinoma of the head and neck. *N Engl J Med.* 1995;332:712–717.)

Un paso más en este enfoque es identificar un defecto génico específico que se asocie con una cierta exposición ambiental. Un ejemplo se observa en los resultados que relacionan un defecto específico del gen p53 con la exposición a la aflatoxina en los pacientes con carcinoma hepatocelular (CHC). En el capítulo 14 se comentó el sinergismo positivo entre el virus de la hepatitis B (VHB) y la exposición a la aflatoxina B_1 a la hora de incrementar el riesgo de CHC. Para determinar si la frecuencia de una mutación específica del gen supresor tumoral p^{53} (una mutación en un «punto caliente» situado en el codón 249) se relacionaba con el riesgo de exposición a la aflatoxina, Bressac y cols. analizaron muestras de CHC de 14 países[30]. La mutación se observó en el 17% (12/72) de las muestras tumorales de cuatro países del sur de África y la costa del sureste de Asia, pero en ninguna de las 95 muestras de otras localizaciones geográficas, incluyendo Norteamérica, Europa, Oriente Medio y Japón. Los cuatro países en los que se encontró la mutación (China, Vietnam, Sudáfrica y Mozambique) tienen la mayor incidencia de CHC del mundo y comparten un clima cálido y húmedo similar, lo que favorece el

crecimiento de hongos que producen aflatoxina. La tasa de portador del VHB era elevada, pero no varió significativamente entre los países estudiados. Sin embargo, el riesgo de exposición a la aflatoxina sí era distinto entre estos países y se observó que la presencia de la mutación se correlacionaba con el riesgo de exposición a las aflatoxinas.

Aguilar y cols. brindaron un apoyo adicional a estos hallazgos al estudiar muestras de hígado sano de tres áreas geográficas con diferencias en cuanto a la prevalencia de exposición a la aflatoxina: niveles insignificantes (Estados Unidos), bajos (Tailandia) y altos (Qidong, China)[31]. La frecuencia de la mutación era paralela al nivel de exposición a la aflatoxina B_1, lo que sugiere que la aflatoxina tiene un papel causal y probablemente precoz en el desarrollo de tumores hepáticos.

Así pues, los estudios que combinan métodos epidemiológicos y moleculares pueden tener una utilidad incalculable para confirmar el papel etiológico de ciertos agentes ambientales mediante la demostración de sus efectos génicos específicos. Por otra parte, estos estudios también pueden sugerir vías y mecanismos biológicos que pueden estar implicados en el desarrollo de ciertos tipos de cáncer y de otras enfermedades. No obstante, los estudios epidemiológicos y moleculares combinados también pueden ayudar a determinar que una enfermedad no está causada principalmente por factores ambientales. Por ejemplo, Harris señaló que la naturaleza exacta de la mutación de p^{53} puede ser útil para indicar que un cierto tipo de cáncer no se debe a un carcinógeno ambiental, sino que está causado por mutagénesis endógena, como se observó en el estudio que se acaba de describir de pacientes con cáncer de cabeza y cuello que eran no bebedores y no fumadores[32]. Las mutaciones de p^{53} en la línea germinal también pueden indicar que una persona tiene una mayor susceptibilidad al cáncer, como propuso inicialmente Knudson en 1971[33].

Sin embargo, debe recordarse que el poder estadístico para detectar los efectos ambientales o genéticos generalmente requerirá un tamaño de muestra mucho mayor en comparación con lo que se requiere para detectar los efectos marginales del factor de riesgo genético o ambiental[28]. La regla de oro es que es necesario aumentar cuatro veces el tamaño de la muestra para detectar interacciones, y en ocasiones incluso las formas simples de interacción pueden ser difíciles de confirmar en los conjuntos de datos. Esto es especialmente cierto cuando hay un error sustancial en la medición del factor de riesgo ambiental. Aun así, si se pudiera obtener evidencia de interacción genético-ambiental, se abrirían nuevas oportunidades para intervenciones de salud pública, porque es mucho más fácil modificar los factores de riesgo ambiental que alterar la estructura genética de una población.

Medicina de precisión

La tecnología de secuenciación está impulsando la era actual de la epidemiología genética. La secuenciación tiene la ventaja de identificar **todas** las variantes en una región determinada del genoma (es decir, variantes de nucleótidos simples raras y de baja frecuencia, así como SNP marcadores comunes). Esta nueva fase ha sido impulsada por dos fuerzas: 1) los avances en la tecnología de secuenciación masiva «de próxima generación» paralela, que se está volviendo asequible incluso para los tamaños de muestras observados en estudios epidemiológicos y 2) la finalización inminente de análisis GWAS combinados para muchas (si no la mayoría) de las enfermedades complejas en estudios extensos, que han identificado numerosos genes que influyen significativamente en el riesgo. No hemos alcanzado nuestro objetivo de comprender completamente las enfermedades complejas (y podríamos estar lejos de ello) si las vías causales implican interacciones mecanicistas gen-gen y gen-ambiente. Hemos pasado de identificar por primera vez la doble hélice a mediados del siglo XX a la capacidad de secuenciar todo el genoma de muchos individuos en el primer cuarto del siglo XXI.

La secuenciación de todo el genoma es particularmente apasionante debido a su promesa de profundizar nuestro conocimiento de los genes y por su potencial para facilitar el desarrollo de «tratamientos personalizados» de pacientes individuales. Anteriormente expusimos algunas limitaciones de los ensayos aleatorizados para desarrollar nuevas modalidades de tratamiento, en particular porque los ensayos habitualmente tratan con grupos en lugar de con individuos y, por lo tanto, por lo general estiman efectos promedio. Por lo tanto, como se señaló anteriormente, los resultados del estudio a menudo se refieren a grupos y dejan al médico tratante sin información sobre la probabilidad de que un paciente determinado al que está tratando se beneficie del nuevo medicamento o si el paciente sufrirá efectos secundarios graves de este. Sin embargo, con el advenimiento de la nueva era actual en genética humana, se tiene esperanza en que podremos desarrollar terapias personalizadas para cada paciente en función de las características de su genoma. Por ejemplo, el fármaco clopidogrel se utiliza después de la colocación de una endoprótesis coronaria tras un ataque al corazón. El clopidogrel deprime la actividad de las plaquetas, lo que reduce el riesgo de ataque cardiaco. La pérdida de la función de las variantes genéticas en el gen CYP2C19 produce una disminución en la producción de metabolitos y se asocia con un mayor riesgo de un ataque cardiaco posterior[34,35], incluso cuando los pacientes son tratados con clopidogrel. Sin embargo, el ajuste de la dosis de clopidogrel en pacientes que poseen estos alelos CYP2C19 (que se encuentran en casi un tercio de la población general estadounidense) o el cambio de la medicación a otro fármaco disminuyen el riesgo de ataques cardiacos posteriores a un nivel similar al de pacientes sin estos alelos[36-38].

Por lo tanto, la información genética generada a partir de poblaciones puede emplearse para tratar a un paciente individual.

Perspectivas de futuro

A pesar de la expectación suscitada por la secuenciación del genoma humano y por los resultados de estudios del genoma completo como los descritos anteriormente, en la mayoría de las enfermedades complejas (en las que se han implicado factores genéticos y ambientales), la información disponible en la actualidad aún no es suficiente para delinear específicamente cómo controlan los genes el riesgo. Una mayor comprensión de las modificaciones moleculares en el cáncer derivada de los estudios de los cambios genéticos en las células cancerosas debería mejorar nuestros conocimientos sobre la susceptibilidad individual a padecer cáncer y facilitar el desarrollo de tratamientos específicos para las vías biológicas involucradas en diferentes tumores. Estos tratamientos «dirigidos» o «individualizados» pueden ayudar al conocimiento y al tratamiento de subgrupos de tumores específicos de tejidos. Al dirigirse a las vías moleculares específicas implicadas en diferentes tumores, así como a los puntos en los que las células tumorales pueden ser especialmente vulnerables a ciertas intervenciones, los tratamientos individualizados deberían ser más eficaces. Dichos tratamientos también pueden tener menos efectos secundarios y que estos sean menos graves que muchos de los tratamientos convencionales, cuyos efectos citotóxicos no son lo bastante específicos, por lo que pueden afectar tanto a las células patológicas como a las sanas.

Childs ha planteado un concepto que abarca no solo las diferentes características de tumores distintos desde el punto de vista histológico, sino también las características genéticas y ambientales únicas de los seres humanos que pueden haber provocado la vulnerabilidad a este tipo de tumores[39]. Como resultado, lo que podría parecer a primera vista una misma enfermedad que se presenta en diferentes individuos, tal vez se deba considerar como diferentes enfermedades con el mismo fenotipo, ya que la enfermedad en una persona es una asociación de anomalías físicas, de laboratorio y de otro tipo, en combinación con un conjunto único de factores del huésped de tipo genético y ambiental, que influyen en la susceptibilidad global. Estas susceptibilidades a menudo pueden incluir factores sociales y psicológicos, además de factores ambientales conocidos. Estos factores pueden estar actuando a nivel del individuo, la familia, la comunidad o de algún otro grupo social amplio. Aunque esta combinación será diferente de un individuo a otro, según las definiciones y clasificaciones actuales de las enfermedades, muchas personas parecen tener la misma enfermedad. La integración del conocimiento de todas estas áreas divergentes podría proporcionar la base para la detección precoz de las personas de alto riesgo y dar lugar a medidas más eficaces de prevención precoz en los próximos años.

Childs y Valle escribieron en el año 2000 lo siguiente:

Los signos y síntomas que presenta un paciente en la actualidad podrían haber sido forjados en la matriz del desarrollo y la maduración del pasado. Al realizar esta caracterización, discernimos la individualidad y la heterogeneidad de aquello que denominamos enfermedad…

En medicina, tenemos problemas para aceptar este tipo de individualidad. Cuando vemos a un paciente, pensamos en primer lugar en el nombre de una enfermedad y después en la variación expresada en el paciente. Esta forma de pensar es tipológica y ha de distinguirse del pensamiento poblacional en el que una población, por ejemplo, de pacientes con la «misma» enfermedad se compone de individuos variables[40].

Dalton y Friend publicaron una presentación esquemática de la naturaleza cíclica del proceso de incorporación de nuevos conocimientos a terapias que son individualizadas para cada paciente (fig. 16.13); el proceso se describe en el pie de esta figura [41]. A pesar de que este planteamiento tiene un gran potencial, en general, sus beneficios aún no se han materializado ampliamente en el tratamiento de la mayoría de las enfermedades complejas. Sin embargo, las nuevas tecnologías a nivel molecular y genético pueden tener profundos efectos sobre la atención sanitaria y el desarrollo de la «medicina personalizada», que incluirá nuevos enfoques para la prevención y el tratamiento de enfermedades, que serán posibles gracias a avances técnicos y a la integración de nueva información procedente de diferentes disciplinas biológicas y sociológicas.



FIG. 16.13 Tratamiento oncológico personalizado como un ciclo continuo. El ciclo comienza con el descubrimiento de alteraciones moleculares específicas en los tumores que luego se relacionan con resultados de pacientes específicos en los ensayos clínicos. La capacidad de recopilar los perfiles moleculares y la información clínica a nivel de los pacientes individuales permite traducir la información en un tratamiento oncológico más personalizado. Las bases de datos relacionales disponibles y los sistemas de información sanitaria garantizan una aplicación más fundamentada de terapias antineoplásicas para futuros pacientes y también pueden guiar el descubrimiento de nuevos tratamientos. (De Dalton WS, Friend SH. Cancer biomarkers—an invitation to the table. *Science.* 2006;312:1165–1168.)

Conclusión

En este capítulo se han descrito algunos de los enfoques epidemiológicos utilizados para evaluar las contribuciones relativas de los factores genéticos y ambientales a la etiología de las enfermedades humanas. La relación de la epidemiología y la genética cada vez es más conocida, y ha surgido un campo denominado *epidemiología genética*[19]. Se han publicado trabajos excelentes sobre el impacto de la era genómica en la investigación epidemiológica[42,43].

La mayoría de los estudios epidemiológicos se dirigen a la identificación de los factores ambientales que controlan el riesgo de las enfermedades, pero a la hora de diseñar y llevar a cabo estudios, así como al interpretar sus resultados, se debe tener en cuenta que los sujetos de los estudios epidemiológicos difieren no solo en cuanto a las exposiciones ambientales, sino también en su dotación genética, lo que también influye sobre el riesgo. Cuando sea oportuno, los estudios epidemiológicos de los factores de riesgo, incluidos los estudios de casos y controles y otros diseños de estudios, deberían ampliarse para incluir la recopilación de los antecedentes familiares y la obtención de muestras biológicas, si es posible. La incorporación de avances genéticos y marcadores genéticos en los estudios epidemiológicos está demostrando ser cada vez más valiosa en la identificación de subgrupos de alto riesgo y en la personalización de terapias específicas para el individuo. Es probable que se vuelvan cada vez más importantes para mejorar la prevención de enfermedades en el futuro.

Bibliografía

1. Massie J, Ioannou L, Delatycki M. Prenatal and preconception population carrier screening for cystic fibrosis in Australia: where are we up to? *Aust N Z J Obstet Gynaecol*. 2014;54(6):503–509.
2. Ashley-Koch A, Yang Q, Olney RS. Sickle hemoglobin (HbS) allele and sickle cell disease: a HuGE review. *Am J Epidemiol*. 2000;151(9):839–845.
3. Therrell Jr BL, Lloyd-Puryear MA, Eckman JR, et al. Newborn screening for sickle cell diseases in the United States: a review of data spanning 2 decades. *Semin Perinatol*. 2015;39(3):238–251.
4. Minkovitz CS, Grason H, Ruderman M, et al. Newborn screening programs and sickle cell disease: a public health services and systems approach. *Am J Prev Med*. 2016;51(1 suppl 1):S39–S47.
5. Mak CM, Lee HC, Chan AY, et al. Inborn errors of metabolism and expanded newborn screening: review and update. *Crit Rev Clin Lab Sci*. 2013;50(6):142–162.
6. Hogben L. Nature and Nurture. New York: WW Norton; 1939.
7. Jörgensen G, Lenz W, Pfeiffer RA, et al. Thalidomide-embryopathy in twins: a collaborative study. *Acta Genet Med Gemellol (Roma)*. 1970;19(1):205–210.
8. Grasby KL, Verweij KJH, Mosing MA, et al. Estimating heritability from twin studies. *Methods Mol Biol*. 2017;1666:171–194.
9. Kaij L. Studies on the Etiology and Sequels of Abuse of Alcohol. Lund: Hakan Ohlssons Boktryckeri; 1960.
10. Hrubec Z, Omenn GS. Evidence of genetic predisposition to alcoholic cirrhosis and psychosis: twin concordances for alcoholism and its biological end points by zygosity among male veterans. *Alcohol Clin Exp Res*. 1981;5:207–215.
11. Murray RM, Clifford C, Gurlin HM, Twin and alcoholism studies. Galanter M, ed. Recent Developments in Alcoholism, Vol. 1. New York: Plenum; 1983:25–47.
12. Pickens RW, Svikis DS, McGue M, et al. Heterogeneity in the inheritance of alcoholism: a study of male and female twins.

Arch Gen Psychiatry. 1991;48:19–28.
13. Lichtenstein P, Holm NV, Verkasalo PK, et al. Environmental and heritable factors in the causation of cancer: analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med*. 2000;343:78–85.
14. Kety SS, Ingraham LJ. Genetic transmission and improved diagnosis of schizophrenia from pedigrees of adoptees. *J Psychiatr Res*. 1992;26:247–255.
15. King MC, Marks JH, Mandell JB, et al. Breast and ovarian cancer risks due to inherited mutations in BRCA1 and BRCA2. *Science*. 2003;302(5645):643–646.
16. Prevalence and penetrance of BRCA1 and BRCA2 mutations in a population-based series of breast cancer cases. Anglian Breast Cancer Study Group. Br J Cancer. 2000;83(10):1301–1308.
17. Malaty HM. Epidemiology of Helicobacter pylori infection. *Best Pract Res Clin Gastroenterol*. 2007;21(2):205–214.
18. Wood H. Multiple sclerosis: latitude and vitamin D influence disease course in multiple sclerosis. *Nat Rev Neurol*. 2017;13(1):3.
19. Khoury MJ, Beaty TH, Cohen BH. Fundamentals of Genetic Epidemiology. New York: Oxford University Press; 1993.
20. Visscher PM, Wray NR, Zhang Q, et al. 10 years of GWAS discovery: biology, function, and translation. *Am J Hum Genet*. 2017;101(1):5–22.
21. Visscher PM, Brown MA, McCarthy MI, et al. Five years of GWAS discovery. *Am J Hum Genet*. 2012;90(1):7–24.
22. Klein RJ, Zeiss C, Chew EY, et al. Complement factor H polymorphism in age-related macular degeneration. *Science*. 2005;308(5720):385–389.
23. Manolio TA. Bringing genome-wide association findings into clinical use. *Nat Rev Genet*. 2013;14(8):549–558.
24. Khoury MJ, Evans JP. A public health perspective on a national precision medicine cohort: balancing long-term knowledge generation with early health benefit. *JAMA*. 2015;313(21):2117–2118.
25. Grant BF. The impact of a family history of alcoholism on the relationship between age at onset of alcohol use and DSM-IV alcohol dependence: results from the National Longitudinal Alcohol Epidemiologic Survey. *Alcohol Health Res World*.

1998;22:144–147.
26. Rosendaal FR, Siscovick DS, Schwartz SM, et al. Factor V Leiden (resistance to activated protein C) increases the risk of myocardial infarction in young women. *Blood*. 1997;89:2817–2821.
27. Austin MA, Schwartz SM. Cardiovascular disease. In: Costa LG, Eaton DL, eds. Gene-Environment Interactions: Fundamentals of Ecogenetics. Hoboken, NJ: John Wiley and Sons; 2006:271–283.
28. Aschard H. A perspective on interaction effects in genetic association studies. *Genet Epidemiol*. 2016;40(8):678–688.
29. Brennan JA, Boyle JO, Koch WM, et al. Association between cigarette smoking and mutation of the p53 gene in squamous-cell carcinoma of the head and neck. *N Engl J Med*. 1995;332:712–717.
30. Bressac B, Puisieux MS, Kew M, et al. p53 mutation in hepatocellular carcinoma after aflatoxin exposure. *Lancet*. 1991;338:1356–1359.
31. Aguilar F, Harris CC, Sun T, et al. Geographic variation of p53 mutational profile in nonmalignant human liver. *Science*. 1994;264:1317–1319.
32. Harris C. p53: at the crossroads of molecular carcinogenesis and risk assessment. *Science*. 1993;262:1980–1981.
33. Knudson Jr AG. Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci USA*. 1971;68:820–823.
34. Shuldiner AR, O'Connell JR, Bliden KP, et al. Association of cytochrome P450 2C19 genotype with the antiplatelet effect and clinical efficacy of clopidogrel therapy. *JAMA*. 2009;302(8):849–857.
35. Mega JL, Simon T, Collet JP, et al. Reduced-function CYP2C19 genotype and risk of adverse clinical outcomes among patients treated with clopidogrel predominantly for PCI: a meta-analysis. *JAMA*. 2010;304(16):1821–1830.
36. Cavallari LH, Lee CR, Beitelshees AL, et al. Multisite investigation of outcomes with implementation of CYP2C19 genotype-guided antiplatelet therapy after percutaneous coronary intervention. *JACC Cardiovasc Interv*. 2018;11(2):181–191.
37. Shuldiner AR, Palmer K, Pakyz RE, et al. Implementation of pharmacogenetics: the University of Maryland Personalized

Anti-platelet Pharmacogenetics Program. *Am J Med Genet C Semin Med Genet*. 2014;166C(1):76–84.
38. Scott SA, Sangkuhl K, Stein CM, et al. Clinical Pharmacogenetics Implementation Consortium. Clinical Pharmacogenetics Implementation Consortium guidelines for CYP2C19 genotype and clopidogrel therapy: 2013 update. *Clin Pharmacol Ther*. 2013;94(3):317–323.
39. Childs B. Genetic Medicine—A Logic of Disease. Baltimore: Johns Hopkins University Press; 1999.
40. Childs B, Valle D. Genetics, biology and disease. *Annu Rev Genomics Hum Genet*. 2000;1:1–19.
41. Dalton WS, Friend SH. Cancer biomarkers—an invitation to the table. *Science*. 2006;312:1165–1168.
42. Millikan R. The changing face of epidemiology in the genomics era. *Epidemiology*. 2002;13:472–480.
43. Willett WC. Balancing life-style and genomics research for disease prevention. *Science*. 2002;296:695–698.

Preguntas de repaso del capítulo 16

1. *La observación de que una mayor proporción de pares de gemelos monocigóticos que dicigóticos son concordantes para una determinada enfermedad sugiere que la enfermedad está causada muy probablemente por:*

a. Factores exclusivamente ambientales.

b. Factores exclusivamente genéticos.

c. Factores hereditarios casi exclusivamente, con la posibilidad de una cierta participación de factores no hereditarios.

d. Factores ambientales y genéticos casi por igual.

e. Diferencias de sexo en los gemelos monocigóticos.

2. *Cuando se estudia la incidencia de una enfermedad en niños adoptados y se compara con su incidencia en familiares biológicos y en familiares adoptivos, todos los siguientes aspectos son relevantes,* excepto:

a. Edad de inicio.

b. Grado de contacto que mantiene el adoptado con sus progenitores biológicos.

c. Estado civil de los progenitores biológicos.

d. Factores de selección relativos a quién es adoptado y quién no lo es.

e. *c* y *d.*

La pregunta 3 se basa en la siguiente información:

En un estudio familiar sobre la esquizofrenia se observaron las siguientes tasas de concordancia en varios pares de familiares:

Par	**Tasa de concordancia (%)**
Marido-mujer	5
Progenitor-hijo	40
Gemelos monocigóticos	65
Gemelos dicigóticos	42
Hermanos no gemelos	40

3. *Una conclusión razonable que puede extraerse de estos datos es:*

a. Los factores genéticos son irrelevantes en la etiología de la esquizofrenia.

b. Los datos sugieren un componente genético potencialmente importante.

c. La incidencia de la esquizofrenia en pares de familiares es mayor en los gemelos monocigóticos.

d. La prevalencia de la esquizofrenia en pares de familiares es mayor en los gemelos monocigóticos.

e. Los gemelos son menos propensos a tener esquizofrenia que los hermanos no gemelos.

La pregunta 4 se basa en la siguiente información:

En un estudio de los inmigrantes japoneses a Estados Unidos se encontraron las siguientes tasas de mortalidad estandarizada (TME) para la enfermedad X:

Grupo	**Tasa de mortalidad estandarizada**
Japoneses nativos que viven en Japón	100
Inmigrantes japoneses	105
Niños de ascendencia japonesa	108
Estadounidenses de raza blanca	591

4. *Estos hallazgos sugieren que:*

a. Los factores ambientales son los principales determinantes de estas TME.

b. Los factores genéticos son los principales determinantes de estas

TME.

c. Los factores ambientales asociados con la cultura del inmigrante probablemente estén implicados.

d. Los inmigrantes están muy seleccionados y no son representativos de la población de su país de origen.

e. Las diferencias internacionales en la codificación de los certificados de defunción para la enfermedad X son un determinante importante de estas TME.

5. *Si se encuentra una asociación entre la incidencia de una enfermedad y una cierta característica determinada genéticamente:*

a. La enfermedad es claramente de origen genético.

b. Los factores genéticos están implicados en todos los casos de la enfermedad.

c. Los factores genéticos están implicados en al menos algunos casos de la enfermedad.

d. Se descarta la implicación de factores ambientales.

e. Es probable que la expresión de la enfermedad sea inevitable.

Glosario de los términos genéticos del capítulo 16

Término	Definición
ADN (ácido desoxirribonucleico)	ADN es el nombre químico de la molécula que contiene las instrucciones genéticas en todos los seres vivos. La molécula de ADN consta de dos cadenas que se enrollan una sobre otra para dar lugar a una forma conocida como doble hélice. Cada hebra consta de un esqueleto compuesto por azúcares (desoxirribosa) alternando con grupos fosfato. Unida a cada azúcar se encuentra una de las cuatro bases: adenina (A), citosina (C), guanina (G) y timina (T). Las dos cadenas se mantienen unidas por enlaces entre las bases: la adenina se une a la timina y la citosina a la guanina. Las secuencias de las bases a lo largo de los esqueletos sirven de instrucciones para el ensamblaje de proteínas y moléculas de ARN
Alelo	Un alelo es una de dos o más versiones de un gen. Un individuo hereda dos alelos para cada gen, uno de cada progenitor. Si los dos alelos son iguales, el individuo es homocigoto para ese gen. Si los alelos son diferentes, el individuo es heterocigoto. Aunque el término alelo se usó originalmente para describir las variaciones entre los genes, ahora también hace referencia a la variación entre las secuencias de ADN no codificantes
Autismo	El autismo es un trastorno del desarrollo cerebral caracterizado por interacciones sociales deficientes, problemas de comunicación y conductas repetitivas. Los síntomas suelen aparecer antes de los 3 años de edad. La causa exacta del autismo no se conoce; sin embargo, es probable que participen factores genéticos. El autismo pertenece a un grupo de trastornos del desarrollo relacionados, llamados trastornos del espectro autista (TEA). Otros TEA incluyen el síndrome de Asperger y el síndrome de Rett
Autosómico dominante	El patrón autosómico dominante es un patrón de herencia característico de algunas enfermedades genéticas. «Autosómico» significa que el gen en cuestión se encuentra en uno de los cromosomas numerados o no sexuales. «Dominante» significa que una sola copia de la mutación asociada a la enfermedad es suficiente para causar la enfermedad. Esto contrasta con los trastornos recesivos, donde se necesitan dos copias de la mutación para causar la enfermedad. La enfermedad de Huntington es un ejemplo común de un trastorno genético autosómico dominante
Consejo genético	El consejo genético es la interacción profesional entre un profesional sanitario con conocimientos especializados de genética y una persona o familia. El asesor genético determina si una enfermedad en la familia puede ser genética y estima las posibilidades de que otro familiar pueda verse afectado. Los asesores genéticos también ofrecen e interpretan pruebas genéticas que pueden ayudar a estimar el riesgo de enfermedad. El asesor genético proporciona información en un esfuerzo por abordar las inquietudes del cliente y brinda asesoramiento psicológico para ayudar a las familias a adaptarse a su enfermedad o riesgo
Cribado de portadores	La prueba de detección de portadores es un tipo de prueba genética que se realiza en personas que no muestran síntomas de un trastorno genético pero que pueden tener riesgo de transmitirlo a su progenie. Un portador de un trastorno genético ha heredado un alelo normal y uno anormal para un gen asociado con el trastorno. La descendencia debe heredar dos alelos anormales para que aparezcan los síntomas. Los futuros padres con antecedentes familiares de un trastorno genético son candidatos para la detección de portadores
Cribado de recién nacidos	Las pruebas de cribado en recién nacidos son pruebas realizadas en bebés recién nacidos para detectar una amplia variedad de trastornos. Normalmente, las pruebas se realizan en una muestra de sangre obtenida de un pinchazo en el talón cuando el bebé tiene 2 o 3 días de edad. En los Estados Unidos, el cribado de recién nacidos es obligatorio para varios trastornos genéticos diferentes, aunque el conjunto exacto de pruebas requeridas difiere de un estado a otro
Cribado genético	El cribado genético es el proceso por el que se evalúa a una población para detectar una enfermedad genética con el fin de identificar a un subgrupo de personas que tienen la enfermedad o la posibilidad de transmitirla a sus descendientes
Cromosoma	Un cromosoma es un paquete organizado de ADN que se encuentra en el núcleo

	celular. Diferentes organismos poseen diferente número de cromosomas. Los seres humanos tienen 23 pares de cromosomas: 22 pares de cromosomas numerados, llamados autosomas, y un par de cromosomas sexuales, X e Y. Cada progenitor contribuye con un cromosoma a cada pareja por lo que la descendencia obtiene la mitad de sus cromosomas de la madre y la otra mitad del padre
Deleción	La deleción es un tipo de mutación que implica la pérdida de material genético. Pueden ser pequeñas, como la pérdida de un único par de bases de ADN; o grandes, involucrando un fragmento de un cromosoma
Dominante	Dominante se refiere a la relación entre dos versiones de un gen. Los individuos reciben dos versiones de cada gen, conocidas como alelos, uno de cada progenitor. Si los alelos de un gen son diferentes, se expresará un alelo; el del gen dominante. El efecto del otro alelo, llamado recesivo, quedará enmascarado
Enfermedad compleja	Una enfermedad compleja es causada por la interacción de múltiples genes y factores ambientales. Las enfermedades complejas también se denominan multifactoriales. Ejemplos de enfermedades complejas incluyen el cáncer y las enfermedades cardiacas
Epidemiología genética	La epidemiología genética es una disciplina médica relativamente nueva que busca comprender cómo los factores genéticos interactúan con el medio ambiente en el contexto de la enfermedad en las poblaciones. Las áreas de estudio incluyen las causas de las enfermedades hereditarias y su distribución y control
Estudios de asociación del genoma completo (GWAS)	Un estudio de asociación del genoma completo (GWAS) es un enfoque utilizado en la investigación genética para asociar variaciones genéticas específicas con enfermedades particulares. El método consiste en estudiar los genomas de muchas personas diferentes y buscar marcadores genéticos que puedan emplearse para predecir la presencia de una enfermedad. Una vez identificados dichos marcadores genéticos, se pueden usar para comprender cómo contribuyen los genes a la enfermedad y desarrollar mejores estrategias preventivas y terapéuticas
Fenotipo	El fenotipo es el conjunto de rasgos observables de un individuo, como la altura, el color de ojos y el grupo sanguíneo. La contribución genética al fenotipo se llama genotipo. Algunos rasgos están determinados en gran medida por el genotipo, mientras que otros están determinados principalmente por factores ambientales
Gemelos fraternales	Los gemelos fraternales también son gemelos dicigóticos. Son el resultado de la fertilización de dos óvulos distintos durante el mismo embarazo. Los gemelos fraternales pueden ser del mismo sexo o diferente. Comparten la mitad de sus genes como cualquier otro hermano. Por el contrario, los gemelos que resultan de la fertilización de un único óvulo que luego se divide en dos se llaman gemelos monocigóticos, o idénticos. Los gemelos idénticos comparten todos sus genes y son siempre del mismo sexo
Gemelos idénticos	Los gemelos idénticos también se conocen como gemelos monocigóticos. Resultan de la fertilización de un solo óvulo, que se divide en dos. Los gemelos idénticos comparten todos sus genes y son siempre del mismo sexo. En comparación, los gemelos fraternales o dicigóticos resultan de la fertilización de dos óvulos distintos durante el mismo embarazo. Comparten la mitad de sus genes, como cualquier otro hermano. Los gemelos fraternales pueden ser del mismo sexo o diferente
Gen	El gen es la unidad física básica de la herencia. Los genes se transmiten de los progenitores a los descendientes y contienen la información necesaria para especificar rasgos. Los genes están dispuestos, uno tras otro, en estructuras llamadas cromosomas. Un cromosoma contiene una sola molécula de ADN larga, de la que solo una parte corresponde a un único gen. Los humanos poseen aproximadamente 20.000 genes dispuestos en sus cromosomas
Gen candidato	Un gen candidato es un gen cuya ubicación cromosómica está asociada con una enfermedad particular u otro fenotipo. Debido a su ubicación, se sospecha que el gen causa la enfermedad u otro fenotipo
Genoma	El genoma es el conjunto completo de instrucciones genéticas que se encuentran en una célula. En los seres humanos, el genoma consta de 23 pares de cromosomas, que se encuentran en el núcleo, así como un pequeño cromosoma que se localiza en las mitocondrias de las células. Cada conjunto de 23 cromosomas contiene aproximadamente 3.100 millones de bases de secuencia de ADN
Genómica	La genómica se refiere al estudio del genoma completo de un organismo, mientras que la genética se refiere al estudio de un gen en particular.
Genotipo	El genotipo es la colección de genes de un individuo. El término también puede referirse a los dos alelos heredados para un gen en particular. El genotipo se expresa cuando la información codificada en el ADN de los genes se utiliza para sintetizar

	moléculas de proteínas y ARN. La expresión del genotipo contribuye a los rasgos observables del individuo, el llamado fenotipo
Hereditario	Un rasgo heredado es aquel que está genéticamente determinado. Los rasgos heredados se pasan de los progenitores a la descendencia según las reglas de la genética mendeliana. La mayoría de los rasgos no están estrictamente determinados por los genes, sino que están influenciados tanto por los genes como por el entorno
Herencia mendeliana	La herencia mendeliana se refiere a los patrones de herencia que son característicos de los organismos que se reproducen sexualmente. El monje austriaco Gregor Mendel realizó miles de cruces con guisantes en su monasterio a mediados del siglo XIX. Mendel explicó sus resultados al describir dos leyes de herencia que introdujeron la idea de genes dominantes y recesivos
Heterocigoto	El término heterocigoto se refiere a haber heredado diferentes formas de un gen particular de cada progenitor. Un genotipo heterocigoto contrasta con un genotipo homocigoto, donde un individuo hereda formas idénticas de un gen particular de cada progenitor
Homocigoto	La homocigosidad es una condición genética en la que un individuo hereda los mismos alelos para un gen particular de ambos progenitores
Interacción genético-ambiental	La interacción genético-ambiental influye en la expresión de un rasgo, que, por tanto, resulta de la interacción entre los genes y el medio ambiente. Algunos rasgos están fuertemente influenciados por los genes, mientras que otros están fuertemente influenciados por el entorno. Sin embargo, la mayoría de los rasgos están influenciados por la interacción compleja entre uno o más genes con el ambiente
Ligado a X	En un rasgo ligado a X el gen en cuestión se localiza en el cromosoma X. Los seres humanos y otros mamíferos tienen dos cromosomas sexuales, el X y el Y. En una enfermedad ligada a X o ligada al sexo, generalmente son los hombres los afectados porque tienen una sola copia del cromosoma X que porta la mutación. En las mujeres, el efecto de la mutación puede estar enmascarado por la segunda copia sana del cromosoma X
Ligado al sexo	El carácter ligado al sexo es un rasgo en el que un gen se localiza en un cromosoma sexual. En los seres humanos, el término generalmente se refiere a rasgos que están influenciados por los genes del cromosoma X. Esto se debe a que el cromosoma X es grande y contiene muchos más genes que el cromosoma Y más pequeño. En una enfermedad ligada al sexo, por lo general son los hombres los que se ven afectados porque tienen una sola copia del cromosoma X que transporta la mutación. En las mujeres, el efecto de la mutación puede estar enmascarado por la segunda copia sana del cromosoma X
Ligamiento	El ligamiento es la asociación cercana de genes u otras secuencias de ADN en el mismo cromosoma. Cuanto más cerca estén los dos genes en el cromosoma, mayor será la probabilidad de que se hereden juntos
Línea germinal	La línea germinal hace referencia a las células sexuales (óvulos y espermatozoides) que utilizan los organismos que se reproducen sexualmente para transmitir los genes de generación en generación. Los óvulos y los espermatozoides se conocen como células germinales, en contraste con el resto de células del organismo que se denominan células somáticas
Locus	Un locus es la ubicación física específica de un gen u otra secuencia de ADN en un cromosoma, como una dirección postal genética. El plural de «locus» es «loci»
Mapa genético	Un mapa genético es un tipo de mapa cromosómico que muestra las ubicaciones relativas de los genes y otras características importantes. El mapa se basa en la idea de vinculación, que significa que cuanto más cerca estén dos genes en el cromosoma, mayor será la probabilidad de que se hereden juntos. Al seguir los patrones de herencia se establecen las ubicaciones relativas de los genes a lo largo del cromosoma
Mapeado genético	El mapeo de genes es el proceso que establece las ubicaciones de los genes en los cromosomas. Los primeros mapas genéticos utilizaron análisis de ligamiento. Cuanto más cerca estén los dos genes en el cromosoma, más probable será que se hereden juntos. Al seguir los patrones de herencia, se pueden determinar las posiciones relativas de los genes. Más recientemente, los científicos han usado técnicas de ADN recombinante (ADNr) para establecer las ubicaciones físicas reales de los genes en los cromosomas
Mapeo	El mapeo consiste en el proceso de elaborar un diagrama representativo que cataloga los genes y otras características de un cromosoma y muestra sus ubicaciones relativas. Los mapas citogenéticos se realizan utilizando fotomicrografías de cromosomas teñidos para revelar variaciones estructurales. Los

	mapas genéticos utilizan la idea de vinculación para estimar las ubicaciones relativas de los genes. Los mapas físicos, elaborados con tecnología de ADN recombinante (ADNr), muestran las ubicaciones físicas reales de puntos de referencia a lo largo de un cromosoma
Marcador genético	Un marcador genético es una secuencia de ADN con una ubicación física conocida en un cromosoma. Los marcadores genéticos pueden ayudar a vincular una enfermedad hereditaria con el gen responsable. Los segmentos de ADN cercanos entre sí en un cromosoma tienden a heredarse juntos. Los marcadores genéticos se utilizan para rastrear la herencia de un gen cercano que aún no se ha identificado, pero cuya ubicación aproximada se conoce. El marcador genético en sí mismo puede ser parte de un gen o puede no tener una función conocida
Mendel, Gregor Johann	Gregor Mendel era un monje austriaco que, en el siglo XIX, elaboró las leyes básicas de la herencia, incluso antes de que se acuñara el término «gen». En el jardín de su monasterio, Mendel realizó miles de cruces con guisantes. Mendel es considerado el fundador de la ciencia de la genética
Mutación	Una mutación es un cambio en una secuencia de ADN. Las mutaciones pueden ser el resultado de errores de copia del ADN cometidos durante la división celular, de la exposición a radiación ionizante, de la exposición a sustancias químicas llamadas mutágenos o de una infección por virus. Las mutaciones de la línea germinal se producen en los óvulos y los espermatozoides y pueden transmitirse a la descendencia, mientras que las mutaciones somáticas ocurren en el resto de las células del cuerpo y no se transmiten
Pedigrí	Un pedigrí es una representación genética del árbol genealógico que ilustra la herencia de un rasgo o enfermedad a través de varias generaciones. El pedigrí muestra las relaciones entre los miembros de la familia e indica qué individuos expresan o son portadores silentes del rasgo en cuestión
Polimorfismo	El polimorfismo implica una de dos o más variantes de una secuencia particular de ADN. El tipo más común de polimorfismo consiste en la variación en un solo par de bases. Los polimorfismos también pueden ser de un tamaño mucho mayor e involucrar tramos largos de ADN
Polimorfismo de nucleótido único (SNP)	Los SNP (*Single Nucleotide Polymorphisms*) son un tipo de polimorfismo consistente en la variación de un solo par de bases. Los científicos están estudiando cómo se correlacionan los SNP del genoma humano con la enfermedad, la respuesta a los fármacos y otros fenotipos
Portador	Un portador es un individuo que transporta y es capaz de transmitir una mutación genética asociada con una enfermedad y puede o no mostrar síntomas de la enfermedad. Los portadores están asociados con enfermedades heredadas como rasgos recesivos. Para presentar la enfermedad, un individuo debe haber heredado alelos mutados de ambos progenitores. Un individuo que tiene un alelo normal y un alelo mutado no sufre la enfermedad. Dos portadores pueden tener hijos con la enfermedad
Rasgo	Un rasgo es una característica específica de un organismo. Los rasgos pueden ser determinados por los genes, el ambiente o, más comúnmente, por las interacciones entre ellos. La contribución genética a un rasgo se denomina genotipo. La expresión exterior del genotipo se llama fenotipo
Rasgo poligénico	Un rasgo poligénico es aquel cuyo fenotipo está influenciado por más de un gen. Los rasgos que muestran una distribución continua, como la altura o el color de la piel, son poligénicos. La herencia de los rasgos poligénicos no muestra las relaciones fenotípicas características de la herencia mendeliana, aunque cada uno de los genes que contribuyen al rasgo se hereda según lo descrito por Gregor Mendel. Muchos rasgos poligénicos también están influenciados por el entorno y se denominan multifactoriales
Secuenciación del ADN	La secuenciación del ADN es una técnica de laboratorio utilizada para determinar la secuencia exacta de bases (A, C, G y T) en una molécula de ADN. La secuencia de bases del ADN transporta la información que necesita una célula para ensamblar proteínas y moléculas de ARN. La información sobre la secuencia del ADN es importante para los científicos que investigan las funciones de los genes. La tecnología de secuenciación del ADN se hizo más rápida y menos costosa como parte del Human Genome Project

Datos del National Human Genome Research Institute. Glossary of genetic terms. https://www.genome.gov/glossary/index.cfm. Consultado el 10 de

diciembre de 2017.

SECCIÓN 3

Aplicación de la Epidemiología a la Evaluación y al Desarrollo de Políticas Sanitarias

Introducción
Capítulo 17: Uso de la epidemiología para evaluar los servicios de salud
Capítulo 18: Enfoque epidemiológico en la evaluación de los programas de cribado
Capítulo 19: Epidemiología y políticas de salud pública
Capítulo 20: Aspectos éticos y profesionales en epidemiología

Introducción

En la sección 2 se revisaron los principales tipos de diseños de estudio utilizados en epidemiología y se analizó cómo se utilizan los resultados de los estudios epidemiológicos para demostrar asociaciones y derivar inferencias causales. Aunque las cuestiones metodológicas comentadas son interesantes e intrigantes, uno de los aspectos más emocionantes de la epidemiología es el hecho de que sus resultados tienen una aplicación directa a los problemas que afectan a la salud humana. Las dificultades que se presentan consisten en derivar inferencias válidas a partir de los datos generados por los estudios epidemiológicos, garantizar una comunicación adecuada y clara de los hallazgos y sus interpretaciones a los responsables políticos y el público en general, y hacer frente a los problemas éticos que surgen debido a la estrecha relación de la epidemiología con la salud humana y con la política de salud pública y clínica.

En esta sección se describe el uso de la epidemiología para evaluar tanto los servicios sanitarios (cap. 17) como los programas de cribado y de detección precoz de enfermedades (cap. 18). En estos dos capítulos también se abordan algunas de las dificultades metodológicas y conceptuales que se presentan comúnmente en ambas situaciones. Luego se pasa a otras cuestiones relacionadas con la aplicación de la epidemiología para el desarrollo de políticas sanitarias (cap. 19), incluida la relación de la epidemiología con la prevención, la evaluación de riesgos, la epidemiología en el ámbito judicial y las fuentes y los efectos de la incertidumbre.

En el último capítulo se abordan algunos de los aspectos éticos y profesionales principales que surgen tanto al realizar investigaciones epidemiológicas como al utilizar los resultados de los estudios epidemiológicos para mejorar la salud de la comunidad. Los estudios epidemiológicos son una estrategia significativa para mejorar la eficacia tanto de la atención clínica como de las intervenciones de salud pública. Algunos de los temas principales de este capítulo son las obligaciones de los investigadores de estudiar a los sujetos, la protección de la privacidad y la confidencialidad, la raza y la etnicidad en los estudios epidemiológicos, los conflictos de intereses y la interpretación de los resultados de los estudios epidemiológicos en su aplicación a los procesos de desarrollo y mejora de la política de salud en diferentes comunidades (cap. 20).

CAPÍTULO 17

Uso de la epidemiología para evaluar los servicios de salud

Palabras clave

medidas del proceso y el resultado
eficacia
efectividad y eficiencia
investigación de resultados
mortalidad evitable
indicadores de salud de Healthy People 2020

Objetivos de aprendizaje

- Distinguir las medidas de proceso de las medidas de resultado y comentar algunas de las medidas de resultado empleadas comúnmente en la investigación sobre servicios de salud.
- Definir eficacia, efectividad y eficiencia en el contexto de los servicios de salud.
- Comparar y contrastar los estudios epidemiológicos sobre la etiología de las enfermedades con los estudios epidemiológicos que evalúan los servicios de salud.
- Comentar la investigación de resultados en el contexto de los datos ecológicos y describir varios sesgos posibles asociados a los estudios epidemiológicos que evalúan los servicios de salud utilizando datos a nivel de grupo.
- Describir varios diseños de estudio posibles que se pueden utilizar para evaluar los servicios de salud

utilizando datos a nivel individual, incluidos los diseños aleatorizados y no aleatorizados.

Tal vez el ejemplo más antiguo de una evaluación es la descripción de la creación que aparece en el libro del Génesis 1:1-4, que se muestra en el original hebreo en la figura 17.1. Su traducción, con la adición de unos encabezamientos, es la siguiente:

DATOS BASALES

En el principio creó Dios los cielos y la tierra. Y la tierra estaba desordenada y vacía, y las tinieblas estaban sobre la faz del abismo.

IMPLEMENTACIÓN DEL PROGRAMA

Y Dios dijo: «Hágase la luz». Y la luz se hizo.

EVALUACIÓN DEL PROGRAMA

Y vio Dios que la luz era buena.

ACTIVIDADES ADICIONALES DEL PROGRAMA

Y separó Dios la luz de las tinieblas.

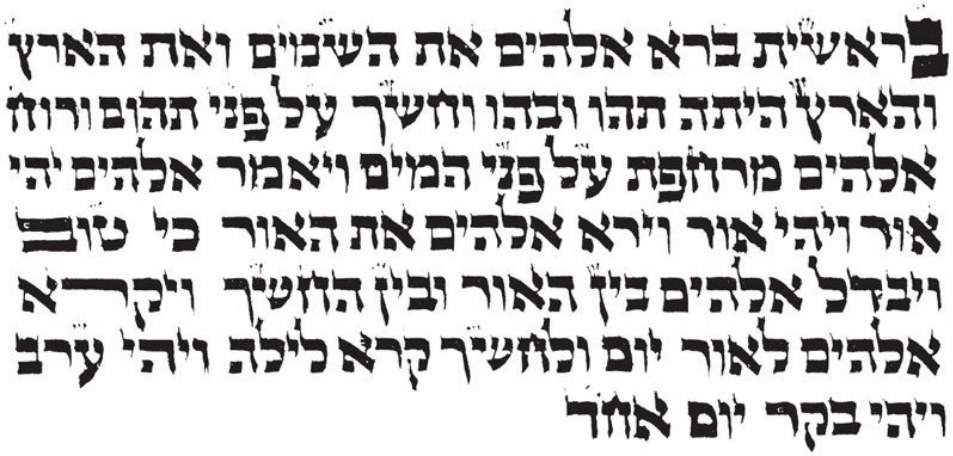

בראשית ברא אלהים את השמים ואת הארץ
והארץ היתה תהו ובהו וחשך על פני תהום ורוח
אלהים מרחפת על פני המים ויאמר אלהים יהי
אור ויהי אור וירא אלהים את האור כי טוב
ויבדל אלהים בין האור ובין החשך ויקרא
אלהים לאור יום ולחשך קרא לילה ויהי ערב
ויהי בקר יום אחד

FIG. 17.1 Evaluación más antigua conocida (Génesis 1:1-4).

Este fragmento incluye todos los componentes básicos del proceso de evaluación: datos basales, implementación del programa, evaluación del programa e implementación de nuevas actividades del programa basándose en los resultados de la evaluación. Sin embargo, surgen dos problemas en esta descripción. En primer lugar, no se ofrecen los criterios precisos que se utilizaron para determinar si el programa era «bueno»; solo se indica que Dios vio que era bueno (lo que, en retrospectiva, puede ser suficiente). En segundo lugar, esta evaluación ejemplifica un problema observado con frecuencia: el director del programa está evaluando su propio programa. Esta evaluación puede dar lugar a sesgos conscientes y subconscientes. Además, aunque el director del programa lo administre magníficamente, puede que no tenga las habilidades específicas que se requieren para llevar a cabo una evaluación del programa rigurosa desde el punto de vista metodológico.

El Dr. Wade Hampton Frost, un líder de la epidemiología de principios del siglo xx, abordó el uso de la epidemiología en la evaluación de programas de salud pública en una presentación ante la American Public Health Association en 1925[1]. Esto es parte de lo que escribió:

El funcionario de salud ocupa el puesto de un agente a quien el público confía algunos de sus recursos de dinero público y cooperación para que se inviertan de manera que puedan obtenerse los mejores rendimientos en salud; en el cumplimiento de las responsabilidades de este puesto, se espera que siga los principios generales de procedimiento que seguiría un agente fiscal en circunstancias similares…

Puesto que su capital proviene en su totalidad de los ciudadanos, es razonable esperar que esté preparado para explicarles las razones de cada inversión y para ofrecerles una cierta estimación de los rendimientos que espera. También debe considerar razonable que los ciudadanos deseen que rinda cuentas de vez en cuando para saber qué contrapartidas están

recibiendo en realidad y cómo se relacionan con las estimaciones previas que les ofreció. Sin duda, cualquier agente fiscal esperaría ser sometido a este escrutinio y ganar o perder la confianza de sus clientes a medida que sus estimaciones se verificasen o no.

Sin embargo, en lo que respecta a este rendimiento de cuentas, el funcionario de salud se encuentra en una posición difícil y posiblemente incómoda, porque, aunque puede explicar con bastante exactitud la cantidad de dinero y de esfuerzo que ha dedicado a cada una de sus actividades, pocas veces o nunca puede ser igual de exacto a la hora de rendir cuentas sobre las contrapartidas de estas inversiones consideradas por separado y de forma individual. Esto, sin lugar a dudas, no es del todo culpa suya, sino que se debe principalmente al carácter de los dividendos del esfuerzo en salud pública y a la manera en que se distribuyen. Dichos rendimientos no se reciben en cuotas separadas de una moneda uniforme, cada uno con indicaciones sobre su origen y registrado según se recibe, sino que se obtienen de forma irregular a lo largo del tiempo, distribuidos en personas no identificadas en toda la comunidad, que no son conscientes individualmente de haberlos recibido. Son beneficios positivos en cuanto a vida añadida y mejora de la salud, pero el único registro que suele llevarse en las estadísticas de morbilidad y mortalidad es el registro parcial y negativo de mortalidad y de enfermedad debidas a ciertos tipos claramente definidos de enfermedades, sobre todo las de tipo transmisible más agudas, que constituyen solo una fracción de la morbilidad[1].

El Dr. Charles V. Chapin comentó lo siguiente sobre la presentación de Frost:

La ferviente demanda del Dr. Frost para que los procedimientos de la medicina preventiva se fundamenten en una base científica sólida es muy oportuna. De hecho, habría sido oportuna en cualquier momento de los últimos 40 años y es de temer que siga siendo igual de necesaria durante los próximos 40 años[2].

Chapin subestimó claramente el número de años; esta necesidad sigue siendo tan crucial hoy en día, casi 90 años después, como lo era en 1925.

Estudios de proceso y de resultado

Avedis Donabedian es ampliamente considerado como el autor del trabajo fundamental sobre la creación de un marco para examinar los servicios de salud en relación con la calidad de la atención. Él identificó tres factores importantes simultáneamente en juego: 1) la estructura, 2) el proceso y 3) el resultado. La estructura se relaciona con las ubicaciones físicas donde se brinda la atención, el personal, el equipo y la financiación. Limitaremos nuestra exposición aquí a los dos componentes restantes, el proceso y el resultado.

Estudios de proceso

En primer lugar, hay que distinguir entre estudios de proceso y de resultado. *Proceso* significa que decidimos lo que constituye los componentes de una asistencia, servicios o acciones preventivas correctos. Esta decisión puede ser tomada en primer lugar por un panel de expertos. A continuación, se puede evaluar una clínica o a un profesional sanitario, mediante la revisión de los registros pertinentes o mediante la observación directa, y determinar en qué medida la asistencia recibida cumple con los criterios establecidos y aceptados. Por ejemplo, en la atención primaria se puede determinar en qué porcentaje de pacientes se midió la presión arterial. El problema con estas medidas de proceso es que no indican si el paciente se encuentra mejor; por ejemplo, la monitorización de la presión arterial no asegura que la presión arterial del paciente esté controlada o que el paciente tomará sistemáticamente fármacos antihipertensivos si se los recetan. En segundo lugar, debido a que las evaluaciones del proceso se basan a menudo en la opinión de expertos, los criterios utilizados en las evaluaciones del proceso pueden cambiar con el tiempo a medida que varía la opinión de los expertos. Por ejemplo, en la década de 1940, el estándar de asistencia aceptado para los bebés prematuros requería que se les administrase oxígeno al 100%. Las incubadoras eran monitorizadas para asegurarse de que se mantenían dichos niveles. Sin embargo, cuando la investigación demostró que una concentración de oxígeno elevada desempeñaba un papel destacado en el desarrollo de retinopatía de la prematuridad (una forma de ceguera en los niños prematuros), las concentraciones de oxígeno elevadas se consideraron posteriormente inaceptables.

Estudios de resultado

Dadas las limitaciones de los estudios de proceso, el resto de este capítulo se centra en las medidas de resultado. *Resultado* indica si un paciente (o una comunidad en general) se beneficia o no de la asistencia médica proporcionada. Los resultados de salud con frecuencia se consideran del

dominio de la epidemiología. Aunque estas medidas han sido tradicionalmente la mortalidad y la morbilidad, el interés por la investigación de resultados en los últimos años ha ampliado las medidas de interés para incluir la satisfacción del paciente, la calidad de vida, el grado de dependencia y discapacidad y otras medidas similares.

Eficacia, efectividad y eficiencia

Tres términos que se encuentran a menudo en la literatura sobre evaluación de los servicios de salud son *eficacia*, *efectividad* y *eficiencia*. Estos términos se usan a menudo en asociación con los hallazgos de los ensayos aleatorizados.

Eficacia

¿El agente o la intervención «funcionan» en condiciones ideales «de laboratorio»? Un nuevo fármaco se prueba en un grupo de pacientes que han aceptado ser hospitalizados y que se observan mientras toman su tratamiento, o una vacuna se prueba en un grupo de sujetos que dan su consentimiento. La eficacia es, por tanto, una medida en una situación en la que se controlan todas las condiciones para maximizar el efecto del agente. En general, las condiciones «ideales» son aquellas que se producen al probar un nuevo agente de intervención mediante un ensayo aleatorizado.

Efectividad

Si administramos el agente en una situación de «la vida real», ¿es efectivo? Por ejemplo, cuando una vacuna se prueba en una comunidad, muchas personas no pueden acudir para ser vacunadas. O un medicamento oral puede tener un sabor tan desagradable que nadie lo tomará (de modo que resultará inefectivo), a pesar de que, en condiciones controladas, cuando se garantizó el cumplimiento, el fármaco demostró ser eficaz.

Eficiencia

Si se demuestra que un agente es efectivo, ¿cuál es la relación coste-beneficio? ¿Es posible alcanzar nuestros objetivos de modo más barato y mejor? El coste incluye no solo dinero, sino también las molestias, el dolor, el absentismo, la discapacidad y el estigma social.

Si no se ha demostrado que una medida de asistencia médica es efectiva, no tiene mucho sentido evaluar su eficiencia, ya que, si no es efectiva, la alternativa más barata tampoco se debe usar. A veces, por supuesto, las presiones políticas y sociales pueden impulsar un programa incluso si no es eficaz (un ejemplo a menudo citado es DARE: Drug Abuse Resistence Education, que nunca se ha demostrado que tenga un impacto en el uso de drogas en adolescentes y adultos jóvenes). No obstante, este capítulo se centrará solo en la ciencia de la evaluación y, en concreto, en la cuestión de la efectividad al evaluar los servicios de salud.

Medidas de resultado

Si se ha demostrado la eficacia de una medida, es decir, si se ha demostrado que los métodos de prevención e intervención que son de interés funcionan, se puede pasar a evaluar la efectividad. ¿Qué pautas se deben utilizar para seleccionar una medida de resultado apropiada que sirva como índice de efectividad? En primer lugar, la medida debe ser claramente cuantificable; es decir, se debe poder expresar su efecto en términos cuantitativos. En segundo lugar, la medida de resultado debería ser relativamente fácil de definir y diagnosticar. Si la medida se va a utilizar en un estudio poblacional, es indudable que no se querrá depender de un procedimiento invasivo para la evaluación de los beneficios. En tercer lugar, la medida seleccionada debe prestarse a la estandarización con fines de estudio. En cuarto lugar, la población atendida (y la población de comparación) debe estar en situación de riesgo para la misma afección para la que se está evaluando una intervención. Por ejemplo, es evidente que tendría poco sentido evaluar la efectividad de un programa de cribado de la drepanocitosis en una población de raza blanca en Norteamérica (ya que la drepanocitosis afecta principalmente a afroamericanos).

El tipo de criterio de valoración del resultado de salud que se seleccione debería depender de la pregunta que se esté planteando. Aunque esto puede parecer obvio, no siempre es evidente de inmediato. En el cuadro 17.1 se muestran los posibles criterios de valoración a la hora de evaluar la efectividad de un programa de vacunación. Sea cual sea el resultado que se seleccione, debe establecerse de forma explícita para que quienes lean el informe de los hallazgos sean capaces de sacar sus propias conclusiones en cuanto a la idoneidad de la medida seleccionada y a la calidad de los datos. El hecho de que la medida que se ha seleccionado sea en realidad apropiada depende de los aspectos clínicos y de salud pública de la enfermedad o del estado de salud en cuestión.

Cuadro 17.1 Algunos criterios de valoración posibles que pueden usarse para medir el éxito de un programa de vacunación

1. Número (o proporción) de personas vacunadas.
2. Número (o proporción) de personas de (alto) riesgo que están vacunadas.
3. Número (o proporción) de personas vacunadas que muestran respuesta serológica.
4. Número (o proporción) de personas vacunadas y posteriormente expuestas en las que la enfermedad clínica no se desarrolla.
5. Número (o proporción) de personas vacunadas y posteriormente

expuestas en las que la enfermedad clínica o subclínica no se desarrolla.

En el cuadro 17.2 se muestran las posibles medidas que pueden escogerse para evaluar la efectividad de un programa de cultivo faríngeo infantil. Las medidas del volumen de servicios prestados, el número de cultivos tomados y el número de visitas a la clínica han sido las opciones tradicionales empleadas, ya que son relativamente fáciles de determinar y son útiles para justificar las solicitudes de incrementos presupuestarios para el programa en el año siguiente. Sin embargo, tales medidas son todas medidas de proceso y no informan sobre la efectividad de una intervención. Por tanto, pasaremos a otras posibilidades enumeradas en este cuadro. Una vez más, las medidas más apropiadas deberían depender de la pregunta que se plantea. Dicha pregunta debe ser específica. No basta solo con preguntar si el programa es bueno.

Cuadro 17.2 Algunos criterios de valoración posibles que pueden usarse para medir el éxito de un programa de cultivo faríngeo

1. Número de cultivos tomados (pacientes sintomáticos o asintomáticos).
2. Número (o proporción) de cultivos positivos para la infección estreptocócica.
3. Número (o proporción) de personas con cultivos positivos para las que se obtiene asistencia médica.
4. Número (o proporción) de personas con cultivos positivos en las que se prescribe y se realiza el tratamiento apropiado.
5. Número (o proporción) de cultivos positivos seguidos de una recidiva.
6. Número (o proporción) de cultivos positivos seguidos de fiebre reumática.

Comparación entre los estudios epidemiológicos sobre la etiología de la enfermedad y la investigación epidemiológica que evalúa la efectividad de los servicios de salud

En los estudios epidemiológicos clásicos sobre la etiología de la enfermedad se analiza la posible relación entre una causa supuesta (la variable independiente o «exposición») y un efecto o efectos adversos para la salud (la variable dependiente o «resultado»). De este modo, se tienen en cuenta otros factores, como la asistencia sanitaria, que pueden modificar la relación o actuar como factores de confusión (fig. 17.2A). En la investigación sobre los servicios de salud, nos centramos en el servicio de salud como variable independiente, con una reducción de los efectos adversos para la salud como el resultado previsto (variable dependiente) si la modalidad de asistencia es efectiva. En esta situación, también se tienen en cuenta los factores ambientales y de otro tipo que puedan influir en la relación (v. fig. 17.2B). Así pues, tanto la investigación epidemiológica etiológica como la investigación sobre los servicios de salud abordan la posible relación entre una variable independiente y una variable dependiente, así como la influencia de otros factores sobre dicha relación. Por tanto, no es sorprendente que muchos de los diseños de estudio comentados sean comunes tanto a la investigación epidemiológica como a la de los servicios de salud, al igual que los problemas metodológicos y los sesgos potenciales que pueden caracterizar a este tipo de estudios.

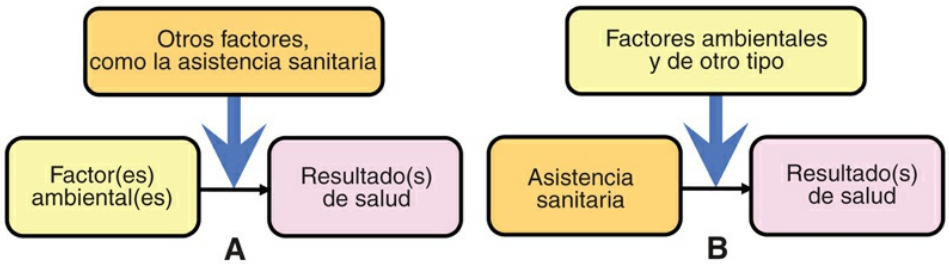


FIG. 17.2 (A) Investigación epidemiológica clásica sobre la etiología, teniendo en cuenta la posible influencia de otros factores, como la asistencia sanitaria. (B) Investigación clásica de servicios de salud sobre la efectividad, teniendo en cuenta la posible influencia de factores ambientales y de otro tipo.

Evaluación mediante el uso de datos de grupo

Los datos que suelen estar disponibles, como las cifras de mortalidad y de hospitalización, a menudo se utilizan en los estudios de evaluación. Estos datos se pueden obtener de diferentes fuentes, que pueden diferir de forma significativa. Por ejemplo, la figura 17.3 muestra los cambios en la proporción estimada de la población estadounidense con enfermedades seudogripales (ILI, *influenza-like illness*) a lo largo del tiempo, *tendencias*, utilizando tres fuentes de datos diferentes: sitios de vigilancia centinela supervisados por los Centers for Disease Control and Prevention (CDC), Tendencias de la Gripe en Google *(Google Flu Trends)* y Gripe Próxima a Usted *(Flu Near You)*[3].

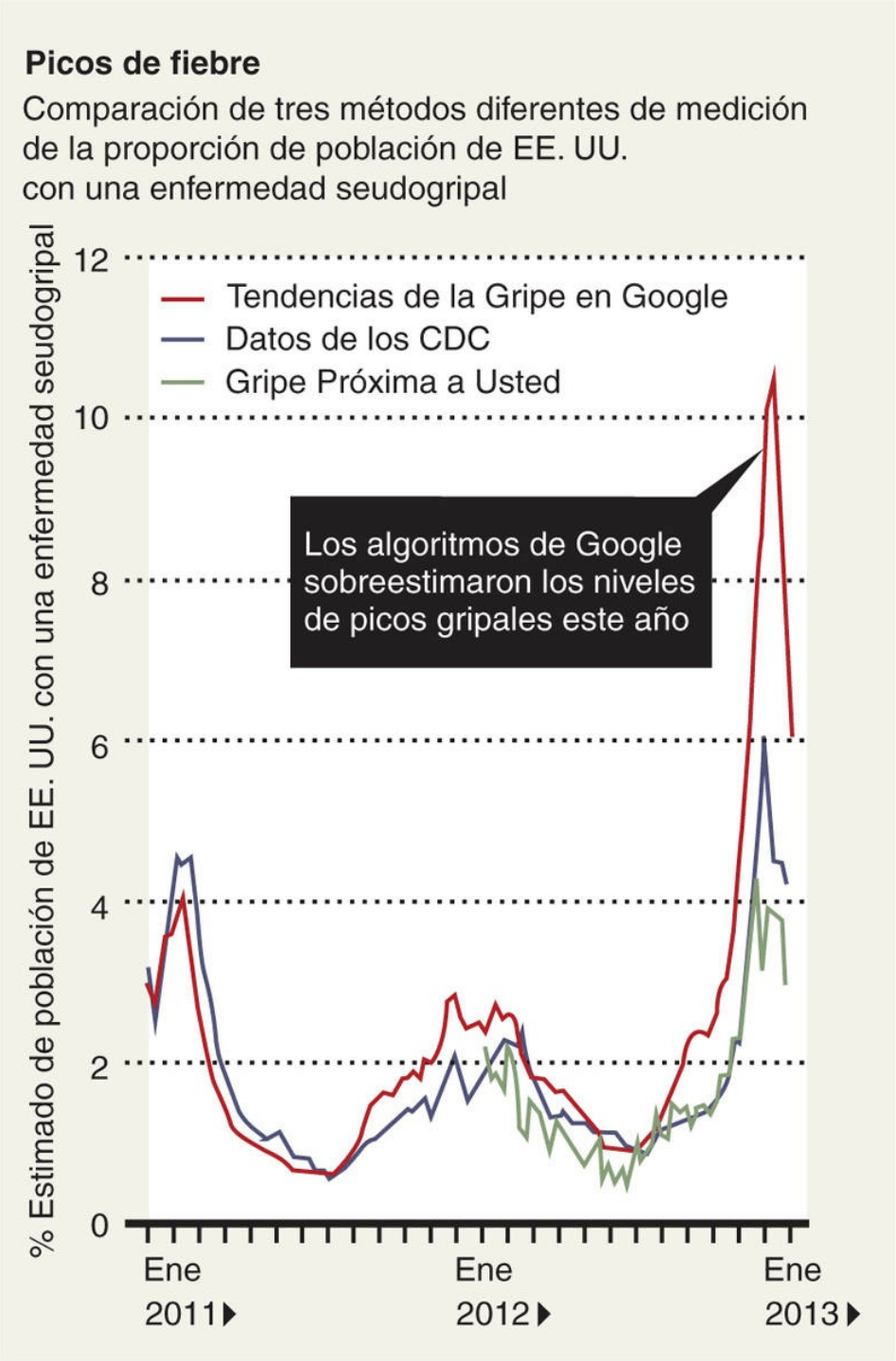


FIG. 17.3 Proporción estimada de la población de EE. UU. con una enfermedad seudogripal durante enero 2011–enero 2013. *CDC,* Centers for Disease Control and Prevention. (De Butler D. When Google got flu wrong. *Nature.* 2013;494:155–156.)

Aunque las tendencias son bastante similares en este periodo de tiempo, se puede observar que las tendencias de la gripe en Google estimaron una mayor proporción de la población de EE. UU. con ILI a finales de 2012, casi el doble que las estimaciones de los CDC. Esto se puede atribuir a la metodología variable de recolección de datos de cada fuente de datos. El CDC genera sus datos a partir de más de 2.700 centros de atención médica que

reciben más de 30 millones de consultas de pacientes cada año. La página Tendencias de la Gripe en Google emplea la metodología de minería de datos y modelado generada a partir de los términos de búsqueda relacionados con la gripe ingresados en el motor de búsqueda de Google. La página Gripe Próxima a Usted utiliza datos ingresados por usuarios de Internet, no necesariamente médicos, que ofrecen información voluntaria, para informar semanalmente si ellos o sus familiares tienen síntomas de ILI. Es posible que no todas las personas que desarrollan síntomas de ILI busquen atención médica y, por lo tanto, no sean capturadas por los datos de los CDC, pero pueden realizar una búsqueda en Google para encontrar formas de aliviar los síntomas de ILI, por ejemplo. Ya que la página Gripe Próxima a Usted depende únicamente del autoinforme voluntario de los síntomas de ILI, es posible que subestime la prevalencia. En una reciente temporada de gripe, el gobernador del estado de Nueva York, Andrew M. Cuomo, declaró una emergencia de salud pública en respuesta a una temporada de gripe severa. Se sugirió que esto podría haber provocado numerosas búsquedas en Google por parte de personas que en realidad no sufrían síntomas de ILI, lo que a su vez podría haber provocado el aumento que vemos en la figura.

Investigación de resultados

El término *investigación de resultados* se ha utilizado cada vez más para referirse a los estudios que comparan los efectos de dos o más intervenciones o modalidades de asistencia sanitaria (como tratamientos, formas de organización de la asistencia sanitaria o el tipo y el alcance de la cobertura del seguro y el reembolso al proveedor) sobre resultados de salud o económicos. Los criterios de valoración de la salud pueden incluir la morbilidad y la mortalidad, así como medidas de calidad de vida, estado funcional y percepciones del paciente de su estado de salud, incluido el reconocimiento de los síntomas y la satisfacción comunicada por el paciente. Las medidas económicas pueden reflejar los costes directos o indirectos y pueden incluir las tasas de hospitalización, la rehospitalización por la misma patología en los 30 días siguientes al alta hospitalaria, visitas a consultas externas y a urgencias, días perdidos de trabajo, cuidado de niños y días de restricción de la actividad. Por tanto, la epidemiología es una de las diversas disciplinas necesarias en la investigación de resultados.

La investigación de resultados a menudo utiliza información procedente de grandes conjuntos de datos derivados de poblaciones extensas. Aunque en los últimos años algunos de los grandes conjuntos de datos se han desarrollado a partir de cohortes que se establecieron originalmente para propósitos de investigación diferentes, muchos de los conjuntos de datos utilizados se iniciaron a menudo originalmente para fines administrativos o fiscales en lugar de para algún objetivo de investigación. Con frecuencia, varios conjuntos de datos extensos, cada uno con información sobre variables diferentes, pueden combinarse o relacionarse (resultando en «metadatos»)

con el fin de tener una muestra de un tamaño suficiente para analizar un tema de interés.

Con el advenimiento del registro médico electrónico (RME), los datos de atención al paciente están cada vez más disponibles para las comunidades de investigación en epidemiología y servicios de salud. El propósito del RME es proporcionar a los proveedores de atención médica toda la información relativa a pacientes individuales: hallazgos de las consultas médicas, utilización de servicios preventivos, medicamentos recetados, intervenciones quirúrgicas, hallazgos radiológicos, resultados de pruebas de laboratorio, continuamente a lo largo del tiempo (es decir, prospectivamente). Sin embargo, el propósito del RME no es servir como base de investigación, sino dirigir la atención médica del paciente. Aprovechar los RME para evaluar las preguntas de investigación acerca de los servicios de salud es muy prometedor, pero hasta la fecha ha resultado difícil de usar y los métodos para maximizar su potencial aún se están desarrollando y probando en el campo.

Las ventajas del uso de grandes conjuntos de datos (a veces referidos como «macrodatos») son que dichos datos se refieren a poblaciones del mundo real, lo que minimiza el problema de la «representatividad» o «generalización». Además, ya que los conjuntos de datos están presentes en el momento de iniciar la investigación, se suele poder completar el análisis y obtener los resultados de forma relativamente rápida. Además, dados los grandes conjuntos de datos utilizados, el tamaño muestral no suele ser un problema, excepto cuando se analizan los subgrupos más pequeños. Teniendo en cuenta estas consideraciones, los costes de utilizar los conjuntos de datos existentes suelen ser menores que los de la recogida de datos primarios.

Los inconvenientes son que, dado que los datos a menudo se recogieron inicialmente para la atención fiscal al paciente y con fines administrativos, puede que no sean muy adecuados para fines de investigación y para responder a la pregunta específica de la investigación planteada en el estudio. Aun cuando los datos fueron recogidos originalmente para la investigación, los conocimientos en ese campo pueden ser más completos en la actualidad y pueden haber surgido nuevas preguntas de investigación que tal vez ni siquiera se habían podido imaginar cuando se inició la recogida de datos original. En general, los datos pueden ser incompletos. Los datos sobre las variables dependientes e independientes pueden ser muy limitados. Puede que falten datos sobre detalles clínicos como la gravedad de la enfermedad y sobre los detalles de las intervenciones, y la codificación diagnóstica puede ser discordante entre los centros y dentro de un mismo centro a lo largo del tiempo. Los datos relacionados con los posibles factores de confusión pueden ser inadecuados o inexistentes, ya que la investigación que se lleva a cabo en la actualidad a menudo ni siquiera era posible cuando se generaron los datos originalmente. Debido a que ciertas variables que hoy en día se consideran relevantes e importantes no se incluyeron en el conjunto de datos original, los investigadores pueden a veces crear variables sustitutas para las que faltan,

usando ciertas variables que se incluyen en el conjunto de datos, pero que tal vez no reflejen directamente la variable de interés. No obstante, dichas variables sustitutas presentan un grado diverso de idoneidad como medida adecuada de la variable de interés. Por todas estas razones, la validez de las conclusiones alcanzadas puede ser dudosa.

Otro problema destacado que puede surgir con los grandes conjuntos de datos es que, debido a que las variables necesarias pueden estar ausentes en el conjunto de los datos disponibles, el investigador puede cambiar de forma consciente o inconsciente la pregunta que había querido plantear inicialmente por otra que tenga menor interés para él, pero para la que las variables que se necesitan para llevar a cabo el estudio están presentes en el conjunto de datos. Por tanto, en lugar de que el investigador decida cuál es la pregunta de investigación, puede que sea el conjunto de datos en sí el que termine determinando qué preguntas se plantean en el estudio.

Por último, cuando se usan grandes conjuntos de datos, los investigadores se alejan progresivamente de los sujetos estudiados. Con los años, las entrevistas directas y revisiones de registros de pacientes han tendido a sustituirse por grandes bases de datos informatizadas. Cuando se usan estas fuentes de datos, muchas de las características personales de los sujetos nunca se investigan y su relevancia para las preguntas que se plantean casi nunca se evalúa.

Una de las áreas en las que las fuentes existentes de datos se utilizan a menudo en los estudios de evaluación es la asistencia prenatal. Los problemas descritos anteriormente se ejemplifican en el uso de los certificados de nacimiento. Estos documentos se usan a menudo, ya que son de fácil acceso y proporcionan ciertos datos sobre la asistencia médica, como el trimestre en el que se comenzó la asistencia prenatal. Sin embargo, los certificados de nacimiento para las mujeres con embarazos de alto riesgo tienen datos incompletos con más frecuencia que los de las mujeres con embarazos de bajo riesgo. La calidad de los datos proporcionados en los certificados de nacimiento también puede diferir a nivel regional e internacional y puede complicar la realización de cualquier comparación.

Un ejemplo de investigación de resultados que utiliza conjuntos extensos de datos es un estudio realizado por Ikuta y cols. sobre los beneficiarios de Medicare en Estados Unidos[4]. Dado que la cobertura sanitaria de Medicare se proporciona a casi todos los ancianos de Estados Unidos (de 65 años de edad y mayores), se supone que, si una población de estudio se limita a quienes tienen cobertura de Medicare, los obstáculos económicos para la asistencia y otras variables, como la edad, el sexo o la pertenencia a subpoblaciones étnicas/raciales, se mantienen constantes entre los diferentes grupos. No obstante, aún persisten grandes disparidades entre personas de raza blanca y afroamericanas en cuanto a la utilización de muchos servicios de Medicare. Los autores estudiaron las tendencias nacionales en el uso del cateterismo de la arteria pulmonar (CAP) entre los beneficiarios de Medicare durante el periodo 1999-2013[4]. El CAP es un procedimiento mediante el cual se inserta

un tubo en una de las grandes venas del cuerpo, que es introducido hasta alcanzar el corazón y es colocado finalmente en la arteria pulmonar. Este procedimiento solía estar indicado como parte del tratamiento de rutina de la insuficiencia cardiaca y del síndrome de dificultad respiratoria aguda relacionado con sepsis, entre muchos otros. Sin embargo, dada la creciente evidencia de que el CAP no mejoraba los resultados de los pacientes, las guías de práctica clínica del Colegio Americano de Cardiología y la Sociedad de Medicina de Cuidados Críticos recomienda en la actualidad *no* utilizar de rutina el CAP. Los autores estudiaron los datos de reclamaciones de pacientes hospitalizados de los Centros de Servicios de Medicare y Medicaid de 1999 a 2013 y estimaron la tasa de uso de CAP por cada 1.000 ingresos, la mortalidad a 30 días y la duración de la estancia. Descubrieron una reducción relativa estadísticamente significativa del 67,8% en el uso de CAP (de 6,28 por 1.000 ingresos en 1999 a 2,02 por 1.000 ingresos en 2013), además de reducciones interanuales en la mortalidad hospitalaria, la mortalidad a 30 días y la duración de la estancia. Sin embargo, los hallazgos también mostraron que dichas tasas variaban sustancialmente por sexo (fig. 17.4), raza (fig. 17.5) y edad (fig. 17.6). Estos resultados mostraron los beneficios adicionales en la restricción del uso de CAP en algunos pacientes. Mientras tanto, los autores admitieron las limitaciones en el uso de conjuntos de datos administrativos y la incapacidad de generalizar a individuos más jóvenes y sin seguro.

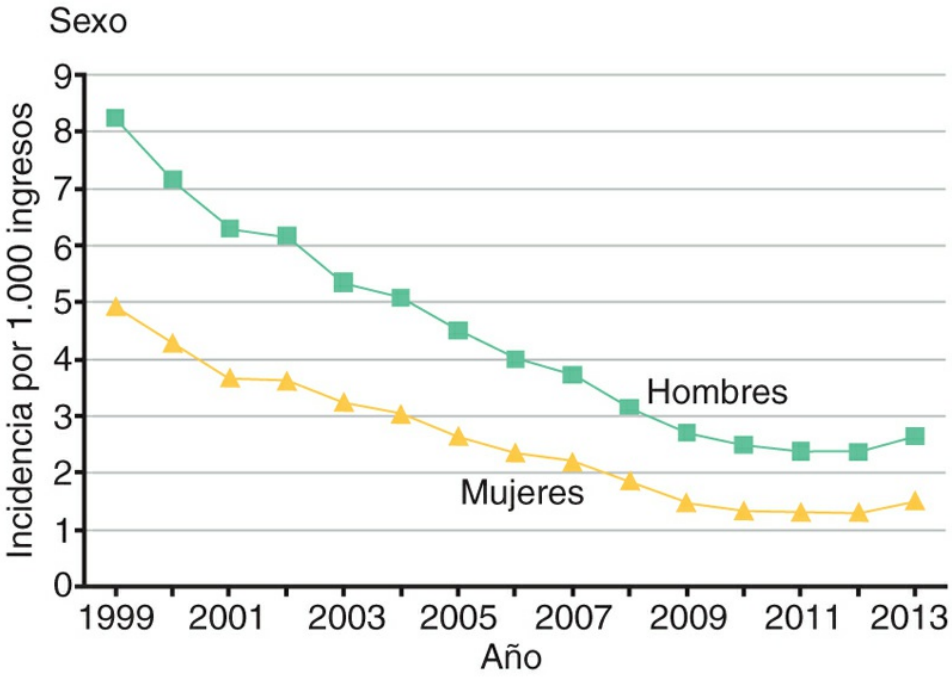


FIG. 17.4 Tasa de uso de catéter en la arteria pulmonar por 1.000 ingresos por género entre 1999 y 2013. (Modificado de Ikuta K, Wang Y, Robinson A, et al. National trends in use and outcomes of pulmonary artery catheters among medicare beneficiaries, 1999–2013. *JAMA Cardiol.* 2017;2:908–913.)

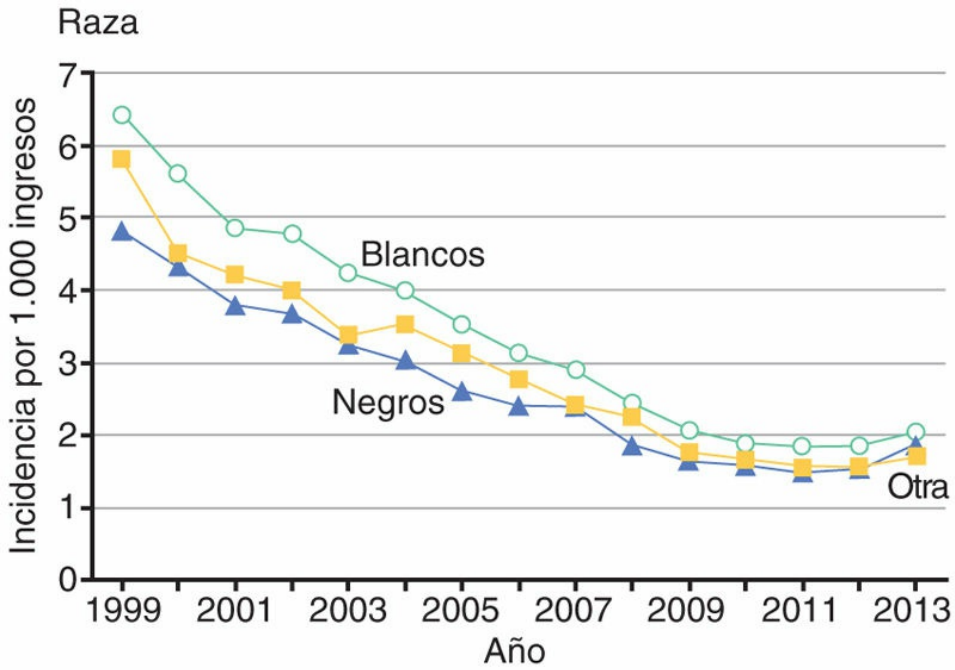


FIG. 17.5 Tasa de uso de catéter en la arteria pulmonar por 1.000 ingresos por raza entre 1999 y 2013. (Modificado de Ikuta K, Wang Y, Robinson A, et al. National trends in use and outcomes of pulmonary artery catheters among medicare beneficiaries, 1999–2013. *JAMA Cardiol.* 2017;2:908–913.)

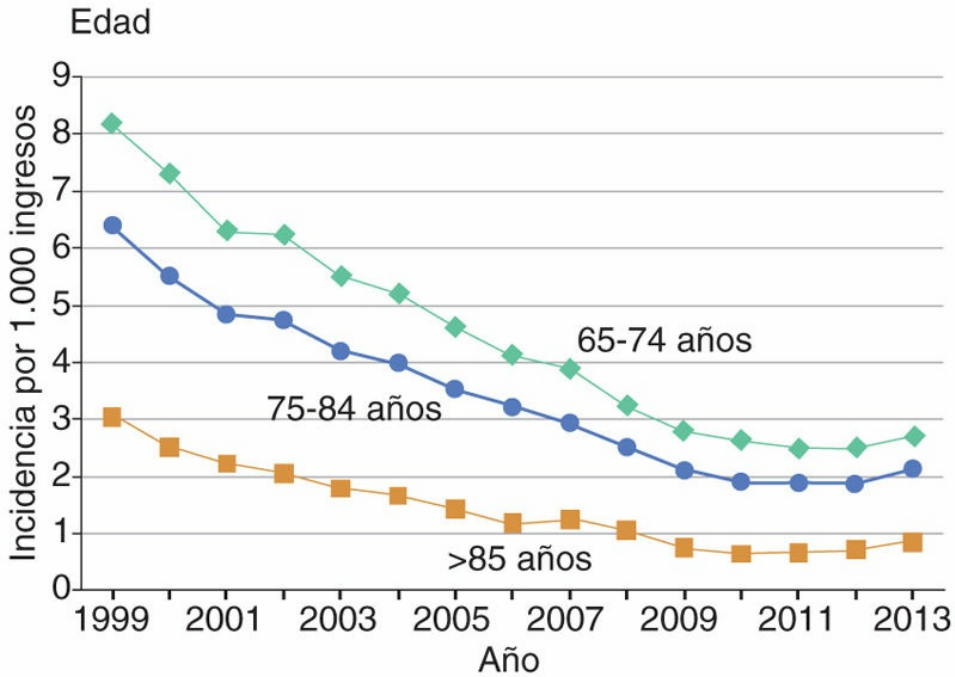


FIG. 17.6 Tasa de uso de catéter en la arteria pulmonar por 1.000 ingresos por grupos de edad entre 1999 y 2013. (Modificado de Ikuta K, Wang Y, Robinson A, et al. National trends in use and outcomes of pulmonary artery catheters among medicare beneficiaries, 1999–2013. *JAMA Cardiol.* 2017;2:908–913.)

Sesgos potenciales en la evaluación de los servicios de salud cuando se usan datos de grupo

Los estudios que evalúan los servicios de salud a partir de datos de grupo son susceptibles de presentar muchos de los sesgos que caracterizan a los estudios etiológicos, como se comentó en el capítulo 15. Además, ciertos sesgos son especialmente importantes para áreas y temas de investigación específicos y pueden ser importantes dependiendo del diseño epidemiológico específico seleccionado. Por ejemplo, los estudios sobre la relación de la asistencia prenatal con los resultados del parto son susceptibles de presentar varios sesgos potenciales importantes. En estos estudios, la pregunta que a menudo se plantea es si la atención prenatal, medida por el número absoluto de visitas prenatales, reduce el riesgo de prematuridad y de bajo peso al nacer. En este tipo de análisis se pueden introducir varios sesgos. Por ejemplo, a igualdad de otras condiciones, una mujer que tenga un parto prematuro tendrá menos visitas prenatales (es decir, al ser el embarazo más corto, habrá tenido menos tiempo en el que haya estado «en situación de riesgo» de realizar consultas prenatales). El resultado sería una relación espuria entre el menor número de visitas prenatales y la prematuridad, debido únicamente a que la gestación fue más corta. Sin embargo, el sesgo también puede actuar en la otra dirección. Una mujer que inicia la asistencia prenatal en el último trimestre del embarazo probablemente no tendrá un parto muy prematuro, ya que su embarazo ya está en el último trimestre. Esto daría lugar a la observación de una asociación entre menos visitas prenatales y un riesgo menor de partos muy prematuros. Además, las mujeres que han tenido complicaciones médicas o un mal resultado del embarazo en una gestación anterior pueden tener tal grado de ansiedad que acudan a más consultas prenatales (donde los problemas fetales pueden detectarse antes) y también pueden tener un riesgo mayor de un mal resultado. Así pues, los posibles sesgos pueden actuar en una o ambas direcciones. Si estas mujeres presentan un riesgo elevado que no sea susceptible de prevención, puede que se observe una asociación aparente entre un mayor número de consultas prenatales y un resultado adverso.

Por último, los estudios sobre resultados prenatales basados en la asistencia prenatal suelen estar sesgados por la autoselección, es decir, las mujeres que deciden comenzar la asistencia prenatal al principio del embarazo suelen tener un nivel educativo y socioeconómico más alto, así como una actitud más positiva hacia la asistencia sanitaria. Por tanto, una población de mujeres, que ya comienza con un menor riesgo de resultados adversos del parto, se selecciona a sí misma para una asistencia prenatal precoz. El resultado es la posibilidad de observar una asociación aparente entre el cuidado prenatal precoz y un menor riesgo de resultado adverso del embarazo, aunque la propia asistencia carezca de un verdadero beneficio para la salud.

Dos índices utilizados en los estudios ecológicos sobre los servicios de salud

Un índice que se utiliza en la evaluación de los servicios de salud mediante estudios ecológicos es la mortalidad evitable. Los análisis de *mortalidad evitable* asumen que la tasa de «muertes evitables» debería variar inversamente con la disponibilidad, accesibilidad y calidad de la asistencia médica en las diferentes regiones geográficas. La UK Office for National Statistics define la mortalidad evitable como:

Las muertes evitables son todas aquellas cuyas causas son prevenibles, tratables o ambas; y donde cada muerte se cuenta una sola vez. Cuando una causa de muerte se encuentra dentro de la definición prevenible y tratable, todas las muertes por esa causa se cuentan en ambas categorías cuando se presentan por separado[5].

Las patologías incluyen tuberculosis, hepatitis C, virus de inmunodeficiencia humana/síndrome de inmunodeficiencia adquirida (VIH/SIDA), neoplasias malignas seleccionadas, trastornos por consumo de drogas ilegales, enfermedades cardiovasculares y respiratorias, lesiones involuntarias e intencionales, entre otras.

En condiciones ideales, la mortalidad evitable serviría como una medida de la accesibilidad, la idoneidad y la eficacia de la asistencia en un área. Las muertes por VIH/SIDA serán menos frecuentes en las comunidades que cuenten con programas de orientación y realicen pruebas de detección de VIH adecuadas en un entorno amistoso y práctico, y cuenten con organizaciones de servicios relacionados con el SIDA de alta calidad, que a menudo se encuentran en áreas urbanas. En las zonas rurales, estos servicios pueden ser menos accesibles y los diagnósticos solo se pueden hacer cuando un paciente se presenta con una enfermedad definitoria del SIDA. Por lo tanto, es más probable que los pacientes tengan una tasa de mortalidad más alta en áreas con una cobertura de servicios más pobre, mortalidad que no habrían sufrido si hubieran vivido en un entorno urbano. Los cambios a lo largo del tiempo podrían representarse en una gráfica, lo que permitiría realizar comparaciones con otras áreas. Por desgracia, a menudo se carece de los datos necesarios para dicho análisis para muchas de las afecciones sugeridas por los análisis de mortalidad evitable. Además, puede que no se disponga de datos sobre los factores de confusión, por lo que las inferencias resultantes pueden ser cuestionables.

Un segundo enfoque es el uso de los *indicadores de salud*. En este caso, se asume que ciertas afecciones centinela reflejan el nivel general de asistencia sanitaria, y los cambios de la incidencia de estas afecciones a lo largo del tiempo se representan gráficamente y se comparan con los datos de otras poblaciones. Los cambios y las diferencias que se encuentran se relacionan

con los cambios en el sector de servicios de salud y se utilizan para derivar inferencias sobre la causalidad. Sin embargo, es difícil saber qué criterios deben cumplirse para que una condición determinada sea aceptable como un indicador de salud válido. Para permitir la identificación e implementación de un indicador de salud válido se debe seguir un proceso sistemático. Cada indicador debe tener los siguientes atributos: ser válido, confiable, relevante, realista, medible y bien conocido, se debe poder utilizar en la evaluación continua y debe poder medir de modo eficaz el éxito y el fracaso. La primera fase de desarrollo de un indicador consiste generalmente en la identificación de una lista de indicadores propuesta por un grupo de expertos en el área, seguida de una preselección de la lista de indicadores que cumplen con la mayoría o todos los atributos descritos anteriormente. La segunda fase incluye pruebas piloto, cuyo objetivo principal es probar la disponibilidad de los datos y estimar el tiempo, el esfuerzo y el coste para recopilar información de este indicador. La tercera fase de desarrollo es la prueba completa de los indicadores a mayor escala y el ajuste del indicador basado en los comentarios de los profesionales sanitarios sobre el uso de estos indicadores. La cuarta y última fase es la implementación completa de los indicadores maduros. En esta etapa, debe haber un mandato para notificar los indicadores y tener sistemas establecidos para la recopilación, tabulación, análisis e interpretación de datos, junto con un mecanismo de retroalimentación a los niveles intermedio y periférico del sistema de atención de salud.

El CDC mantiene 26 indicadores de salud principales (LHI, *Leading Health Indicators*) bajo 12 temas. Los LHI del programa Healthy People 2020 se muestran en el cuadro 17.3 y se puede acceder a ellos en su página web (healthypeople.gov). Como se mencionó anteriormente, las muertes evitables son todas aquellas definidas como debidas a causas prevenibles, tratables o ambas, donde cada muerte se cuenta una sola vez. Cuando una causa de muerte se encuentra dentro de la definición prevenible y tratable, todas las muertes por esa causa se cuentan en ambas categorías cuando se presentan por separado.

Cuadro 17.3 Los indicadores de salud principales de la iniciativa Healthy People 2020 están compuestos por 26 indicadores organizados bajo 12 temas

Acceso a servicios de salud

- Personas con seguro médico (AHS-1.1).
- Personas con un proveedor de atención primaria habitual (AHS-3).

Servicios preventivos clínicos

- Adultos que reciben pruebas de cribado de cáncer colorrectal según las

pautas más recientes (C-16).
- Adultos con hipertensión cuya presión arterial está controlada (HDS-12).
- Personas con diabetes diagnosticada cuyo valor de A1c es superior al 9% (D-5.1).
- Niños que reciben las dosis recomendadas de vacunas DTaP, polio, SRP, Hib, HepB, varicela y VNC entre los 19 y los 35 meses (IID-8).

Calidad del medio ambiente

- Índice de calidad del aire >100 (EH-1).
- Niños fumadores pasivos (TU-11.1).

Lesiones y violencia

- Muertes por lesiones (IVP-1.1).
- Homicidios (IVP-29).

Salud materna, del lactante y pediátrica

- Todas las muertes infantiles (MICH-1.3).
- Total de nacimientos prematuros vivos (MICH-9.1).

Salud mental

- Suicidio (MHMD-1).
- Adolescentes con un episodio depresivo mayor en los últimos 12 meses (MHMD-4.1).

Nutrición, actividad física y obesidad

- Adultos que cumplen con la actividad física aeróbica y los objetivos de fortalecimiento muscular (PA-2.4).
- Obesidad entre adultos (NWS-9).
- Obesidad entre niños y adolescentes (NWS-10.4).
- Ingesta diaria promedio de vegetales totales (NWS-15.1).

Salud bucal

- Niños, adolescentes y adultos que acudieron al dentista en el último año (OH-7).

Salud reproductiva y sexual

- Mujeres sexualmente activas que reciben servicios de salud reproductiva (FP-7.1).
- Conocimiento del estado serológico entre personas VIH positivas (VIH-13).

Determinantes sociales

- Estudiantes que se gradúan de la escuela secundaria 4 años después de comenzar el noveno grado (AH-5.1).

Abuso de sustancias

- Adolescentes que han consumido alcohol o drogas ilícitas en los últimos 30 días (SA-13.1).
- Consumo excesivo de alcohol en el último mes: adultos (SA-14.3).

Tabaco

- Tabaquismo en adultos (TU-1.1).
- Tabaquismo en adolescentes en los últimos 30 días (TU-2.2).

Evaluación utilizando datos individuales

Debido a las limitaciones inherentes al análisis de estudios que utilizan datos agrupados (es decir, estudios en los que no se tienen datos sobre la asistencia sanitaria [exposición] ni sobre los resultados de salud particulares para cada individuo), suelen ser preferibles los estudios que utilizan datos individuales. Si se desea comparar dos poblaciones, una que recibe la atención que se está evaluando (quizás un nuevo tratamiento) y otra que no la recibe (pacientes que siguen el «tratamiento habitual»), hay que hacer las dos preguntas siguientes para poder derivar inferencias acerca de la efectividad de la asistencia:

1. ¿Son las características de los dos grupos comparables desde los puntos de vista demográfico y médico y en cuanto a los factores relacionados con el pronóstico?
2. ¿Son los métodos de medición comparables (p. ej., los métodos de diagnóstico y la forma en la que se clasifica la enfermedad) en ambos grupos?

Ambas consideraciones se han comentado en los capítulos anteriores, ya que también se aplican igualmente bien a cuestiones de etiología, prevención y tratamiento, por lo que deben tenerse en cuenta en cualquier tipo de diseño de estudio.

Una cuestión importante respecto al uso de la epidemiología para estudiar los resultados de la evaluación de los servicios de salud es la necesidad de abordar la estratificación pronóstica. Si se observa un cambio en los resultados de salud después de proporcionar un cierto tipo de asistencia, ¿se puede concluir necesariamente que el cambio se debe a la (nueva) asistencia sanitaria prestada, o podría deberse a diferencias del pronóstico basadas en enfermedades concurrentes (enfermedades preexistentes que pueden o no estar relacionadas específicamente con la enfermedad que se está estudiando), en la gravedad o en otras afecciones asociadas que influyen en el pronóstico? Para abordar estas cuestiones, los estudios médicos de resultados deben llevar a cabo una estratificación pronóstica mediante el estudio de la casuística y la caracterización cuidadosa de los individuos estudiados en función de la gravedad de la enfermedad.

A continuación, se comentarán algunos *diseños de estudio* utilizados en la evaluación de los servicios de salud.

Diseños aleatorizados

La aleatorización elimina el problema del sesgo de selección secundario a la autoselección por parte del paciente o a la selección del paciente por parte del

profesional sanitario. Por lo general, los participantes en el estudio son asignados para recibir un tipo de asistencia frente a otra en lugar de para recibir atención frente a no recibirla (fig. 17.7). Por muchas razones, tanto éticas como prácticas, la asignación aleatoria de los pacientes para no recibir ninguna asistencia no suele ser una alternativa que se tenga en cuenta.



FIG. 17.7 Diseño de un estudio aleatorizado que compara las asistencias A y B.

Examinemos un estudio en el que se utilizó un diseño aleatorizado para evaluar las distintas estrategias de la asistencia sanitaria a pacientes de edad avanzada que han sufrido un ictus. En los pacientes con accidente cerebrovascular se ha recomendado encarecidamente un tratamiento precoz, organizado y basado en la asistencia hospitalaria. Sin embargo, se dispone de pocos datos de estudios controlados y con una metodología adecuada para comparar la asistencia hospitalaria con la asistencia especializada a domicilio (asistencia domiciliaria). Por otra parte, una alternativa a las unidades de ictus hospitalarias es un equipo de ictus especializado que pueda proporcionar asistencia en cualquier lugar del hospital donde puedan tratarse pacientes con ictus. Esta consideración tiene importancia práctica, ya que puede que no todos los hospitales tengan la posibilidad de ofrecer asistencia en una unidad especializada para todos los pacientes que sufren un accidente cerebrovascular debido a las limitaciones de espacio y otros problemas administrativos y económicos, de ahí la formación de equipos de ictus «itinerantes».

Con el fin de identificar la estructura organizativa óptima para la asistencia de pacientes con accidente cerebrovascular, Kalra y cols.[7] realizaron un ensayo controlado y aleatorizado en el que se comparó la eficacia de tres formas de asistencia (fig. 17.8). Los pacientes fueron asignados aleatoriamente

a uno de los siguientes grupos: 1) asistencia prestada en una unidad de ictus hospitalaria por un médico especialista en ictus y un equipo multidisciplinar, 2) asistencia médica prestada por un equipo multidisciplinar de ictus con experiencia en el tratamiento de accidentes cerebrovasculares, o 3) asistencia domiciliaria proporcionada por un equipo de especialistas. El resultado fue la mortalidad o el ingreso en una institución, y se evaluó a los 3, 6 y 12 meses después del inicio de un accidente cerebrovascular. Los datos fueron analizados por intención de tratar. En cada uno de los tres puntos temporales, los pacientes tratados en la unidad de ictus hospitalaria tenían menos probabilidades de fallecer o de ser ingresados en una institución que los del grupo tratado por el equipo de ictus o los del grupo que recibió asistencia domiciliaria. La supervivencia acumulada en los tres grupos se muestra en la figura 17.9. El estudio respalda el uso de unidades de ictus especializadas para la asistencia de pacientes con accidente cerebrovascular.

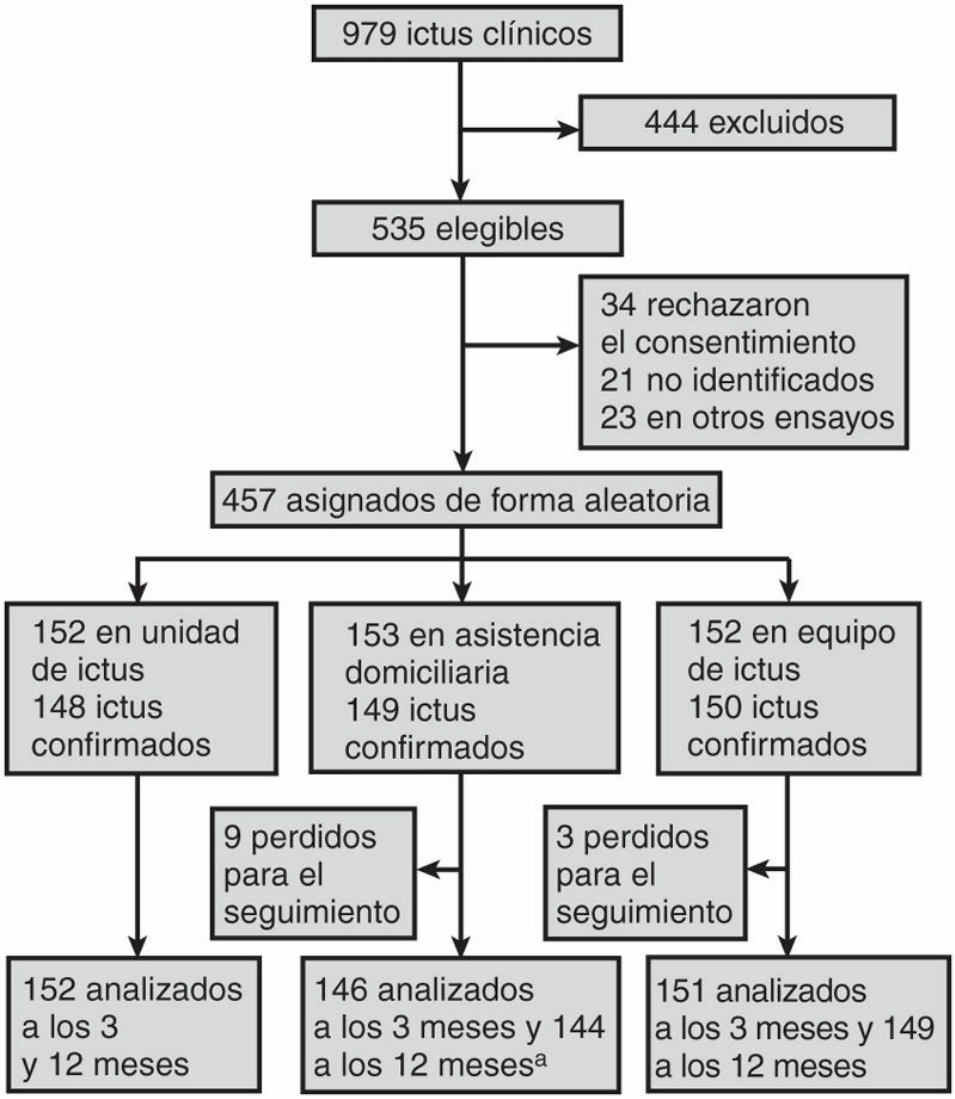
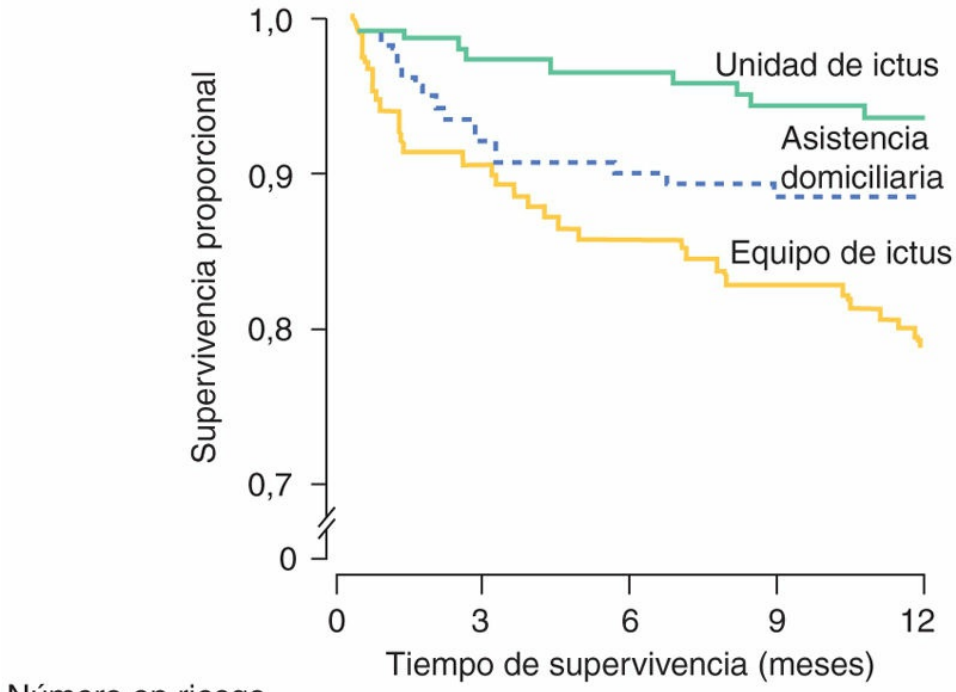


FIG. 17.8 Perfil de un ensayo aleatorizado de las estrategias para la asistencia del ictus. [a]Cincuenta y un pacientes de este grupo fueron ingresados en el hospital en las 2 semanas siguientes a la asignación aleatoria, pero se incluyen en

el análisis por intención de tratar. (Modificada de Kalra L, Evans A, Perez I, et al. Alternative strategies for stroke care: A prospective randomized controlled trial. *Lancet.* 2000;356:894–899.)



Número en riesgo					
Unidad de ictus	152	146	142	140	139
Asistencia domiciliaria	153	132	125	124	123
Equipo de ictus	152	133	124	122	115

FIG. 17.9 Curvas de supervivencia de Kaplan-Meier para diferentes estrategias de asistencia después de un ictus agudo. (De Kalra L, Evans A, Perez I, et al: Alternative strategies for stroke care: A prospective randomized controlled trial. *Lancet.* 2000;356:894–899.)

Como se observa en la figura 17.9, un resultado interesante y algo sorprendente en este estudio es que la supervivencia fue mejor en los pacientes que fueron asignados de forma aleatoria para recibir asistencia domiciliaria que en los que recibieron asistencia en el hospital por un equipo de ictus.

Una posible explicación de esta observación es que los pacientes del grupo de asistencia domiciliaria cuya afección empeoró o que habían desarrollado nuevos problemas fueron retirados de la asistencia domiciliaria e ingresaron en una unidad de ictus. Estos pacientes aún se analizaron con el grupo de asistencia domiciliaria, porque se utilizó un análisis por intención de tratar, que analiza el resultado de acuerdo con la asignación aleatoria inicial. Estos pacientes pueden haber recibido asistencia en la unidad de ictus, y, si es así, sus resultados tenderían a mejorar los resultados del grupo de asistencia domiciliaria debido al análisis por intención de tratar.

Drummond y cols.[8] llevaron a cabo un seguimiento de 10 años de un ensayo clínico controlado y aleatorizado sobre la asistencia en una unidad de rehabilitación de ictus. Estos autores observaron que el tratamiento en una unidad de rehabilitación de ictus proporcionaba beneficios de supervivencia incluso 10 años después del accidente cerebrovascular. Las razones exactas no están claras, pero los autores sugieren que una explicación podría ser que la supervivencia a largo plazo se relaciona con la reducción precoz de la discapacidad.

Diseños no aleatorizados

No es posible evaluar muchas intervenciones de asistencia sanitaria mediante ensayos aleatorizados por varias razones. En primer lugar, tales ensayos suelen ser complejos desde el punto de vista logístico y muy caros. Debido a que en cualquier momento dado se están usando tantas medidas diferentes de la asistencia sanitaria, no es posible someter a todas a una evaluación aleatorizada. En segundo lugar, pueden plantearse problemas éticos en los estudios de evaluación de servicios sanitarios. En concreto, la asignación aleatoria puede considerarse un proceso inaceptable tanto por muchos pacientes como por los encargados de su atención médica. En tercer lugar, los ensayos aleatorizados a menudo requieren mucho tiempo para completarse y, dado que los programas de atención sanitaria y los problemas de salud cambian con el tiempo, cuando se obtienen finalmente los resultados del estudio y se analizan, es posible que ya no sean totalmente relevantes. Por estas razones, muchos investigadores de la asistencia sanitaria están buscando estrategias alternativas que puedan proporcionar al menos algo de información. Una de estas estrategias que se comentó anteriormente *(la investigación de resultados)* suele referirse a la utilización de los datos de estudios no aleatorizados que a menudo utilizan conjuntos extensos de datos existentes (los denominados «macrodatos»).

Diseño antes-después (controles históricos)

Si la aleatorización no es posible o no se va a utilizar por algún motivo, un posible diseño de estudio para evaluar un programa consiste en comparar a las personas que recibieron asistencia antes de que se estableciese un programa (o antes de que se dispusiese de la medida de la asistencia sanitaria) con las que recibieron asistencia del programa después de su creación (o después de que se dispusiese de la medida). ¿Cuáles son los problemas de este tipo de diseño *antes-después*? En primer lugar, los datos obtenidos en cada uno de los dos periodos no suelen ser comparables en términos de calidad o integridad. Cuando se desarrolla una nueva forma de prestación de servicios de salud, los evaluadores del programa pueden querer incluir a personas que fueron tratadas en el pasado, antes de que comenzara el programa, como grupo de comparación. Los datos sobre personas tratadas después de que se iniciase el programa pueden ser

recopilados mediante un instrumento de investigación bien diseñado, mientras que los datos de los pacientes previos pueden incluir solo a aquellos disponibles a partir de registros de asistencia sanitaria que habían sido diseñados y utilizados solo con fines clínicos o administrativos. Si encontramos una diferencia en el resultado, puede que no se sepa si la diferencia observada se debe al efecto del programa o a las diferencias de calidad de los datos de los dos periodos de tiempo.

En segundo lugar, si se observa una diferencia (p. ej., la mortalidad es más baja después de comenzar un programa que antes de su inicio), no sabemos si la diferencia se debe al programa en sí o a otros factores que pueden haber cambiado con el tiempo, como la vivienda, la nutrición, otros aspectos del estilo de vida o el uso de otros servicios de salud.

En tercer lugar, existe un problema de selección. A menudo, es difícil saber si la población estudiada después de instaurar un programa en realidad es similar a la observada antes de implantar el programa en lo que respecta a otros factores que podrían afectar a los resultados.

¿Significa esto que los estudios *antes-después* carecen de utilidad? No, en absoluto. Pero sí quiere decir que estos estudios solo proporcionan una sugerencia (y raramente son concluyentes) a la hora de demostrar la eficacia de un nuevo servicio de salud.

Un diseño *antes-después* se utilizó en un estudio para evaluar el impacto del sistema de pago prospectivo (SPP) de Medicare en Estados Unidos sobre la calidad de la asistencia[9]. El estudio fue impulsado por la preocupación de que el SPP, con su estricta regulación de la duración de la estancia hospitalaria y los incentivos para la reducción de costes, podría haber afectado negativamente a la calidad asistencial. Se seleccionó el diseño *antes-después* debido a que el SPP fue instituido a nivel nacional, por lo que no se podía utilizar un diseño de cohortes prospectivo. Los datos de casi 17.000 pacientes de Medicare que fueron hospitalizados en 1981-82 antes de que se implantase el SPP se compararon con los datos de los pacientes hospitalizados en 1985-86 después de que se instaurase el SPP. Se evaluó la calidad asistencial para cinco enfermedades: 1) insuficiencia cardiaca congestiva, 2) infarto de miocardio, 3) neumonía, 4) accidentes cerebrovasculares y 5) fracturas de cadera. Los resultados se ajustaron por el nivel de enfermedad del paciente al ingreso en el hospital. Aunque no se observó que el SPP se asociase con un aumento de la mortalidad a 30 días o a los 6 meses, se apreció un aumento de la inestabilidad al alta (definida como la presencia de afecciones en el momento del alta que los médicos consideraban que debían corregirse antes del alta o monitorizarse después de ella, y que podrían dar lugar a malos resultados si no se corregían)[10]. Los autores señalaron que hay otros factores que también pueden haber cambiado durante el periodo anterior y posterior a la institución del SPP. Aunque el diseño *antes-después* era probablemente el único posible para el tema abordado en este estudio, este es susceptible de presentar algunos de los problemas de este tipo de diseño que se comentaron anteriormente.

Cuando el cambio en el riesgo del resultado es espectacular, el diseño *antes-después* es similar al llamado *experimento natural* (v. cap. 14, sección titulada «Enfoques de la etiología en las poblaciones humanas»). Por ejemplo, sería difícil explicar la marcada disminución en las tasas de hospitalización por diabetes y meningitis por razones distintas a la introducción de la insulina y la estreptomicina, respectivamente.

Diseño no aleatorizado simultáneo (programa-ausencia de programa)

Una opción para evitar los problemas de los cambios que se producen con el tiempo es llevar a cabo una comparación simultánea de dos poblaciones sin asignación aleatoria en la que una población es atendida por el programa y la otra no. Este tipo de diseño es, en efecto, un estudio de cohortes en el que el tipo de asistencia sanitaria estudiada representa la «exposición». Al igual que en cualquier estudio de cohortes, se plantea el problema de cómo seleccionar los grupos expuesto y no expuesto para su estudio.

En los últimos años ha surgido un gran interés por saber si el mayor volumen de pacientes atendidos en un hospital y por un cirujano se relaciona con mejores resultados en los pacientes y mejores costes, habiéndose realizado muchos estudios sobre estos temas. Un ejemplo de un estudio no aleatorizado simultáneo sobre el volumen de pacientes de un hospital es el publicado por Wallenstein y cols.[11]. En este estudio se evaluó si las diferencias en cuanto a los resultados de los pacientes en distintos hospitales se relacionaban con el volumen de procedimientos hospitalarios realizados. Los autores estudiaron las hospitalizaciones de pacientes sometidas a histerectomía laparoscópica, el procedimiento ginecológico mayor más común (600.000 cirugías por año) en Estados Unidos. Examinaron la relación de las complicaciones intrahospitalarias (intraoperatorias, del sitio quirúrgico y médicas), así como la duración de la estancia y el costo durante la hospitalización índice con el volumen de cirugías realizadas por los médicos y en general en el hospital[11]. Como se observa en la tabla 17.1, se encontró una relación dosis-respuesta: las complicaciones hospitalarias más altas, la duración del ingreso más prolongada y los mayores costes ocurrieron en los hospitales que tenían el volumen más bajo de histerectomías por año. El hallazgo de que los hospitales que realizan más histerectomías tienen menor duración de los ingresos y menores costes tiene importantes implicaciones políticas potenciales y es un argumento a favor de la regionalización de los servicios quirúrgicos ginecológicos.

Tabla 17.1

Asociación entre el volumen hospitalario de histerectomías laparoscópicas realizadas por año y la morbilidad, la mortalidad y la utilización de recursos

	NÚMERO DE INTERVENCIONES		
	<46,4/año	46,4-105/año	>105/año
Cualquier complicación (%)	5,8	5,0	4,7
Complicaciones intraoperatorias (%)	2,4	2,2	2,1
Complicaciones en el sitio quirúrgico (%)	2,60	2,3	1,8
Complicaciones médicas (%)	1,4	1,1	1,2
Ingreso superior a 2 días (%)	10,0	7,8	5,3
Coste (dólares)	6.527,00 $	5.809,00 $	5.561,00 $
Muertes (%)	0,02	0,01	0,01

Modificada de Wallenstein ME, Ananth CV, Kim JH, et al. Effect of surgical volume on outcomes for laparoscopic hysterectomy for benign indications. *Obstet Gynecol*. 2012;119:709–716.

Es posible que las conclusiones que relacionan unos volúmenes hospitalarios mayores con unos resultados mejores en los pacientes puedan deberse a la realización de un mayor volumen de procedimientos por los cirujanos de estos hospitales en lugar de a los volúmenes totales de los procedimientos realizados en ellos. Birkmeyer y cols. analizaron esta cuestión[12]; utilizaron los datos de reclamaciones a Medicare de 1998 y 1999 para analizar la mortalidad entre los 474.108 pacientes que fueron sometidos a uno de cuatro procedimientos cardiovasculares o de cuatro procedimientos de resección oncológica (fig. 17.10). Observaron que, para la mayoría de los procedimientos, la mortalidad era mayor en los pacientes operados por cirujanos que atendían un bajo volumen de casos que en los operados por cirujanos con un alto volumen de casos. Esta relación se mantuvo con independencia del volumen quirúrgico del hospital en el que se realizó la cirugía.

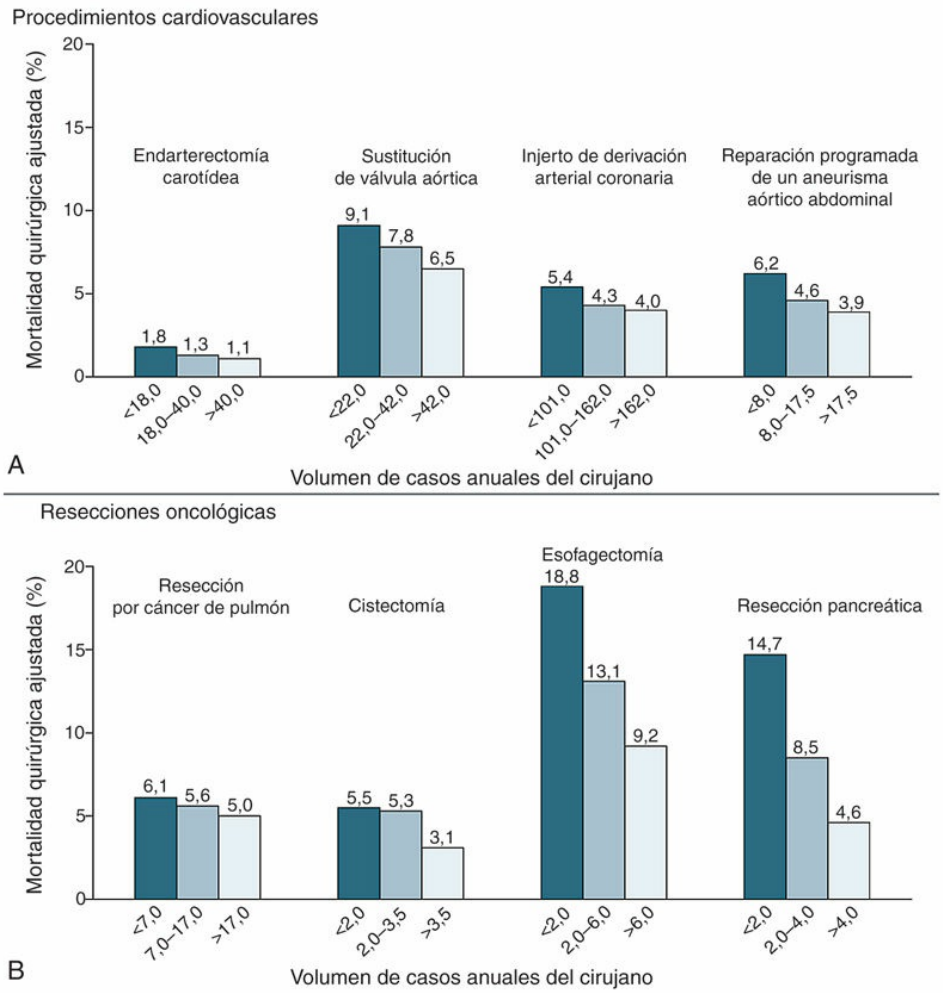


FIG. 17.10 Mortalidad quirúrgica ajustada entre los pacientes de Medicare en 1998 y 1999 según el volumen de casos del cirujano para cuatro procedimientos cardiovasculares (gráfica A) y cuatro procedimientos de resección oncológica (gráfica B). La mortalidad quirúrgica se definió como la mortalidad antes del alta hospitalaria o en los 30 días posteriores al procedimiento índice. El volumen de casos del cirujano se basa en el número total de procedimientos realizados. (De Birkmeyer JD, Stukel TA, Siewers AE, et al. Surgeon volume and operative mortality in the United States. *N Engl J Med.* 2003;349:2117–2127.)

Comparación entre usuarios y no usuarios

Una estrategia para realizar un estudio no aleatorizado simultáneo es comparar un grupo de personas que utilizan un servicio de salud con un grupo de personas que no lo hacen (fig. 17.11).

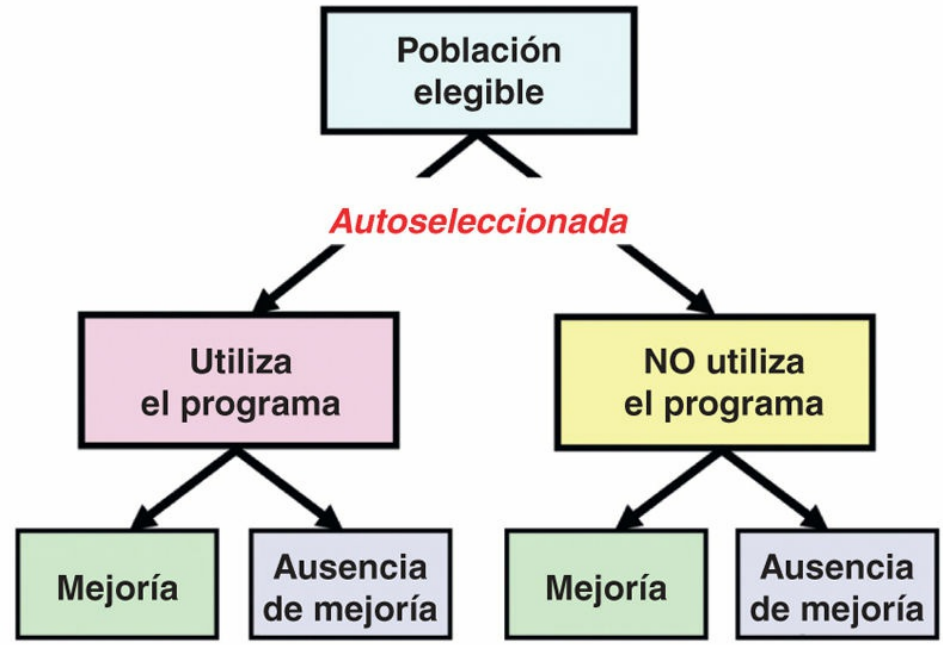


FIG. 17.11 Diseño de un estudio de cohortes no aleatorizado que compara a los usuarios con los no usuarios de un programa.

Los problemas de autoselección inherentes a este tipo de diseño se conocen desde hace mucho tiempo. Haruyama y cols. estudiaron la asociación entre la utilización personal de los chequeos de salud generales (CSG) y los gastos médicos (GM) en una población japonesa de mediana edad (tabla 17.2)[13].

Tabla 17.2

Cociente de posibilidades (e intervalos de confianza del 95%) de cualquier consulta médica (definida como la consulta a un médico en un periodo de 1 año) de acuerdo con los subgrupos de utilización del chequeo de salud general en la población japonesa de mediana edad, 2010

	No usuarios	Baja frecuencia	Alta frecuencia
Ambulatoria	1,00	2,90 (2,61-3,22)	4,37 (3,88-4,92)
Hospitalaria	1,00	0,79 (0,71-0,88)	0,75 (0,67-0,83)

Modificada de Haruyama Y, Yamazaki T, Endo M, et al. Personal status of general health checkups and medical expenditure: a large-scale community-based retrospective cohort study. *J Epidemiol*. 2017;27(5):209–214.

En este estudio, los autores incluyeron a 33.417 residentes de la ciudad de Soka, prefectura de Saitama, Japón, y estudiaron la utilización de CSG de 2008 a 2010. La utilización de CSG se dividió en cero veces (no usuarios), de una a tres veces (usuarios de baja frecuencia), y de cuatro a seis veces (usuarios de alta frecuencia). En comparación con los no usuarios, los

usuarios de alta frecuencia mostraron GM ambulatorios más elevados y estadísticamente significativos. Además, los usuarios de baja y alta frecuencia mostraron menores GM por ingreso y GM totales, con carácter estadísticamente significativo, que los no utilizadores. Los autores concluyeron que los GM ambulatorios aumentaron con la frecuencia de la asistencia a los CSG, y que el diagnóstico temprano facilitado por la consulta ambulatoria temprana es más probable que conduzca a un ligero aumento en los GM ambulatorios, pero a una disminución en los GM hospitalarios por enfermedades graves, lo que resulta en una disminución en el coste total de la asistencia sanitaria.

Otro ejemplo de las diferencias entre las características de los grupos comparados lo proporciona un estudio realizado por Gierisch y cols. sobre la no adherencia al cribado del cáncer de mama con exámenes mamográficos periódicos[14]. En este estudio, las mujeres no cumplidoras con el cribado tenían más probabilidades que las mujeres cumplidoras de tener entre 40 y 49 años, y de tener una salud autoinformada como aceptable o deficiente, así como dificultades para hacerse mamografías. Como estas variables están relacionadas con el cáncer de mama (y la mortalidad por todas las causas), deben tenerse en cuenta al usar un diseño no aleatorizado para examinar la efectividad del cribado del cáncer de mama.

Aunque se puede tratar de resolver el problema de la selección mediante la caracterización del perfil pronóstico de quienes utilizan la asistencia y quienes no, siempre que los grupos no se hayan aleatorizado persistirá la incertidumbre de si no se identificaron algunos factores en el estudio que podrían haber diferenciado a los usuarios y a los no usuarios y, por tanto, haber afectado al resultado de salud.

Comparación de las poblaciones elegible y no elegible

Debido al problema de los posibles sesgos de selección a la hora de comparar los grupos de usuarios y no usuarios, otra estrategia es la de comparar a las personas que son elegibles para recibir la asistencia que se está evaluando con un grupo de personas que no son elegibles (fig. 17.12).

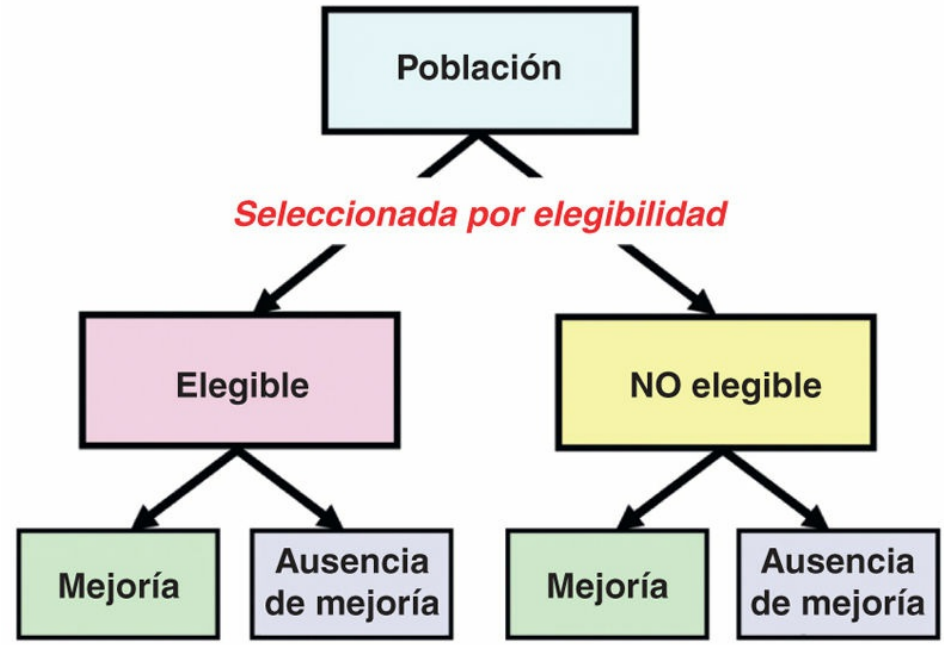


FIG. 17.12 Diseño de un estudio de cohortes no aleatorizado que compara a las personas elegibles con las no elegibles para un programa.

En este caso, se supone que la elegibilidad o no elegibilidad no se relaciona ni con el pronóstico ni con el resultado; por tanto, no se introduce un sesgo de selección que pueda afectar a las inferencias del estudio. Por ejemplo, los criterios de elegibilidad pueden incluir el tipo de empleador o la zona de residencia. Sin embargo, incluso con este diseño, se debe estar alerta ante los factores que pueden introducir sesgos de selección. Por ejemplo, es evidente que la zona de residencia puede relacionarse con el nivel socioeconómico. El problema de encontrar una población no elegible apropiada para la comparación puede ser crítico. Sin embargo, las personas no elegibles pueden seleccionarse de vecindarios similares, lo que podría compensar el interés por asegurar la comparabilidad del estatus socioeconómico. Además, como las diferencias entre individuos elegibles y no elegibles también pueden afectar la validez externa, en ocasiones el ajuste por las variables que difieren entre estos individuos mejora la validez externa.

Diseños combinados

En la figura 17.13 se muestra un resultado hipotético de un estudio no aleatorizado que comparó el nivel de morbilidad en un grupo que *no* ha recibido un servicio de salud (grupo X, en rojo) con el nivel de morbilidad en un grupo que *sí* ha recibido el servicio de salud (grupo Y, en negro). Debido a que el nivel de morbilidad observado es menor para el grupo Y que para el grupo X, podría surgir la tentación de concluir a partir de estos resultados que el servicio de salud reduce la morbilidad. Sin embargo, como se observa en la figura 17.13 *(izquierda)*, para llegar a esta conclusión se debe asumir que los niveles originales de morbilidad en ambos grupos eran comparables antes de que se proporcionase la asistencia al grupo Y. Si los niveles de morbilidad

de X_1 e Y_1 eran similares, el hallazgo de un menor nivel de morbilidad en el grupo Y (Y_2) que en el grupo X (X_2) después de haber administrado la asistencia podría interpretarse como un probable efecto de la asistencia recibida.

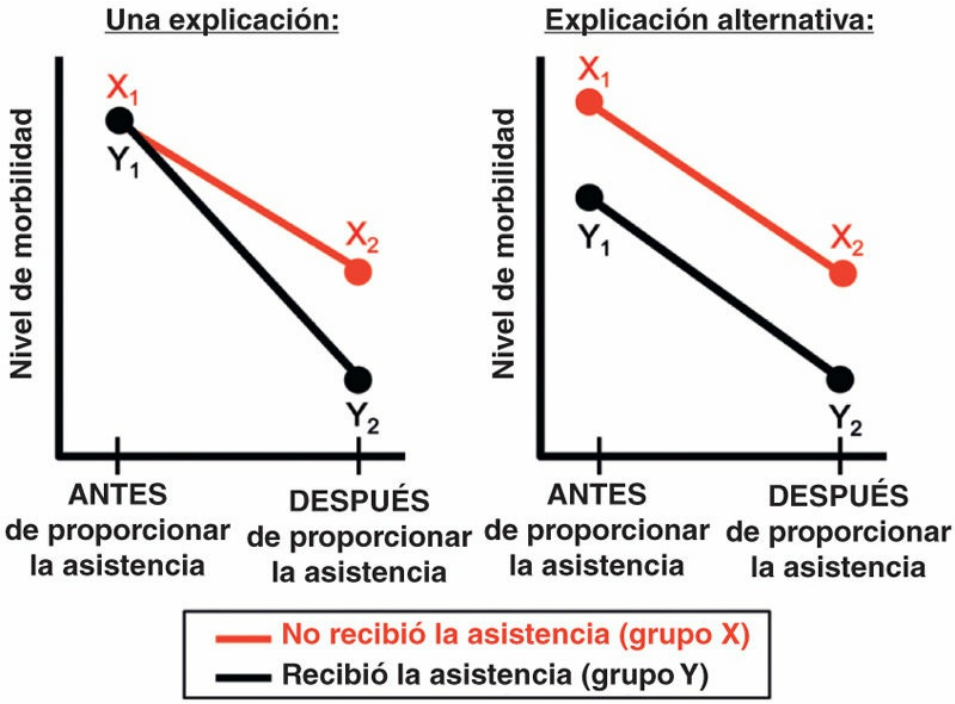


FIG. 17.13 Dos posibles explicaciones que se traducirían en la observación de una diferencia de morbilidad entre el grupo X y el grupo Y después de que el grupo Y (en *negro*) haya recibido un servicio de asistencia sanitaria.

Sin embargo, como se observa en la figura 17.13 *(derecha)*, puede que los grupos hubiesen sido diferentes en un principio y que sus pronósticos pudiesen ser distintos en ese momento, incluso antes de prestar cualquier tipo de asistencia. En tal caso, cualquier diferencia de la morbilidad observada después de la asistencia (es decir, $Y_2 < X_2$) tal vez solo refleje las diferencias originales en el momento previo a la administración de la asistencia, e informaría necesariamente sobre la efectividad de la asistencia proporcionada. Sin datos sobre los niveles de morbilidad en ambos grupos antes de la administración de la asistencia (situación «basal»), esta última explicación de las observaciones no se puede descartar.

En vista de este problema, otra estrategia para la evaluación del programa es utilizar un *diseño combinado*, que asocia un diseño *antes-después* con un diseño *programa-ausencia de programa*. Esta estrategia se demuestra en el siguiente ejemplo, en el que se evaluó la asistencia ambulatoria para los dolores faríngeos infantiles.

El estudio se diseñó para evaluar la eficacia de la asistencia ambulatoria para los dolores faríngeos infantiles mediante la determinación de si los niños que eran elegibles para esta asistencia experimentan tasas más bajas de

complicaciones de faringitis estreptocócica no tratada, como la glomerulonefritis (inflamación del riñón) o los trastornos neuropsiquiátricos pediátricos asociados con infecciones estreptocócicas (PANDAS, *pediatric neuropsychiatric disorders associated with streptococcal infection*), como tics, que los niños que no eran elegibles. El fundamento era el siguiente: las faringitis estreptocócicas son frecuentes en la infancia. Las faringitis estreptocócicas no tratadas pueden dar lugar a complicaciones como infecciones renales. Si las faringitis estreptocócicas reciben un tratamiento adecuado, las complicaciones se pueden prevenir. Por tanto, si estos programas son eficaces a la hora de tratar una faringitis estreptocócica, los niños que recibieron el tratamiento deberían tener menos complicaciones.

Se pueden identificar y comparar varios subgrupos de niños y adolescentes y comparar sus tasas de complicaciones por infecciones estreptocócicas no tratadas. Los grupos podrían incluir residentes de las zonas que cumplen los criterios de elegibilidad para la asistencia integral y residentes de zonas que no cumplen estos criterios de elegibilidad para la asistencia integral. Ambos podrían compararse con la ciudad o el pueblo en su conjunto.

Un ejemplo histórico de la investigación del Dr. Gordis[15] muestra otra complicación de la faringitis estreptocócica: la fiebre reumática, que era mucho más común en el siglo pasado que en la actualidad. En la figura 17.14 se muestra una comparación de tipo *programa-ausencia de programa* de las tasas de fiebre reumática en niños afroamericanos en la ciudad de Baltimore. En los niños elegibles para la asistencia integral, basada en la zona de residencia, la tasa de fiebre reumática era de 10,6 por 100.000, en comparación con 14,9 por 100.000 en los que no eran elegibles. Aunque la tasa era inferior en el grupo elegible en esta comparación simultánea, la diferencia no era llamativa.

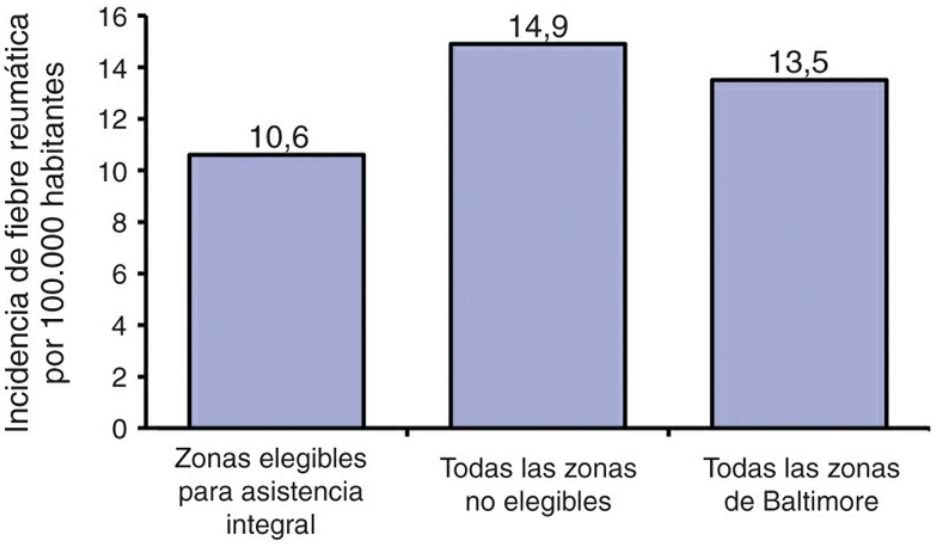


FIG. 17.14 Asistencia integral e incidencia de fiebre reumática por 100.000 habitantes, 1968-1970; Baltimore, población afroamericana de 5-14 años. (Modificada de Gordis L. Effectiveness of comprehensive-care programs in

preventing rheumatic fever. N Engl J Med 1973;289:331–335.)

El siguiente análisis de este diseño combinado examinó las variaciones de las tasas de fiebre reumática a lo largo del tiempo en las poblaciones elegible y no elegible.

Como se observa en la figura 17.15, la tasa de fiebre reumática se redujo un 60% en las zonas elegibles de 1960 a 1964 (antes de la implantación de los programas) a 1968-1970 (después de la implantación). En las zonas no elegibles, la incidencia de fiebre reumática prácticamente no varió (+2%). Por tanto, ambas partes del diseño combinado son concordantes con una disminución relacionada con la asistencia disponible.

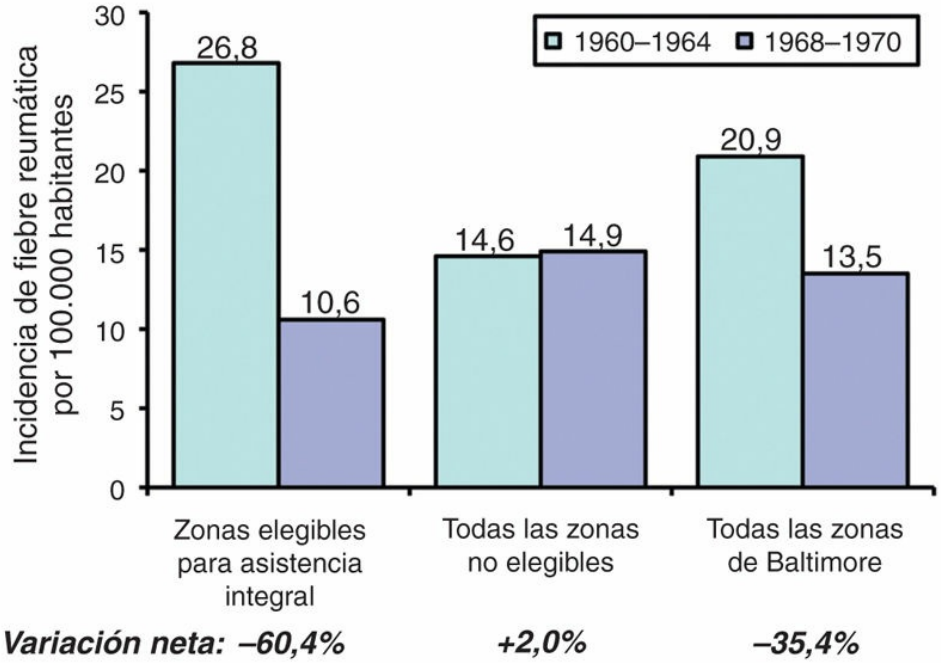


FIG. 17.15 Asistencia integral y variaciones de la incidencia de fiebre reumática por 100.000 habitantes, 1960-1964 y 1968-1970; Baltimore, población afroamericana de 5-14 años. (Modificada de Gordis L. Effectiveness of comprehensive-care programs in preventing rheumatic fever. *N Engl J Med.* 1973;289:331–335.)

Sin embargo, debido a que se habían producido muchos cambios en la ciudad de Baltimore durante este tiempo, no era seguro si la asistencia proporcionada por los programas era en realidad la responsable de la disminución de la fiebre reumática. Se llevó a cabo, entonces, otro análisis. En un niño, una infección faríngea por estreptococos puede ser sintomática o asintomática. Es evidente que solo un niño con dolor faríngeo sintomático habría consultado con el médico. Si se plantea la hipótesis de que la asistencia médica fue responsable de la reducción de la incidencia de fiebre reumática, sería de esperar que la disminución de la incidencia se limitase a los niños con dolor faríngeo sintomático que hubiesen solicitado asistencia, y que no se hubiese producido en niños asintomáticos que tenían infecciones sin

manifestaciones clínicas.

Como se observa en la figura 17.16, toda la disminución se limitaba a los niños que habían tenido una infección sintomática; no hubo modificaciones de la incidencia de fiebre reumática en los niños con faringitis estreptocócica asintomática. Por tanto, estos resultados son muy concordantes con la sugerencia de que fue la asistencia médica (o algún factor estrechamente asociado con ella) la responsable de la disminución de la incidencia de fiebre reumática.

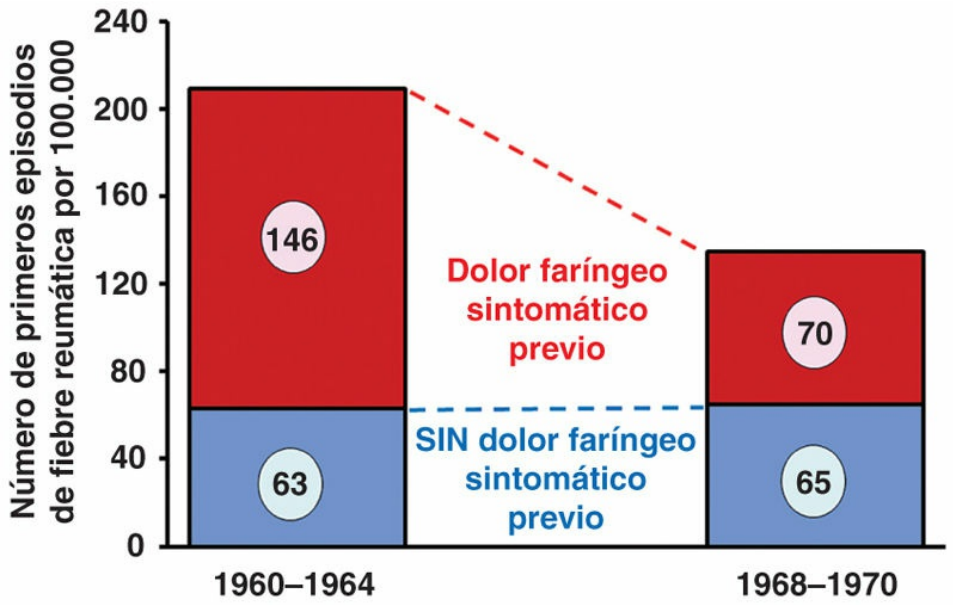


FIG. 17.16 Variaciones de la incidencia anual de los primeros episodios de fiebre reumática en relación con la presencia o ausencia de dolor faríngeo sintomático previo. Como se observa en la figura, toda la disminución de los primeros episodios de fiebre reumática se debió a una reducción de los primeros episodios de fiebre reumática precedidos de dolor faríngeo sintomático. (Modificada de Gordis L. Effectiveness of comprehensive-care programs in preventing rheumatic fever. *N Engl J Med.* 1973;289:331–335.)

Estudios de casos y controles

El uso del diseño de casos y controles para evaluar los servicios de salud, incluidas las vacunas y otras formas de programas de prevención y de cribado, ha suscitado un interés creciente en el campo de la salud pública. Aunque el diseño de casos y controles se ha aplicado sobre todo a los estudios etiológicos, cuando se pueden obtener los datos adecuados, este diseño puede servir como un sustituto útil, aunque limitado, de los ensayos aleatorizados. Sin embargo, debido a que este diseño requiere la definición y la especificación de los casos, es más aplicable a los estudios de prevención de enfermedades concretas. En tal caso, la «exposición» es la medida preventiva específica u otra medida de salud que se está evaluando. Al igual que en la mayoría de las investigaciones sobre servicios de salud, la estratificación por

gravedad de la enfermedad y por otros factores pronósticos posibles es esencial para una interpretación adecuada de los hallazgos. Los problemas metodológicos asociados a este tipo de estudios (se comentaron ampliamente en el cap. 7) también aparecen cuando se utiliza el diseño de casos y controles para evaluar la efectividad. En especial, estos estudios deben enfrentarse a la selección de los controles y a los problemas relacionados con los factores de confusión.

Conclusión

En este capítulo se ha revisado la aplicación de diseños básicos de estudios epidemiológicos para la evaluación de los servicios de salud. Muchas de las cuestiones que se plantean son similares a las que surgen en los estudios etiológicos, aunque a veces presentan un giro diferente. En los estudios etiológicos, el principal interés es la posible asociación de un factor causal potencial con una enfermedad específica, y factores como la accesibilidad a los servicios de salud a menudo son posibles factores de confusión que deben tenerse en cuenta. Por ejemplo, en el Multi-Ethnic Study of Atherosclerosis, la evaluación de los determinantes de la fibrilación auricular tuvo que tener en cuenta el posible efecto de confusión del estado del seguro de salud (un marcador del acceso a la atención médica), ya que el diagnóstico de esta patología a menudo se realiza durante una revisión médica[16].

En los estudios de evaluación de la asistencia sanitaria, el principal interés son las posibles asociaciones de una asistencia sanitaria o de una actividad preventiva con el resultado de una enfermedad concreta, y factores como las enfermedades preexistentes y otros factores pronósticos y de riesgo se convierten en posibles factores de confusión que deben tenerse en cuenta. Por tanto, aunque persisten muchos de los mismos problemas de diseño, la investigación sobre evaluación suele centrarse en temas diferentes de medición y evaluación. El ensayo aleatorizado sigue siendo el método óptimo para demostrar la efectividad de una intervención de salud. Sin embargo, las cuestiones éticas deben ser tenidas en cuenta, ya que puede ser poco ético retener un tratamiento conocido o efectivo en un diseño de ensayo aleatorizado. Al iniciar cualquier estudio de evaluación de la asistencia sanitaria, hay que preguntarse en primer lugar si, con los conocimientos actuales, es plausible desde los puntos de vista biológico y clínico esperar un beneficio específico de la asistencia que se está evaluando.

Por razones prácticas, las observaciones no aleatorizadas también son necesarias y deben aprovecharse para intentar ampliar los esfuerzos de evaluación de los servicios sanitarios. Los críticos de los ensayos aleatorizados han señalado que estos estudios han incluido (y solo pueden incluir) una pequeña fracción de todos los pacientes que reciben asistencia en el sistema sanitario, de modo que la generalización de los resultados es un posible problema. Aunque esto es cierto, la generalización es un problema en cualquier estudio, con independencia del tamaño de la población estudiada. Sin embargo, aunque se perfeccione más la metodología de los ensayos clínicos, también se requieren métodos optimizados para mejorar la información que se puede obtener de las evaluaciones no aleatorizadas de los servicios de salud.

El estudio de los componentes específicos de la asistencia, en lugar de un programa asistencial en sí mismo, es esencial. De esta manera, si se puede identificar un elemento efectivo en un conjunto de muchas modalidades, las

otras pueden eliminarse y la calidad de la asistencia se puede mejorar de una manera rentable.

En el capítulo 18, la exposición sobre la evaluación se extiende a un tipo específico de programa de servicios de salud: el cribado (detección precoz) de enfermedades en las poblaciones humanas.

Bibliografía

1. Frost WH. Rendering account in public health. *Am J Public Health*. 1925;15:394–397.
2. Chapin CV. Comments on "Rendering An Account on Public Health," by Frost. *Am J Public Health*. 1925;15:397–398.
3. Butler D. When Google got flu wrong. *Nature*. 2013;494(7436):155–156.
4. Ikuta K, Wang Y, Robinson A, et al. National trends in use and outcomes of pulmonary artery catheters among medicare beneficiaries, 1999-2013. *JAMA Cardiol*. 2017;2(8):908–913.
5. Office of National Statistics. *Review of Avoidable Mortality Definition.* Cardiff: Government of the United Kingdom; 2015.
6. Khoja T, Farag MK. Synopsis of Indicators: Monitoring, Evaluation, and Supervision of Healthcare Quality. Kingdome of Saudi Arabia: Ministry of Health; 1995.
7. Kalra L, Evans A, Perez I, et al. Alternative strategies for stroke care: a prospective randomized controlled trial. *Lancet*. 2000;356:894–899.
8. Drummond AE, Pearson B, Lincoln NB, et al. Ten year follow-up of a randomized controlled trial of care in a stroke rehabilitation unit. *BMJ*. 2005;331:491–492.
9. Kahn KL, Rubenstein LV, Draper D, et al. The effects of DRG-based prospective payment system on quality of care for hospitalized Medicare patients: an introduction to the series. *JAMA*. 1990;264:1953–1955.
10. Kosecoff J, Kahn KL, Rogerts WH, et al. Prospective payment system and impairment at discharge: the "quicker and sicker" story revisited. *JAMA*. 1990;264:1980–1983.
11. Wallenstein ME, Ananth CV, Kim JH, et al. Effect of surgical volume on outcomes for laparoscopic hysterectomy for benign indications. *Obstet Gynecol*. 2012;119:709–716.
12. Birkmeyer JD, Stukel TA, Siewers AE, et al. Surgeon volume and operative mortality in the United States. *N Engl J Med*. 2003;349:2117–2127.
13. Haruyama Y, Yamazaki T, Endo M, et al. Personal status of general health checkups and medical expenditure: a large-scale community-based retrospective cohort study. *J Epidemiol*. 2017;27(5):209–214.

14. Gierisch JM, Earp JA, Brewer NT, et al. Longitudinal predictors of nonadherence to maintenance of mammography. *Cancer Epidemiol Biomarkers Prev*. 2010;19(4):1103–1111.
15. Gordis L. Effectiveness of comprehensive-care programs in preventing rheumatic fever. *N Engl J Med*. 1973;289:331–335.
16. Lin GM, Colangelo LA, Lloyd-Jones DM, et al. Association of sleep apnea and snoring with incident atrial fibrillation in the Multi-Ethnic Study of Atherosclerosis. *Am J Epidemiol*. 2015;182:49–57.

Preguntas de repaso del capítulo 17

1. *Todas las siguientes son medidas de proceso de la asistencia sanitaria en una clínica, excepto:*

a. Proporción de pacientes en los que se mide la presión arterial.

b. Proporción de pacientes que tienen complicaciones de una enfermedad.

c. Proporción de pacientes a quienes se aconseja dejar de fumar.

d. Proporción de pacientes en quienes se mide la talla y el peso.

e. Proporción de pacientes cuya factura se reduce debido a necesidades económicas.

2. *El grado en el que un tratamiento, servicio, procedimiento, programa u otra intervención específica de asistencia sanitaria hace aquello para lo que está diseñado cuando se utiliza en una población que vive en la comunidad se denomina:*

a. Eficacia.

b. Efectividad.

c. Modificación del efecto.

d. Eficiencia.

5. Ninguna de las anteriores.

3. *El grado en el que un tratamiento, servicio, procedimiento, programa u otra intervención específica de asistencia sanitaria produce un resultado beneficioso en condiciones controladas ideales se denomina:*

a. Eficacia.

b. Efectividad.

c. Modificación del efecto.

d. Eficiencia.

e. Ninguna de las anteriores.

4. *Un problema principal cuando se utiliza un diseño de controles histórico para la evaluación de un servicio de salud usando la letalidad como resultado es que, si la letalidad es más baja después de iniciar la prestación del servicio de salud, entonces:*

a. La menor letalidad podría deberse a la variación de la prevalencia de la enfermedad.

b. La menor letalidad puede ser secundaria a la disminución de la incidencia.

c. La menor letalidad puede ser un efecto indirecto del nuevo servicio de salud.

d. La letalidad puede haberse visto afectada por modificaciones de factores que no están relacionados con el nuevo servicio de salud.

e. Ninguna de las anteriores.

La pregunta 5 se basa en la siguiente información:

Letalidad intrahospitalaria de 100 varones no tratados en una unidad de cuidados coronarios (UCC) y de 100 varones tratados en una UCC, según tres grados clínicos de gravedad del infarto de miocardio (IM)

	NO UCC (N.° DE PACIENTES)			UCC (N.° DE PACIENTES)		
Grado Clínico	Total	Fallecidos	Letalidad (%)	Total	Fallecidos	Letalidad (%)
Leve	60	12	20	10	3	30
Grave	36	18	50	60	18	30
Shock	4	4	100	30	13	43

Los resultados mostrados se basan en la comparación de los últimos 100 pacientes tratados antes de instaurar la unidad de cuidados coronarios (UCC) y los primeros 100 pacientes tratados en la UCC. Los 200 pacientes fueron ingresados durante el mismo mes.

Se puede asumir que este es el *único hospital* de la ciudad y que la historia natural del IM *no se modificó* durante ese periodo.

5. *Los autores concluyeron que la UCC fue muy beneficiosa para los varones con IM grave y para aquellos en estado de shock, ya que la letalidad*

intrahospitalaria para estas categorías fue mucho menor en la UCC. Esta conclusión:

a. Es correcta.

b. Puede ser incorrecta porque se utiliza la letalidad en lugar de la mortalidad.

c. Puede ser incorrecta debido a un sesgo de derivación de los pacientes a este hospital a partir de los hospitales de ciudades lejanas.

d. Puede ser incorrecta debido a las diferencias al asignar el grado de gravedad clínica antes y después de la apertura de la UCC.

e. Puede ser incorrecta debido a que no se identifica una posible disminución de la incidencia anual de infarto de miocardio en los últimos años.

CAPÍTULO 18

Enfoque epidemiológico en la evaluación de los programas de cribado

Palabras clave

detección precoz de la enfermedad
cribado
fase preclínica
fase preclínica detectable
progresión de la enfermedad
sesgo de derivación o del voluntario
muestreo con sesgo de longitud
sesgo de tiempo de adelanto y supervivencia
sesgo de sobrediagnóstico
análisis coste-beneficio

Entre todas las palabras tristes pronunciadas o escritas

Las más tristes son estas:

«Podría haber sido»[1].

—J.G. Whittier, 1856

Si, de entre todas las palabras tristes pronunciadas o escritas,

Las más tristes son «Podría haber sido»,

Más tristes son estas que escuchamos a diario:

«Es, pero no tendría que ser»[2].

—Bret Harte, 1871

Objetivos de aprendizaje

- Profundizar en la descripción de la validez y la fiabilidad de las pruebas de cribado iniciada en el capítulo 5.
- Revisar la historia natural de la enfermedad e introducir los conceptos de tiempo de adelanto y punto crítico.
- Describir las principales fuentes de sesgo que deben tenerse en cuenta en la evaluación de los resultados del estudio que comparan las poblaciones cribada y no cribada, como el sesgo de derivación, el muestreo con sesgo de longitud, el sesgo de tiempo de adelanto, la supervivencia a 5 años y el sesgo de sobrediagnóstico.
- Comentar los diferentes diseños de estudio para la evaluación de los programas de cribado, como los estudios no aleatorizados y aleatorizados, así como las dificultades a la hora de interpretar los resultados de estos estudios.
- Comentar los problemas al evaluar la sensibilidad y especificidad de las pruebas de cribado desarrolladas a nivel comercial.
- Plantear los problemas relacionados con los análisis de rentabilidad del cribado.

En el capítulo 1 se hizo la distinción entre prevención primaria, secundaria y terciaria. En la sección 2 se comentó el diseño y la interpretación de los estudios dirigidos a identificar los factores de riesgo o factores etiológicos de las enfermedades para poder prevenir por completo la aparición de la enfermedad *(prevención primaria)*. En este capítulo se abordará cómo se utiliza la epidemiología en la evaluación de la eficacia de los programas de cribado para la detección precoz de la enfermedad *(prevención secundaria)*. Este tema es especialmente importante tanto en la práctica clínica como en la salud pública, ya que hay una aceptación creciente de la obligación que tiene el

médico de incluir la prevención, junto con el diagnóstico y el tratamiento, entre sus responsabilidades principales en la asistencia clínica de los pacientes.

La validez y fiabilidad de las pruebas de cribado se comentaron en el capítulo 5. En este capítulo se abordarán algunos de los problemas metodológicos que deben tenerse en cuenta a la hora de derivar inferencias sobre los beneficios que pueden obtener las personas en las que se realizan pruebas de cribado.

La cuestión de si los pacientes se benefician de la detección temprana de una enfermedad consta de los siguientes componentes:

1. ¿La enfermedad se puede detectar de forma precoz?
2. ¿Cuáles son la sensibilidad y la especificidad de la prueba?
3. ¿Cuál es el valor predictivo de la prueba?
4. ¿Cuál es la gravedad del problema de los resultados falsos positivos?
5. ¿Cuál es el coste de la detección precoz en términos económicos, de recursos y de impacto emocional?
6. ¿Las pruebas de detección suponen algún perjuicio a las personas?
7. ¿Las personas en las que la enfermedad se detecta precozmente se benefician de ello? ¿Existe algún beneficio global para quienes se someten al cribado?

Este capítulo se centra sobre todo en la última pregunta. Algunos de los otros aspectos de la lista anterior solo se consideran en el contexto de esta pregunta.

El término *detección precoz de la enfermedad* hace referencia al diagnóstico de una enfermedad en una etapa anterior a lo que normalmente sucede en la práctica clínica estándar. Esto generalmente denota la detección de la enfermedad en un estadio presintomático, cuando el paciente no tiene ningún dato clínico (ni síntomas ni signos) y, por tanto, no hay razón para solicitar asistencia médica para la enfermedad. En el cribado se parte de la premisa de que se dispone de una intervención apropiada para la enfermedad que se detecta y que la intervención médica se puede aplicar de manera más eficaz si la enfermedad se detecta en una etapa más temprana.

A primera vista, la cuestión de si las personas se benefician de la detección precoz de la enfermedad puede parecer un tanto sorprendente. De forma intuitiva, parecería obvio que la detección precoz es beneficiosa y que la intervención en una fase más temprana del proceso de la enfermedad es más eficaz y/o más fácil de implementar que una intervención posterior. En realidad, estos supuestos son una visión «quirúrgica»; por ejemplo, todas las lesiones malignas son de tipo localizado en una etapa precoz, en la que se pueden extirpar con éxito antes de que se produzca la diseminación regional o antes de que produzcan metástasis generalizadas. Sin embargo, el atractivo intuitivo de este concepto no debe hacernos perder de vista el hecho de que, a lo largo de la historia de la medicina, las convicciones más profundas han

resultado a menudo ser erróneas cuando no se sustentaban en datos obtenidos en estudios diseñados de forma adecuada y realizados con rigor. Por tanto, a pesar del atractivo que ofrece la idea de los aspectos beneficiosos de la detección temprana de la enfermedad, tanto para los médicos implicados en la prevención y el tratamiento como para los que intervienen en los programas de prevención basados en la comunidad, la evidencia que sustenta la validez de este concepto debe analizarse rigurosamente.

Al igual que en la evaluación de cualquier tipo de servicio de salud, el cribado puede evaluarse utilizando medidas de proceso o de resultado. En el cuadro 18.1 se recoge una lista de las medidas operativas que incluye medidas de proceso, así como medidas de rendimiento y la información producida por el programa de cribado.

Cuadro 18.1 Evaluación de la efectividad de los programas de cribado mediante medidas operativas

1. Número de personas sometidas a cribado.
2. Proporción de la población diana sometida a cribado y número de veces que se realiza el cribado.
3. Prevalencia detectada de enfermedad preclínica.
4. Costes totales del programa.
5. Costes por caso detectado.
6. Costes por caso detectado hasta entonces desconocido.
7. Proporción de personas con cribado positivo en las que se establece el diagnóstico definitivo y se instaura un tratamiento.
8. Valor predictivo de una prueba positiva en la población sometida a cribado.

Modificado de Hulka BS. Degrees of proof and practical application. Cancer 1988;62:1776–1780. Copyright © 1988 American Cancer Society. Reproducida con autorización de Wiley-Liss, Inc., una filial de John Wiley & Sons, Inc.

Existe un interés especial por la cuestión de qué beneficio obtienen las personas que se someten a un programa de cribado. Sin embargo, como sucede en la evaluación de los servicios de salud (se comenta en el cap. 17), mejorar el *proceso* de cribado proporciona una utilidad escasa si las personas sometidas a este no obtienen ninguna ventaja. Es decir, si la detección temprana no conduce a una mejora en la supervivencia, ¿cuál es la ganancia para los pacientes si son diagnosticados antes? ¡Quizás solo sea un tiempo más largo para preocuparse por la mala calidad de vida! Por ello, se analizarán algunos de los problemas asociados con la determinación de si la detección precoz de la enfermedad confiere beneficios a las personas que se someten al cribado (dicho de otro modo, si el cribado mejora el resultado).

¿Qué significa *resultado*? Para responder a la pregunta de si los pacientes se

benefician, hay que definir con precisión lo que se quiere decir con beneficio y qué resultado o resultados son considerados como evidencia de beneficio para el paciente. Algunas de las posibles medidas de resultado que podrían utilizarse se muestran en el cuadro 18.2.

Cuadro 18.2 Evaluación de la efectividad de los programas de cribado mediante medidas de resultado

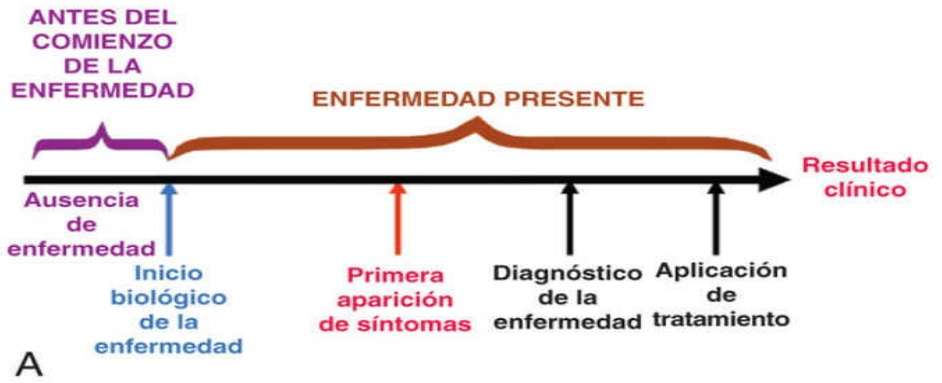
1. Reducción de la mortalidad en la población sometida a cribado.
2. Reducción de la letalidad en las personas sometidas a cribado.
3. Aumento porcentual de los casos detectados en etapas más precoces.
4. Reducción de las complicaciones.
5. Prevención o reducción de las recidivas o metástasis.
6. Mejora de la calidad de vida en las personas sometidas a cribado.

Historia natural de la enfermedad

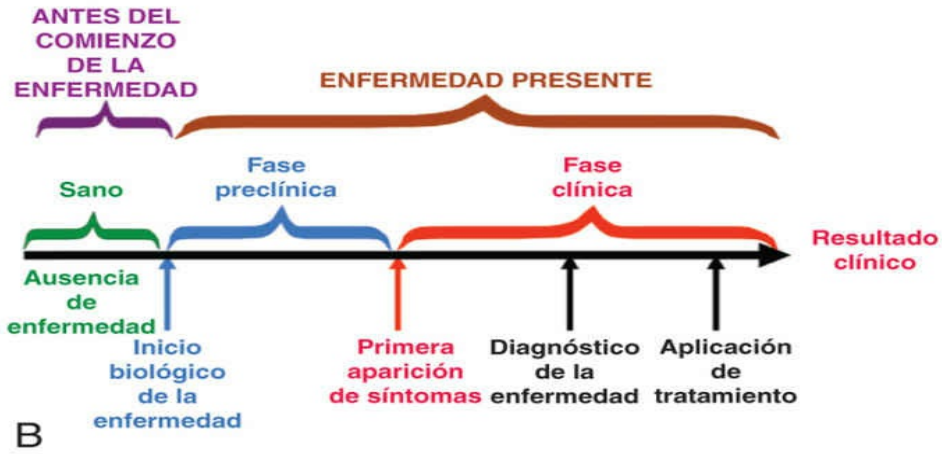
Para comentar las cuestiones metodológicas implicadas en la evaluación del beneficio del cribado, analizaremos con más detalle la historia natural de la enfermedad (se comentó por primera vez en el cap. 6).

Comenzaremos colocando el cribado en su lugar apropiado en la línea de tiempo de la historia natural de la enfermedad y lo haremos en relación con las diferentes estrategias de prevención expuestas en el capítulo 1.

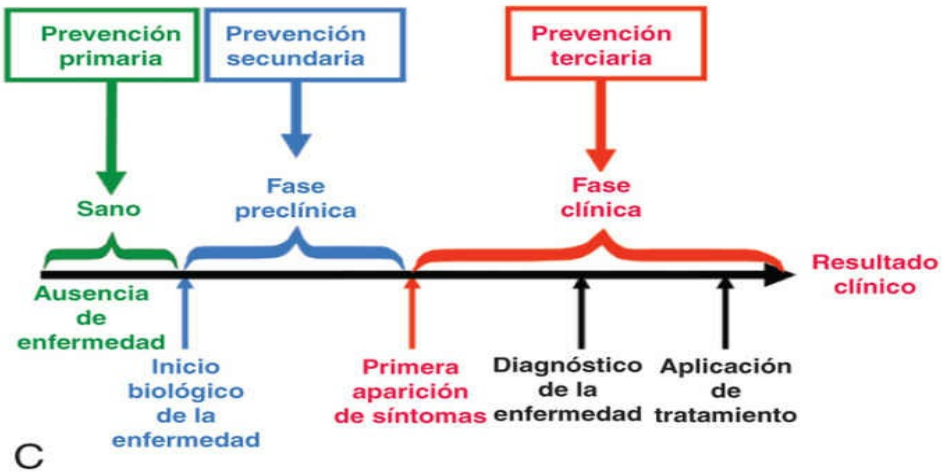
En la figura 18.1A se muestra una representación esquemática de la historia natural de una enfermedad en un individuo. En algún momento se produce el inicio biológico de la enfermedad. Esto puede ser un cambio subcelular, como una alteración en el ADN, lo que suele ser indetectable. En algún momento posterior, la enfermedad se vuelve sintomática o desarrolla signos clínicos (es decir, la enfermedad entra ahora en una fase clínica). Los síntomas y signos clínicos (p. ej., sangre en heces) hacen que el paciente solicite asistencia médica, tras lo que se establece un diagnóstico y se instaura un tratamiento apropiado, cuyo resultado final puede ser la curación, el control de la enfermedad, la discapacidad o el fallecimiento.


ANTES DEL COMIENZO DE LA ENFERMEDAD
ENFERMEDAD PRESENTE
Resultado clínico
Ausencia de enfermedad
Inicio biológico de la enfermedad
Primera aparición de síntomas
Diagnóstico de la enfermedad
Aplicación de tratamiento

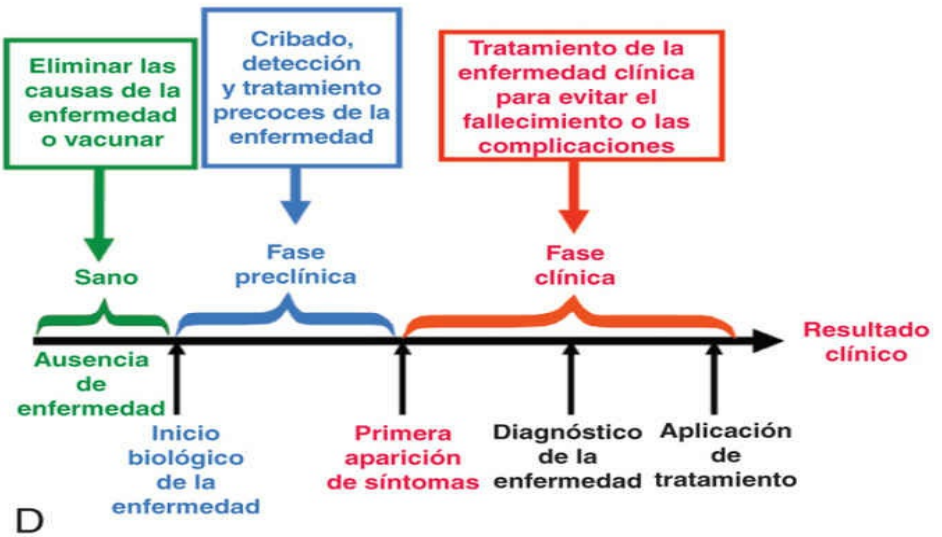
A


ANTES DEL COMIENZO DE LA ENFERMEDAD
ENFERMEDAD PRESENTE
Sano
Fase preclínica
Fase clínica
Resultado clínico
Ausencia de enfermedad
Inicio biológico de la enfermedad
Primera aparición de síntomas
Diagnóstico de la enfermedad
Aplicación de tratamiento

B


Prevención primaria
Prevención secundaria
Prevención terciaria
Sano
Fase preclínica
Fase clínica
Resultado clínico
Ausencia de enfermedad
Inicio biológico de la enfermedad
Primera aparición de síntomas
Diagnóstico de la enfermedad
Aplicación de tratamiento

C


Eliminar las causas de la enfermedad o vacunar
Cribado, detección y tratamiento precoces de la enfermedad
Tratamiento de la enfermedad clínica para evitar el fallecimiento o las complicaciones
Sano
Fase preclínica
Fase clínica
Resultado clínico
Ausencia de enfermedad
Inicio biológico de la enfermedad
Primera aparición de síntomas
Diagnóstico de la enfermedad
Aplicación de tratamiento

D

FIG. 18.1 (A) Historia natural de una enfermedad. (B) Historia natural de una enfermedad con las fases preclínica y clínica. (C) Historia natural de una enfermedad con puntos para prevención primaria, secundaria y terciaria. (D) Historia natural de una enfermedad con medidas específicas de prevención primaria, secundaria y terciaria.

Como se observa en la figura 18.1B, el inicio de los síntomas marca un punto importante en la historia natural de la enfermedad. El periodo en el que la enfermedad está presente se puede dividir en dos fases. El periodo comprendido entre el inicio biológico de la enfermedad y el desarrollo de los signos y síntomas es la *fase preclínica* de la enfermedad, que tiene lugar antes de la fase clínica de la enfermedad.

El periodo desde que se desarrollan los signos y los síntomas hasta el resultado final, que puede ser la curación, el control de la enfermedad o el fallecimiento, se conoce como la *fase clínica* de la enfermedad. Como se observa en las figuras 18.1C y D, la *prevención primaria* (es decir, la prevención del desarrollo de la enfermedad mediante la prevención o la reducción de la exposición a sus agentes causales), consiste en la intervención antes de que se haya desarrollado la enfermedad. (La prevención de la exposición a los factores de riesgo, como la inmunización y la prevención del inicio del hábito de fumar, también se conoce como *prevención primordial*.) La *prevención secundaria*, es decir, la detección de la enfermedad en una etapa más precoz de lo habitual, por ejemplo, mediante cribado, se lleva a cabo durante la *fase preclínica* de la enfermedad (es decir, después de que la enfermedad se haya desarrollado, pero antes de la aparición de signos y síntomas clínicos). La *prevención terciaria* hace referencia al tratamiento de las personas con enfermedad clínica para prevenir las complicaciones de la enfermedad (p. ej., la rehabilitación en un ictus), incluido el fallecimiento del paciente.

Si se quiere detectar la enfermedad antes de lo habitual mediante programas de educación para la salud, se podría animar a las personas sintomáticas a que soliciten asistencia médica cuanto antes. Sin embargo, la identificación de las personas que tienen la enfermedad pero que son asintomáticas plantea una gran dificultad. Este capítulo se centra en la identificación de la enfermedad en las personas que aún no han desarrollado síntomas y que están en la fase preclínica de la enfermedad.

A continuación, analizaremos más en detalle la *fase preclínica* de la enfermedad (fig. 18.2). En algún momento durante la fase preclínica, es posible detectar la enfermedad mediante el uso de pruebas disponibles en la actualidad (v. fig. 18.2A). El intervalo desde este momento hasta la aparición de signos y síntomas es la *fase preclínica detectable* de la enfermedad (v. fig. 18.2B). Cuando la enfermedad se detecta mediante una prueba de cribado, el momento del diagnóstico se adelanta a un punto anterior en la historia natural de la enfermedad respecto al que hubiera ocurrido si no se hubiera realizado la prueba de cribado. El *tiempo de adelanto* se define como el intervalo de tiempo que se anticipa el diagnóstico mediante el cribado para la detección precoz de la enfermedad en comparación con el momento habitual

del diagnóstico (v. fig. 18.2C). El concepto de tiempo de adelanto es inherente a la idea de cribado y de detección de una enfermedad antes de lo que se diagnosticaría normalmente.

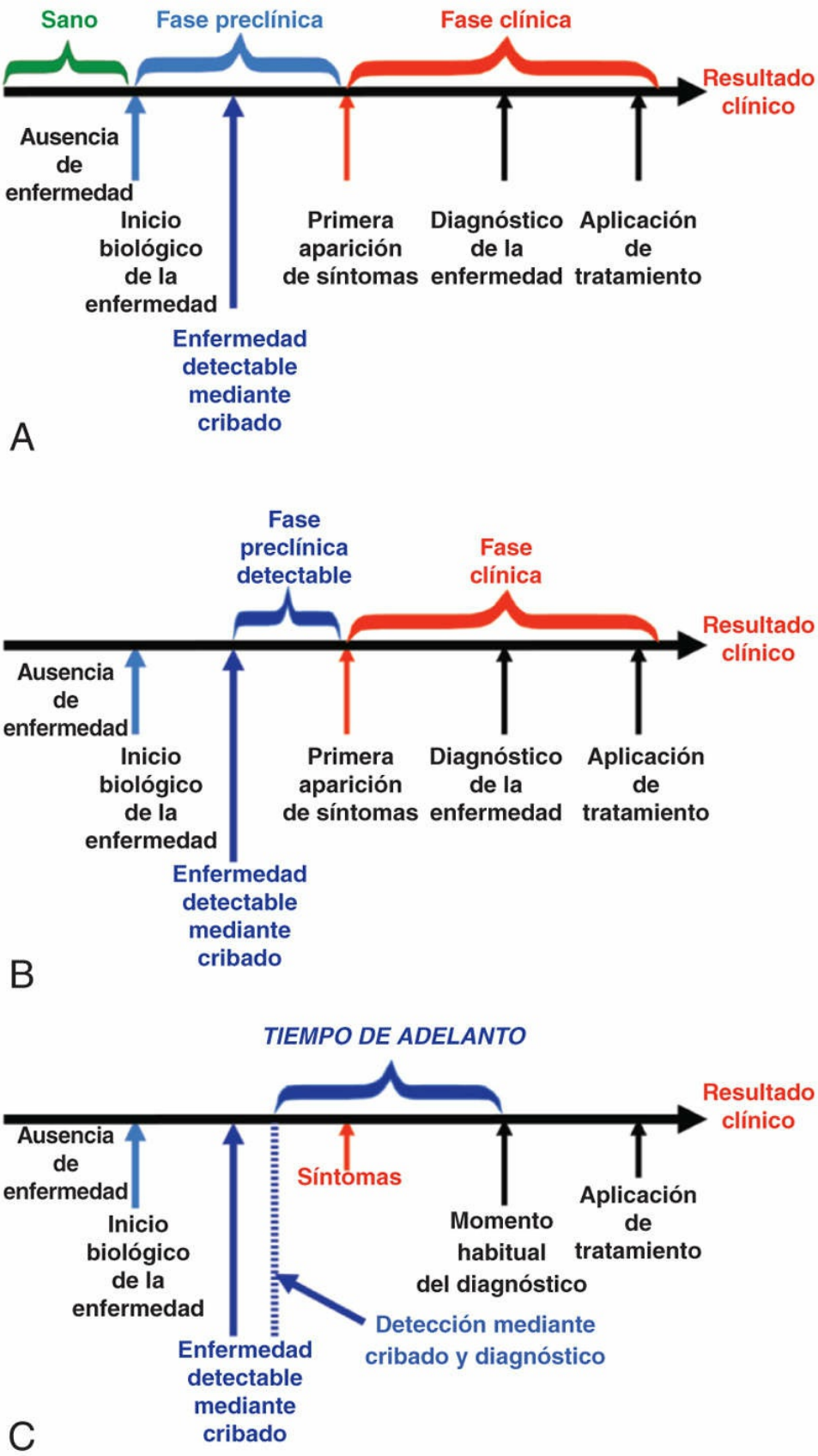


FIG. 18.2 Fase preclínica de la enfermedad. (A) Historia natural con un punto en el que la enfermedad es detectable mediante cribado. (B) Historia natural con la fase preclínica detectable. (C) Historia natural con tiempo de adelanto.

Otro concepto importante en el cribado es si existe un *punto crítico* en la

historia natural de una enfermedad[3] (fig. 18.3A). Este es un punto de la historia natural antes del cual el tratamiento es más eficaz y/o menos difícil de administrar. Si una enfermedad es potencialmente curable, la curación puede ser posible antes de este punto, pero no después. Por ejemplo, en una mujer con cáncer de mama, un punto crítico sería aquel en el que la enfermedad se disemina de la mama a los ganglios linfáticos axilares. Si la enfermedad se detecta y se trata antes de su extensión, el pronóstico es mucho mejor que después de haberse producido la diseminación a los ganglios.

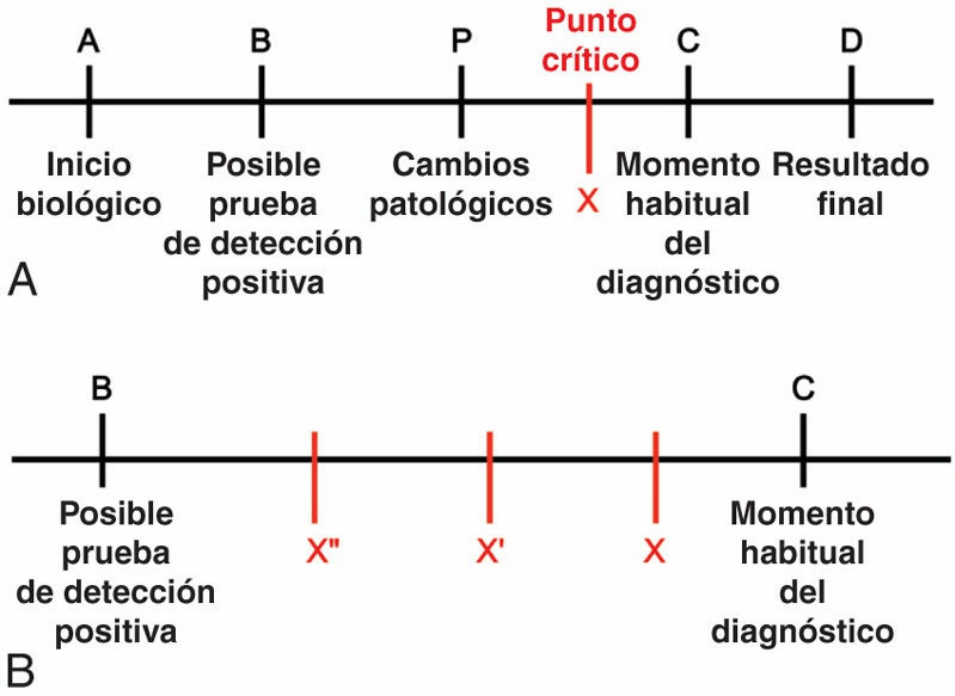


FIG. 18.3 (A) Un único punto crítico en la historia natural de una enfermedad. (B) Múltiples puntos críticos en la historia natural de una enfermedad. (Modificada de Hutchison GB. Evaluation of preventive services. *J Chronic Dis.* 1960;11:497–508.)

Como se muestra en la figura 18.3B, puede haber múltiples puntos críticos en la historia natural de una enfermedad. Por ejemplo, en la paciente con cáncer de mama, un segundo punto crítico puede ser aquel en el que la enfermedad se disemina desde los ganglios axilares a otras partes más distantes del cuerpo. El pronóstico aún es mejor cuando la enfermedad se limita a los ganglios linfáticos axilares que cuando se ha producido la diseminación sistémica, pero no tan bueno como cuando la enfermedad se limita a la mama. El concepto de puntos críticos múltiples sugiere que cuanto más temprano sea el diagnóstico, mejor será el pronóstico.

Sin embargo, el punto crítico es un concepto un tanto teórico, porque por lo general no se puede identificar cuándo se alcanza dicho punto crítico. Sin embargo, es un concepto fundamental en el cribado. Si no es posible identificar uno o más puntos críticos en la historia natural de una

enfermedad, es evidente que no hay justificación para el cribado y la detección precoz. La detección precoz supone que existe un punto biológico en la historia natural de una enfermedad antes del cual el tratamiento beneficiará a una persona más que si recibe el tratamiento después de ese punto.

Patrón de progresión de la enfermedad

Es de esperar que el cribado y la detección temprana sean beneficiosos si se cumplen las dos suposiciones siguientes:

1. Todos o la mayoría de los casos clínicos de una enfermedad pasan primero por una fase preclínica detectable.
2. En ausencia de intervención, todos o la mayoría de los casos progresan de la fase preclínica a la fase clínica.

Ambas suposiciones son bastante obvias. Por ejemplo, si ninguno de los casos preclínicos progresa a una fase clínica, no hay ninguna razón para realizar pruebas de cribado. Alternativamente, si ninguno de los casos clínicos pasa por una fase preclínica, no hay ninguna razón para llevar a cabo dichas pruebas. Por tanto, ambas suposiciones son importantes para evaluar cualquier posible beneficio del cribado.

Veamos el ejemplo del cribado del cáncer cervical. Han pasado unos 80 años desde que se desarrolló la prueba de Papanicolaou (Pap) para probar la presencia de células precancerosas o cancerosas del cuello uterino, la abertura del útero. Durante este procedimiento de rutina se obtienen células cervicales mediante raspado alrededor del cuello uterino y luego se examinan. La biología del cáncer cervical ha sido bien documentada, pasando por una serie de pasos desde la displasia al carcinoma in situ, hasta el cáncer cervical invasivo que a menudo tarda años en desarrollarse. Así, la detección temprana a menudo permite que el tratamiento detenga la progresión de este cáncer. Más recientemente, con la documentación del origen vírico del cáncer cervical (infección por el virus del papiloma humano [VPH]), la detección selectiva del cáncer cervical se realiza en la actualidad mediante la detección del VPH, y menos con la prueba de Papanicolaou. La figura 18.4A muestra la progresión de un cuello uterino normal a un cáncer cervical. Podríamos esperar que la detección y el tratamiento de más casos en la etapa in situ (no invasiva) se refleje en una reducción proporcional en el número de casos que progresan a una enfermedad invasiva.

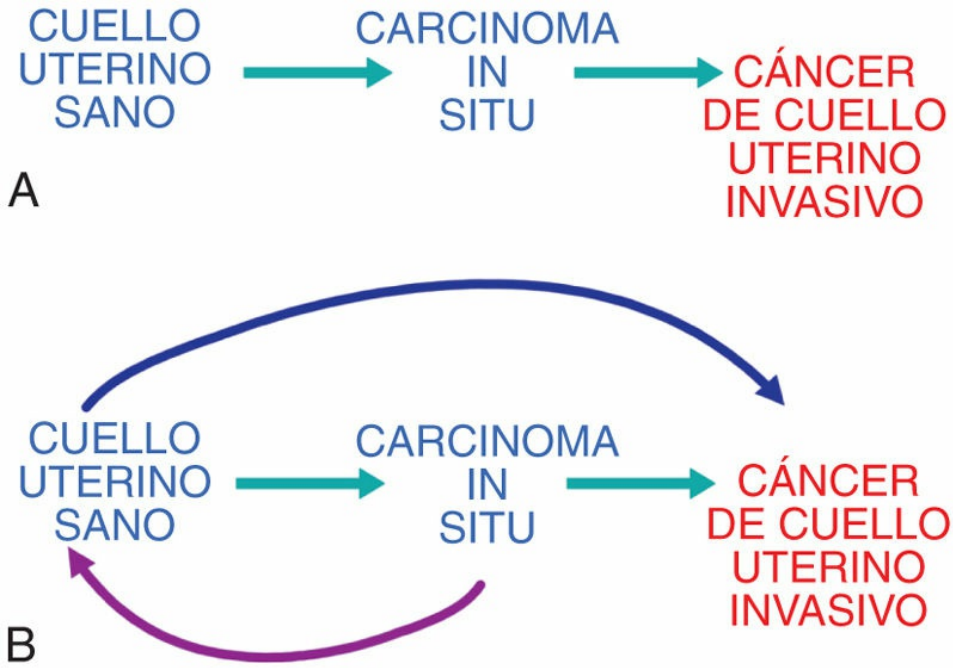


FIG. 18.4 (A) Historia natural del cáncer de cuello uterino: I. Progresión del cuello uterino sano a un cáncer invasivo. (B) Historia natural del cáncer de cuello uterino: II. Progresión extremadamente rápida y regresión espontánea.

Sin embargo, los dos supuestos asociados con la detección precoz son cuestionables. En ciertas situaciones, y a diferencia de lo que sucede con el cáncer cervical, la fase preclínica puede ser tan corta que es poco probable detectar la enfermedad mediante cualquier programa de cribado periódico. Además, cada vez hay más evidencia de que se puede producir una remisión espontánea en algunas enfermedades; por tanto, no todos los casos preclínicos avanzan inexorablemente a una enfermedad clínica. Es importante destacar que, en el caso de la detección del VPH en mujeres,la mayoría de los tipos de VPH detectados en el cribado de rutina generalmente revierten (desaparecerán espontáneamente) en los siguientes 6 meses.

Sin embargo, la evaluación de los beneficios del cribado del cáncer de cuello uterino se complica por el problema de que algunos casos pasan por el estadio in situ tan deprisa y la fase preclínica es tan breve que para todos los fines prácticos no hay etapa preclínica durante la que se pueda detectar la enfermedad mediante cribado. Además, los estudios de cuantificación del ADN nuclear sugieren que pueden existir anomalías intraepiteliales cervicales, bien como un estado reversible o bien como un precursor irreversible del cáncer invasivo. Los datos también sugieren que algunos casos de neoplasia intraepitelial cervical detectados mediante un frotis de Papanicolaou remiten de forma espontánea, sobre todo en las primeras etapas, pero también en la última etapa (carcinoma in situ). En un estudio, un tercio de las mujeres con anomalías en la citología vaginal que rechazaron someterse a una intervención tuvieron más adelante citologías normales. Además, existen datos que sugieren que la mayoría, o todas, las neoplasias in situ de cuello uterino se asocian con diferentes tipos de virus del papiloma. Solo las neoplasias asociadas con ciertos tipos de virus del papiloma de alto

riesgo progresan a un cáncer invasivo, por lo que puede que nos enfrentemos a una heterogeneidad tanto del agente causal como de la enfermedad.

El modelo simple de progresión desde un cuello uterino sano al cáncer de cuello uterino invasivo visto en la figura 18.4A sugiere que la detección precoz seguida por una intervención eficaz se refleja en una reducción proporcional del número de lesiones invasivas que se desarrollan posteriormente, en la figura 18.4B se muestra una representación más precisa de la historia natural del cáncer cervical. La magnitud de ambos fenómenos (la remisión espontánea y la progresión muy rápida) influye claramente en el grado de disminución de la enfermedad invasiva que se podría esperar gracias a la detección e intervención precoces y, por tanto, debe tenerse en cuenta al evaluar los beneficios del cribado. Aunque estas cuestiones se han demostrado para el cáncer de cuello uterino, tienen una relevancia evidente para evaluar los beneficios del cribado de muchas enfermedades.

Problemas metodológicos

Cuando se van a interpretar los resultados de un estudio diseñado para evaluar los beneficios del cribado, se deben tener en cuenta ciertos problemas metodológicos. La mayoría de los estudios de los programas de cribado que se han publicado no han sido ensayos aleatorizados, debido a las dificultades de la asignación aleatoria de una población para el cribado. Por tanto, la pregunta es: ¿podemos analizar a un grupo de personas que han sido sometidas a cribado y comparar su mortalidad con la de un grupo de personas sin cribado (es decir, utilizar un diseño de cohortes para evaluar la eficacia del cribado)?

Supongamos que comparamos una población de personas que han sido sometidas a cribado para una enfermedad con una población de personas que no lo han sido. Supongamos, además, que se dispone de un tratamiento viable y efectivo y se va a utilizar eficazmente en quienes se detecta la enfermedad. Si se observa una mortalidad menor por la enfermedad en las personas en las que la enfermedad se identificó mediante el cribado que en aquellas en las que la enfermedad no se ha detectado de esta manera, ¿es posible concluir que el cribado y la detección precoz de la enfermedad han sido beneficiosos? A continuación, se comentarán algunos de los problemas metodológicos implicados.

Sesgos de selección

Sesgo de derivación (sesgo del voluntario)

A la hora de alcanzar una conclusión acerca de los beneficios del cribado, la primera pregunta que debemos plantearnos es si hubo un sesgo de selección en términos de quién se ha sometido al cribado y quién no. Nos gustaría ser capaces de asumir que quienes se sometieron al cribado tenían las mismas características que quienes no lo hicieron (es decir, eran similares entre sí en todos los aspectos, excepto en el antecedente de cribado). Sin embargo, hay muchas diferencias en las características entre quienes participan en el cribado o aprovechan otros programas de salud y quienes no lo hacen. En muchos estudios se ha demostrado que los voluntarios tienen un mejor estado de salud que la población general y que son más propensos a cumplir las recomendaciones médicas (son más cumplidores). Si, por ejemplo, las personas cuya enfermedad tenía un mejor pronóstico desde el principio fueron derivadas para realizar el cribado o fueron autoseleccionadas, podría observarse una mortalidad menor en el grupo del cribado, aunque la detección precoz no tuviese ningún papel a la hora de mejorar el pronóstico. Por supuesto, también es posible que en los voluntarios se incluyan muchas personas de alto riesgo y que se presenten voluntarios para someterse al cribado porque tienen temores debido a unos antecedentes familiares

positivos o a sus propias características de estilo de vida. El problema es que no sabemos en qué dirección podría actuar el sesgo de selección y cómo podría afectar a los resultados del estudio.

El problema del sesgo de selección que afecta en mayor medida a nuestra interpretación de los resultados se aborda mejor realizando la comparación con un estudio experimental aleatorizado en el que se procure que los dos grupos tengan perfiles pronósticos iniciales comparables (fig. 18.5).



FIG. 18.5 Diseño de un ensayo aleatorizado sobre los beneficios del cribado.

Muestreo con sesgo de longitud (selección por pronóstico)

El segundo tipo de problema que surge al interpretar los resultados de una comparación de un grupo sometido a cribado con otro sin cribado es un posible sesgo de selección que no se relaciona con quién acude para el cribado, sino más bien con el tipo de enfermedad que se detecta mediante dicho cribado. La pregunta es: ¿el cribado identifica selectivamente los casos de la enfermedad que tienen un mejor pronóstico? Dicho de otro modo, ¿los casos encontrados mediante el cribado tienen una historia natural mejor, con independencia de la precocidad con la que se inicie el tratamiento? Si el resultado de las personas en las que la enfermedad se detecta mediante cribado es mejor que el de las personas en las que no se realiza el cribado y en quienes la enfermedad se identificó durante el curso normal de la asistencia clínica, ¿podría deberse el mejor resultado en las personas sometidas a cribado a la identificación selectiva por este de personas con un mejor pronóstico? ¿Podría el mejor resultado no estar relacionado con el momento en el que se establece el diagnóstico y se llevan a cabo las intervenciones terapéuticas?

¿Cómo podría suceder esto? Hay que recordar la historia natural de la enfermedad, con las fases clínica y preclínica (v. fig. 18.1B). Sabemos que la

duración de la fase clínica de la enfermedad difiere entre las distintas personas (es decir, hay una distribución natural de los parámetros de la enfermedad clínica en cada población). Por ejemplo, algunos pacientes con cáncer de colon fallecen poco después del diagnóstico, mientras que otros sobreviven muchos años. Lo que parece ser la misma enfermedad puede incluir a pacientes con fases clínicas de distinta duración.

¿Qué sucede con la fase preclínica en estas personas? En realidad, la enfermedad de cada paciente tiene una sola historia natural continua, que se divide en las fases preclínica y clínica (fig. 18.6) en función del momento en el que se desarrollan los signos y síntomas. En algunas, la historia natural es breve y en otras es prolongada. Esto sugiere que si una persona tiene una historia natural lentamente progresiva con una fase clínica larga, la fase preclínica también será larga. Por el contrario, si una persona tiene un proceso de enfermedad rápidamente progresivo y una historia natural breve, es probable que la fase clínica sea corta, y parece razonable concluir que la fase preclínica también será corta. De hecho, hay datos que apoyan la idea de que una fase clínica larga se asocia con una fase preclínica larga y que una fase clínica corta se asocia con una fase preclínica corta. El cáncer de pulmón sirve como ejemplo: tiene una fase clínica corta y muy probablemente también una fase preclínica corta, como lo sugieren los resultados inconsistentes de los ensayos clínicos de fumadores cribados por tomografía computarizada, de modo que algunos ensayos muestran una efectividad de aproximadamente el 15 al 20%. y otros no muestran efectividad alguna[4].

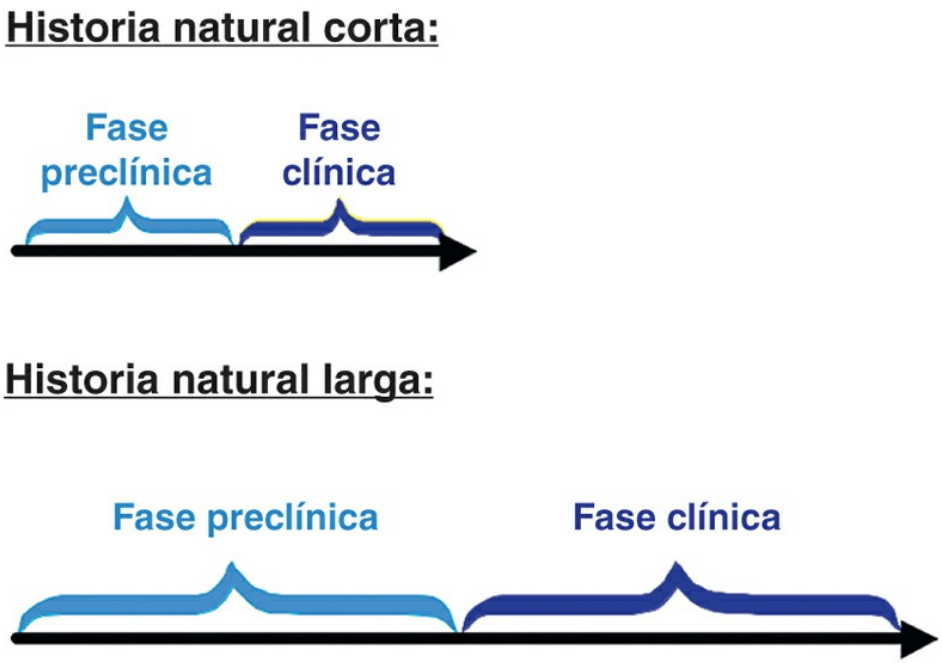


FIG. 18.6 Historias naturales corta y larga de la enfermedad: relación de la duración de la fase clínica con la duración de la fase preclínica.

Hay que recordar que nuestro propósito al realizar el cribado es detectar la

enfermedad durante la fase preclínica, ya que durante la fase clínica el paciente ya es consciente del problema e, incluso sin cribado, probablemente solicitará asistencia médica para los síntomas. Si se organiza un programa de cribado puntual en una comunidad, ¿qué grupo de pacientes es probable que se identifiquen: los que tienen una fase preclínica corta o una fase preclínica larga?

Para responder a esta pregunta, consideremos una pequeña población que se somete a cribado para una determinada enfermedad (fig. 18.7). Como se muestra en la figura 18.7, cada caso tiene una fase preclínica y una fase clínica. En la figura se representa cada fase preclínica con la misma longitud que su fase clínica asociada. Los pacientes en la fase clínica se identificarán en el curso habitual de asistencia médica, por lo que el propósito del cribado es identificar los casos en el estado preclínico (es decir, antes de la aparición de algún signo o síntoma). Hay que tener en cuenta que las duraciones de las fases preclínicas de los casos representados aquí varían. Cuanto más larga sea la fase preclínica, más probable será que el programa de cribado detecte el caso cuando todavía está en la fase preclínica. Por ejemplo, si se realiza el cribado una vez al año para una enfermedad cuya fase preclínica es de solo 24 horas de duración, es evidente que se pasarán por alto la mayoría de los casos durante la fase preclínica. Sin embargo, si la fase preclínica es de 1 año, muchos más casos se identificarán durante ese periodo. El cribado tiende a identificar selectivamente aquellos casos que tienen fases preclínicas de la enfermedad más largas. Por tanto, incluso si el tratamiento posterior no tuviese ningún efecto, el cribado aún identificaría selectivamente a las personas con una fase preclínica larga y, por ende, con una fase clínica larga (es decir, aquellas con un mejor pronóstico). Estas personas tendrían un pronóstico mejor incluso si no hubiera un programa de cribado o incluso si el cribado no tuviese verdaderos beneficios.

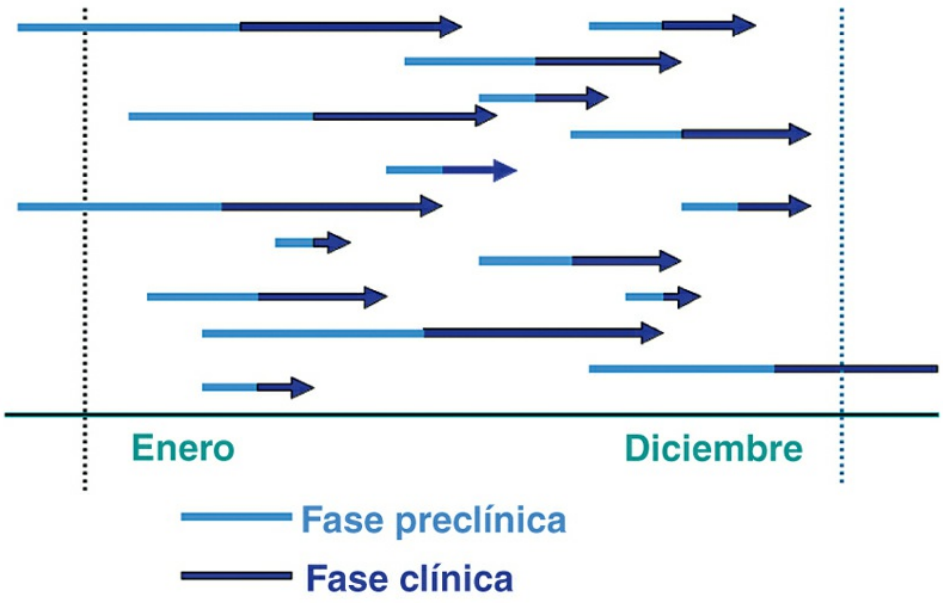


FIG. 18.7 Población hipotética de personas con historias naturales larga y corta.

Este problema puede abordarse de varias maneras. Una estrategia consiste en utilizar un diseño aleatorizado experimental en el que se procura mantener los grupos comparables en cuanto a las duraciones de la fase preclínica detectable de la enfermedad. Sin embargo, puede que esto no sea fácil. Además, debería analizarse la supervivencia de todos los miembros de cada grupo (es decir, tanto los que se someten a cribado como los que no). En el grupo del cribado, se debe calcular la supervivencia de aquellos en los que la enfermedad se detecta mediante cribado y de aquellos en los que la enfermedad se detecta en el periodo entre las pruebas de cribado, los denominados *casos de intervalo.* Se volverá sobre la importancia de los casos de intervalo más adelante en este capítulo.

Sesgo de tiempo de adelanto

Otro problema que surge al comparar la supervivencia de las personas sometidas a cribado con la de aquellas en las que no se realiza el cribado es el sesgo de tiempo de adelanto (ilustrado inicialmente en la fig. 18.2C). ¿Con cuánta antelación se puede establecer el diagnóstico si la enfermedad se detecta mediante cribado en comparación con el momento habitual del diagnóstico si no se realiza el cribado?

Consideremos cuatro personas con una determinada enfermedad representadas por las cuatro líneas de tiempo de la figura 18.8. La parte más gruesa de cada línea horizontal indica la supervivencia aparente que se observa. La primera línea de tiempo (A) muestra el momento habitual del diagnóstico y el momento habitual del fallecimiento. La segunda línea de tiempo (B) muestra un momento de diagnóstico más precoz, pero el mismo momento del fallecimiento. La supervivencia *parece* mejor porque el intervalo desde el diagnóstico hasta el fallecimiento es más largo, pero el paciente no está mejor porque el fallecimiento no se ha retrasado. La tercera línea de tiempo (C) muestra un diagnóstico precoz y un retraso del fallecimiento por la enfermedad, lo que es claramente un beneficio para el paciente (suponiendo que la calidad de vida posterior sea buena). Por último, la cuarta línea de tiempo (D) muestra un diagnóstico más precoz, con la subsiguiente prevención del fallecimiento por la enfermedad.

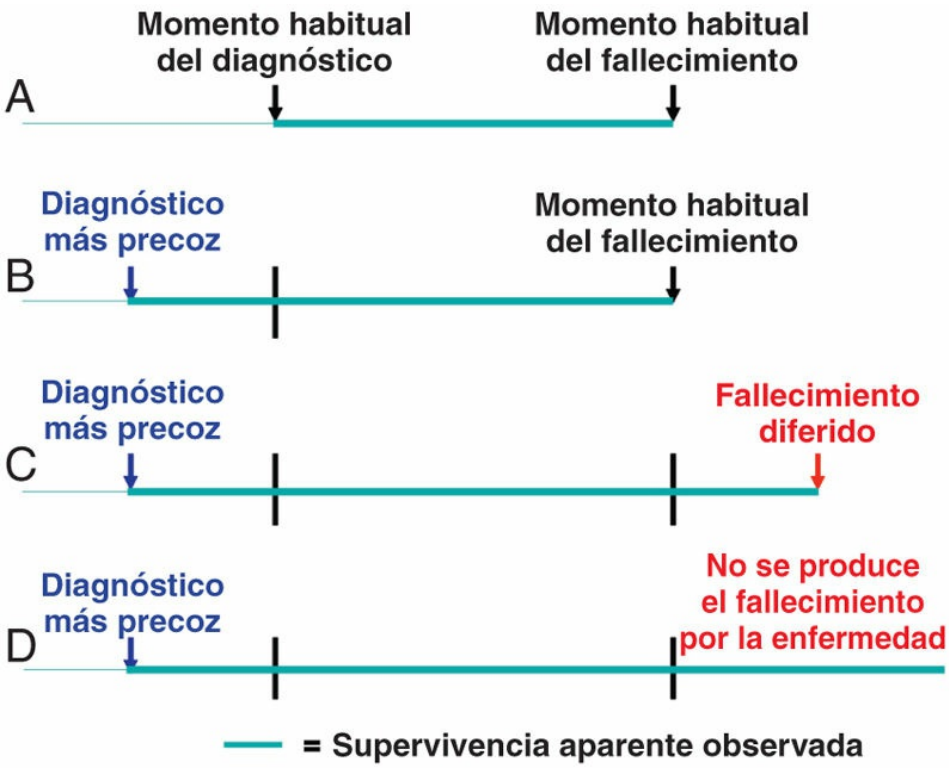


FIG. 18.8 (A) Resultado del diagnóstico en el momento habitual, sin cribado. (B-D) Tres posibles resultados de un diagnóstico más precoz como resultado de un programa de cribado.

Los beneficios que se buscan con el cribado son el retraso o la prevención del fallecimiento. Aunque en este capítulo hemos optado por centrarnos en la mortalidad, también podríamos haber utilizado parámetros de morbilidad, las recidivas, la calidad de vida o la satisfacción del paciente como medidas válidas de resultado.

Tiempo de adelanto y supervivencia a 5 años

La supervivencia a cinco años es una medida de uso frecuente del éxito terapéutico, sobre todo en el tratamiento del cáncer. A continuación, se analizará el posible efecto del tiempo de adelanto sobre la supervivencia a 5 años aparente.

En la figura 18.9A se muestra la historia natural de la enfermedad en un paciente hipotético con cáncer de colon, que fue diagnosticado en el contexto clínico habitual sin ningún tipo de cribado. La aparición biológica de la enfermedad fue en 2008. El paciente advirtió los síntomas en 2016, y se realizó un estudio diagnóstico que culminó en el diagnóstico de cáncer de colon. El paciente fue intervenido quirúrgicamente en 2016, pero falleció de cáncer de colon en 2018. Este paciente ha sobrevivido 2 años (2016-18) y es evidente que no es un superviviente a 5 años. Si se utiliza la supervivencia a 5 años como índice de éxito del tratamiento, este paciente representa un fracaso del tratamiento.

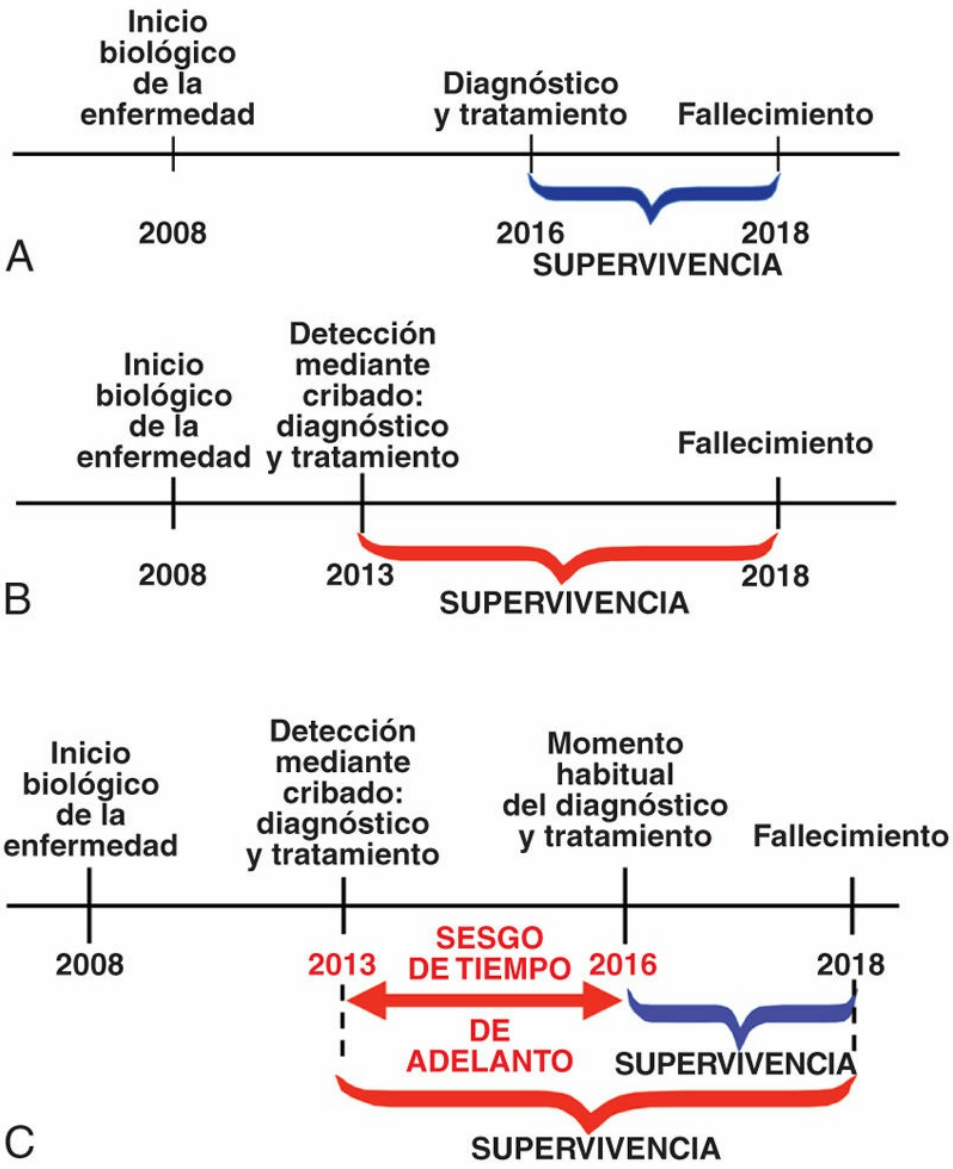


FIG. 18.9 (A) Historia natural de un paciente con cáncer de colon sin cribado. Enfermedades diagnosticadas y tratadas en 2008. (B) Enfermedad detectada por cribado 3 años antes en 2013 (tiempo de adelanto). (C) Sesgo de tiempo de adelanto debido al cribado 3 años antes.

Ahora, consideremos lo que le podría suceder a este paciente si reside en una comunidad en la que se inicia un programa de cribado (v. fig. 18.9B). Para este ejemplo hipotético solamente, supongamos que la detección precoz no proporciona en realidad ningún beneficio (es decir, la historia natural del cáncer de colon no se ve afectada por la intervención precoz). En este caso, el paciente se encuentra asintomático, pero se somete a una prueba de cribado rutinaria en 2013, cuyo resultado es positivo. En 2013 se realiza la cirugía, pero el paciente fallece en 2018. El paciente ha sobrevivido 5 años y ahora es claramente un superviviente a 5 años, aunque no porque la muerte se haya retrasado, sino porque el diagnóstico se ha adelantado. Al comparar este ejemplo de cribado con el ejemplo sin cribado (v. fig. 18.9A), es evidente que el paciente no ha obtenido ningún beneficio de la detección más precoz en términos de haber vivido más tiempo. De hecho, el paciente puede haber

salido perdiendo en términos de calidad de vida, ya que la detección más precoz de la enfermedad mediante el cribado le ha proporcionado un añadido de 3 años de postoperatorio y de otro tipo de asistencia médica y puede haberle privado de 3 años de vida normal. Este problema de una ilusión de una mejor supervivencia por una detección más precoz se denomina *sesgo de tiempo de adelanto*, como se muestra en la figura 18.9C.

Por tanto, aunque no se obtenga un verdadero beneficio por la detección precoz de una enfermedad, *parecerá* que existe un beneficio asociado con el cribado, incluso si el fallecimiento no se retrasa, debido a que el diagnóstico se establece en un momento más precoz respecto a aquel en el que se mide la supervivencia. Esto no significa que la detección precoz no conlleve ningún beneficio, sino que, incluso sin ningún beneficio, el tiempo de adelanto asociado con la detección temprana sugiere la aparición de un beneficio en forma de aumento de la supervivencia. Así pues, el tiempo de adelanto debe tenerse en cuenta al interpretar los resultados de las evaluaciones no aleatorizadas.

En la figura 18.10 se muestra el efecto del sesgo resultante del tiempo de adelanto sobre las estimaciones cuantitativas de la supervivencia. En la figura 18.10A se muestra una situación en la que no se está llevando a cabo ninguna actividad de cribado. Cinco años después del diagnóstico, la supervivencia es del 30%. Si se instaura un programa de cribado con un tiempo de adelanto de 1 año, todo el marco temporal se desplaza hacia la izquierda (v. fig. 18.10B). Si ahora se calcula la supervivencia a los 5 años a partir del nuevo momento del diagnóstico (v. fig. 18.10C), la supervivencia parece ser del 50%, pero solo como consecuencia del sesgo del tiempo de adelanto. El problema es que la supervivencia aparentemente mejor no se debe a que las personas sometidas al cribado vivan más tiempo, sino más bien a que el diagnóstico se realiza en un momento más precoz de la historia natural de su enfermedad. En muchas enfermedades, como el cáncer, el paciente no puede morir antes del inicio de la fase clínica y, por lo tanto, el tiempo anterior al diagnóstico precoz y habitual (es decir, el tiempo de espera) refleja lo que también se conoce como «sesgo de tiempo inmortal».

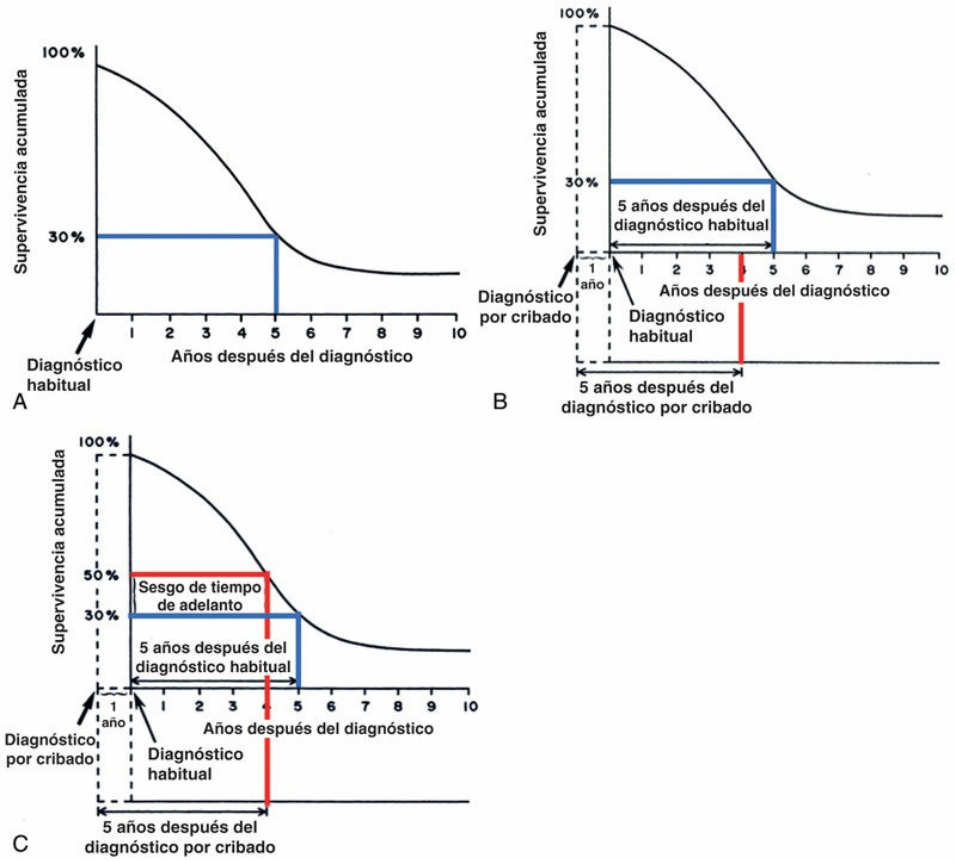


FIG. 18.10 (A) Sesgo de tiempo de adelanto: I: supervivencia a 5 años cuando el diagnóstico se realiza sin cribado. (B) Sesgo de tiempo de adelanto: II: desplazamiento del periodo de 5 años mediante el cribado y la detección precoz (tiempo de adelanto) (C) Sesgo de tiempo de adelanto: III: sesgo en el cálculo de la supervivencia debido a la detección precoz. (Modificada de Frank JW. Occult-blood screening for colorectal carcinoma: The benefits. *Am J Prev Med.* 1985;1:3–9.)

Por tanto, en cualquier comparación entre la población en la que se realiza el cribado y aquella en la que no se lleva a cabo, hay que tener en cuenta un tiempo de adelanto estimado para intentar identificar cualquier prolongación de la supervivencia más allá de la que se debe al efecto artificial del tiempo de adelanto. Si la detección precoz se asocia en realidad con una mayor supervivencia, la supervivencia en el grupo en el que se realiza el cribado debería ser mayor que la del grupo control *más el tiempo de adelanto*. Hay que realizar, por tanto, una cierta estimación del tiempo de adelanto para la enfermedad que se está estudiando[5].

Otra estrategia consiste en comparar la mortalidad por la enfermedad en todo el grupo sometido a cribado con la del grupo sin cribado en lugar de solo la supervivencia acumulada o su recíproca, la tasa de letalidad en aquellos en los que la enfermedad se detectó mediante cribado.

Sesgo de sobrediagnóstico

Otro sesgo potencial es el de sobrediagnóstico. En ocasiones, las personas que inician un programa de cribado tienen un entusiasmo casi ilimitado por este.

Incluso los citopatólogos que leen las citologías vaginales para el cáncer de cuello uterino pueden llegar a mostrar tal entusiasmo que pueden tender a hacer una *sobrelectura* de estos (es decir, a hacer diagnósticos falsos positivos). Si se realiza una sobrelectura, algunas mujeres sanas se incluirán en el grupo donde el frotis de Papanicolaou se considera positivo. Por tanto, el grupo patológico se diluirá con mujeres que no tienen cáncer. Si las personas sanas del grupo de cribado tienen una mayor probabilidad de ser diagnosticadas erróneamente como positivas que las personas sanas del grupo sin cribado (es decir, clasificadas como pacientes con cáncer cuando en realidad no lo son), se podría obtener una impresión falsa de unas mayores tasas de detección y diagnóstico del cáncer en estadio precoz como resultado del cribado. Además, dado que muchas de las personas con un diagnóstico de cáncer en el grupo de cribado en realidad no tendrían cáncer y, por tanto, su supervivencia sería buena, los resultados representarían una estimación exagerada de la supervivencia después del cribado de personas que se cree que tienen cáncer, lo que da lugar a la conclusión errónea de que se ha demostrado que el cribado mejora la supervivencia de los pacientes con cáncer en esta población.

El posible impacto cuantitativo del sobrediagnóstico secundario al cribado se demuestra en el ejemplo hipotético ilustrado en la figura 18.11. En la figura 18.11A se muestra la situación 1, en la que no se realiza cribado. En esta situación, 1.000 pacientes con cáncer de pulmón clínico son seguidos durante 10 años. En ese momento, 900 personas han fallecido y 100 están vivas. Por tanto, la supervivencia a los 10 años para los 1.000 pacientes es de 100/1.000, es decir, del 10%.

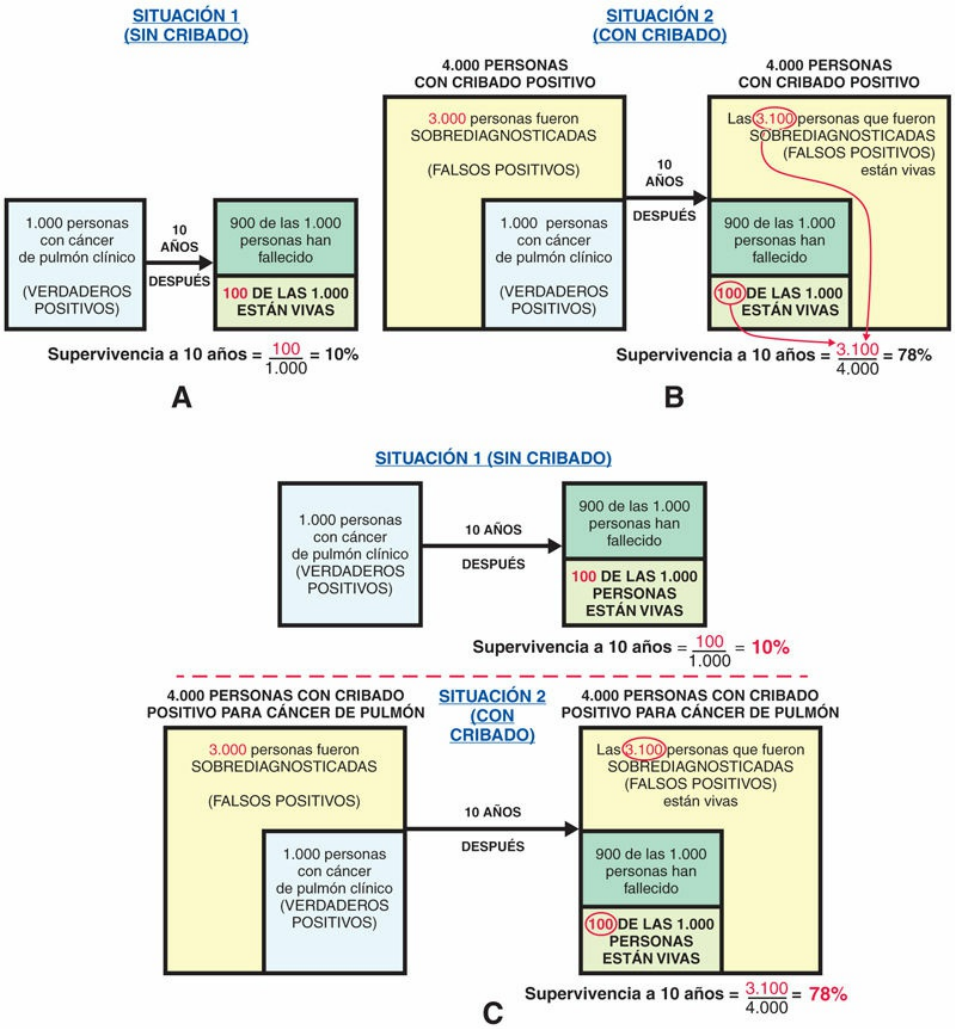


FIG. 18.11 Impacto del sobrediagnóstico debido al cribado sobre la estimación de la supervivencia. (A) Situación 1. Supervivencia sin cribado. (B) Situación 2. El cribado causa un sobrediagnóstico: supervivencia a los 10 años. (C) Comparación de la supervivencia a los 10 años entre la situación 1 y la situación 2. (Modificada de Welch HG, Woloshin S, Schwartz LM. Overstating the evidence for lung cancer screening: The International Early Lung Cancer Action Program [I-ELCAP] study. *Arch Intern Med.* 2007;167:2289–2295.)

En la figura 18.11B se muestra la situación 2, en la que el cribado da lugar a un sobrediagnóstico. En esta situación, 4.000 personas tienen un resultado positivo en el cribado del cáncer de pulmón. De ellas, 1.000 son los mismos pacientes con cáncer de pulmón clínico vistos en la figura 18.11A, y los otros 3.000 son personas que no tienen cáncer de pulmón, pero que son sobrediagnosticadas por la prueba de cribado como positivas para el cáncer de pulmón (falsos positivos).

Después de 10 años, estas 3.000 personas aún están vivas, al igual que las 100 personas que tenían cáncer de pulmón clínico y que sobrevivieron, como se muestra en la figura 18.11A. El resultado es que, de las 4.000 personas con un cribado positivo inicialmente, 3.100 han sobrevivido durante 10 años. Como se muestra en la comparación de las situaciones 1 y 2 en la figura 18.11C, la supervivencia a 10 años en la situación 2 es ahora del 78% frente al 10% en la situación 1 en la población original de 1.000 pacientes que

tenían cáncer de pulmón clínico. Sin embargo, la supervivencia aparentemente «mejor» que se observa en la situación 2 se debe por completo a la inclusión de 3.000 personas que no tenían cáncer de pulmón, pero que se han sobrediagnosticado por el método de cribado.

Este es un ejemplo de sesgo de clasificación errónea, como se describió en el capítulo 15. En este ejemplo, 3.000 personas sin cáncer de pulmón han sido clasificadas erróneamente por la prueba de cribado como pacientes con cáncer de pulmón. Es esencial, por tanto, que, en estos estudios de supervivencia, el proceso de diagnóstico esté rigurosamente estandarizado para minimizar el problema potencial del sobrediagnóstico.

Diseños de estudio para la evaluación del cribado: estudios no aleatorizados y aleatorizados

Estudios no aleatorizados

Al comentar los problemas metodológicos implicados en los estudios no aleatorizados del cribado, se han descrito esencialmente estudios observacionales de personas que han sido sometidas a cribado y de otras que no, es decir, un diseño de cohortes (fig. 18.12).



FIG. 18.12 Diseño de un estudio de cohortes no aleatorizado sobre los beneficios del cribado.

El diseño de casos y controles también ha sido utilizado como método para evaluar la eficacia del cribado (fig. 18.13). En este diseño, los «casos» son las personas con enfermedad avanzada (el tipo de enfermedad que se espera evitar mediante cribado). Se han realizado varias propuestas para escoger los controles apropiados para este tipo de estudio. Está claro que deben ser «no casos» (es decir, personas que no tienen enfermedad avanzada). Aunque los «controles» que se usaron en los primeros estudios de casos y controles para evaluar el cribado fueron personas con la enfermedad en una etapa precoz, muchos investigadores creen que las personas seleccionadas de la población de la que se derivan los casos son mejores controles. A continuación, se determina la prevalencia de antecedentes de cribado tanto en los casos como en los controles, por lo que el cribado se considera una «exposición». Si el

cribado es eficaz, sería de esperar encontrar una mayor prevalencia de antecedentes de cribado entre los controles que entre las personas con enfermedad avanzada, y se puede calcular el cociente de posibilidades *(odds ratio)*, que será inferior a 1,0 si el cribado es eficaz.

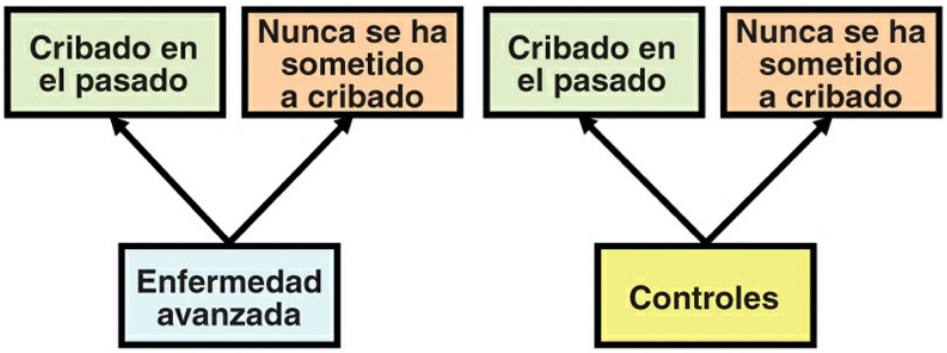


FIG. 18.13 Diseño de un estudio de casos y controles sobre los beneficios del cribado.

Estudios aleatorizados

En este tipo de estudio, una población se distribuye de forma aleatoria, la mitad para someterse al cribado y la otra mitad no. Este tipo de estudio es difícil de organizar y de llevar a cabo y puede estar cargado de consideraciones éticas. Quizás el ensayo aleatorizado más conocido sobre el cribado sea el ensayo del cribado de cáncer de mama mediante mamografía que se llevó a cabo en el Health Insurance Plan (HIP) de Nueva York[6]. Shapiro y cols. realizaron un ensayo aleatorizado en mujeres inscritas en el programa prepago HIP, una organización de mantenimiento de la salud temprana (HMO, *Health Maintenance Organization*) en Nueva York. Este estudio se ha convertido en un clásico de la literatura al presentar una evaluación de los beneficios del cribado mediante un diseño de ensayo aleatorizado, y sirve como modelo para futuros estudios de este tipo.

El estudio se inició en 1963. Fue diseñado para determinar si el cribado periódico mediante exploración física de la mama por parte de un médico y mamografía reducía la mortalidad por cáncer de mama en mujeres de 40-64 años. Unas 62.000 mujeres se asignaron de forma aleatoria a un grupo de estudio y a un grupo control, de alrededor de 31.000 personas cada uno (fig. 18.14). A las mujeres del grupo de estudio se les ofreció la posibilidad de realizar las exploraciones de cribado; el 65% de las mujeres acudieron para la primera exploración, y se les ofrecieron exploraciones adicionales a intervalos anuales. La mayoría de estas mujeres acudieron al menos a una de las tres exploraciones anuales de cribado que se les ofrecieron. El cribado consistió en una exploración física de la mama, mamografía y anamnesis. Las mujeres del grupo control recibieron la asistencia médica habitual del programa médico

prepago. Se han publicado muchos artículos de este excelente estudio; a continuación, analizaremos solo algunos de los resultados.



FIG. 18.14 Diseño del ensayo controlado aleatorizado del Health Insurance Plan *(HIP)* que comenzó en 1963 para estudiar la eficacia del cribado mamográfico. (Datos de Shapiro S, Venet W, Strax P, et al., eds. *Periodic Screening for Breast Cancer: The Health Insurance Plan Project and Its Sequelae, 1963–1986.* Baltimore: Johns Hopkins University Press; 1988.)

En la figura 18.15 se muestra el número de fallecimientos por cáncer de mama y las tasas de mortalidad tanto en el grupo de estudio (mujeres a quienes se les ofreció la mamografía de cribado) como en el grupo control después de 5 años de seguimiento.

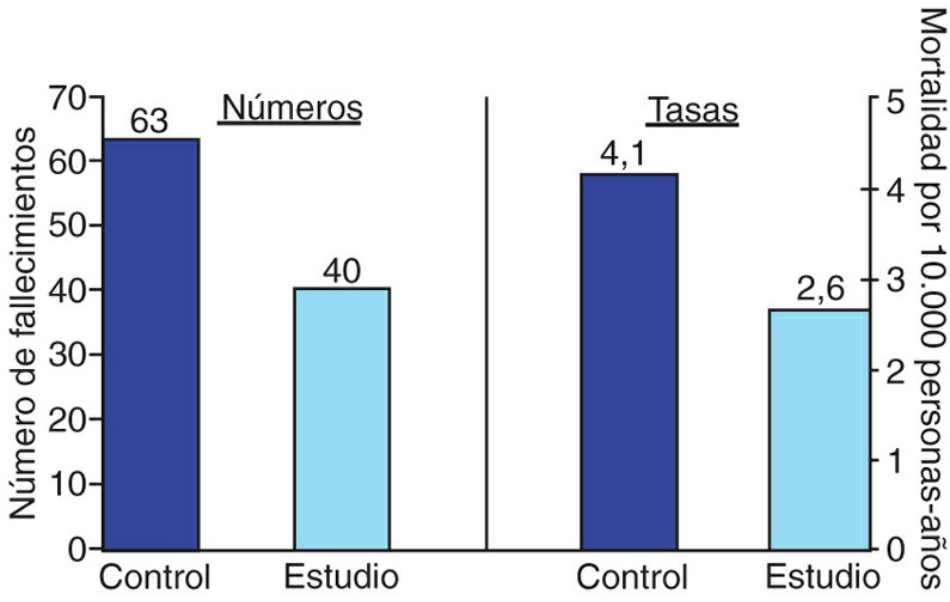


FIG. 18.15 Número de muertes por cáncer de mama y mortalidad por cáncer de mama en los grupos control y de estudio; 5 años de seguimiento después de la inclusión en el estudio. Los datos del grupo de estudio incluyen las muertes de las mujeres sometidas a cribado y de las que rechazaron el cribado. (Datos de Shapiro S, Venet W, Strax P, et al. Selection, follow-up, and analysis in the Health Insurance Plan Study: a randomized trial with breast cancer screening. *Natl Cancer Inst Monogr.* 1985;67:65–74.)

Obsérvese que los datos del grupo de estudio incluyen los fallecimientos de las mujeres que aceptaron el cribado y de las que lo rechazaron. Debe recordarse la presentación del problema del cruce no planificado en ensayos aleatorizados. En ese contexto, se señaló que el procedimiento estándar de análisis de los datos fue analizar de acuerdo con la asignación aleatoria original, estrategia denominada «por intención de tratar». Eso es precisamente lo que se hizo aquí. Una vez que una mujer se asignaba de forma aleatoria para la realización de la mamografía, se mantenía en ese grupo para los fines del análisis, incluso si posteriormente rechazaba el cribado. A pesar de esto, se observa que las muertes por cáncer de mama son mucho más elevadas en el grupo control que en el grupo de estudio.

En la figura 18.16 se muestra la letalidad a 5 años de las mujeres que desarrollaron cáncer de mama en ambos grupos. La letalidad en el grupo control fue del 40%. En el grupo de estudio *total* (mujeres que fueron asignadas de forma aleatoria para la realización de una mamografía, con independencia de si se sometían en realidad o no al cribado), la letalidad fue del 29%. Shapiro y cols. dividieron a continuación este grupo en las que se sometieron al cribado y las que lo rechazaron. En las que rechazaron el cribado, la letalidad fue del 35% y en las que lo aceptaron, del 23%.

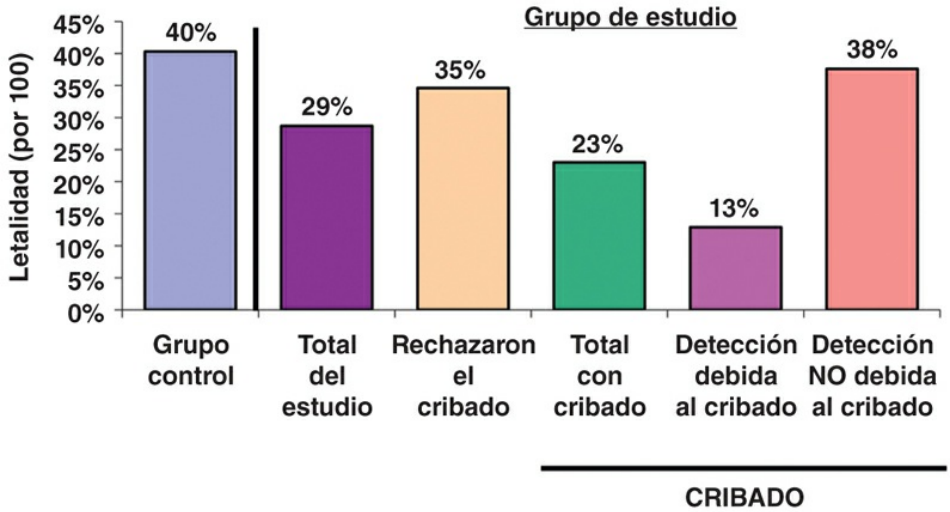


FIG. 18.16 Letalidad a los cinco años en pacientes con cáncer de mama. La letalidad en las mujeres en las que el tumor se detectó mediante el cribado presenta un tiempo de adelanto de 1 año. (Datos de Shapiro S, Venet W, Strax P, et al. Ten- to 14-year effect of screening on breast cancer mortality. *J Natl Cancer Inst.* 1982;69:349–355.)

Shapiro y cols. compararon después la supervivencia en mujeres cuyo cáncer de mama se detectó en la exploración de cribado con la de las mujeres cuyo cáncer de mama se diagnosticó entre las exploraciones de cribado (es decir, en quienes no se detectó cáncer de mama en el cribado y que, antes de la siguiente exploración un año después, desarrollaron síntomas que llevaron al diagnóstico de cáncer de mama). Si el cáncer había sido detectado por mamografía, la letalidad era tan solo del 13%. Sin embargo, si el cáncer de mama era un *caso de intervalo* (es decir, diagnosticado entre las exploraciones), la letalidad era del 38%. ¿Qué podría explicar esta diferencia de letalidad? La explicación más probable es que la enfermedad diagnosticada entre las exploraciones mamográficas periódicas era rápidamente progresiva. No era detectable en la exploración mamográfica periódica, pero se identificó antes de la siguiente exploración programada periódicamente un año después porque era muy agresivo (Otra posibilidad es que al menos algunos casos de intervalo aparente fueran en realidad casos que no se habían detectado en el examen de cribado anterior [es decir, eran falsos negativos].)

Estas observaciones también respaldan el concepto comentado anteriormente en este capítulo según el cual es probable que una fase clínica larga se asocie con una fase preclínica larga. Las mujeres en las que se detectaron signos de cáncer en el cribado tenían una fase preclínica larga y una letalidad de solo el 13%, lo que indica también una fase clínica larga. Las mujeres que habían tenido mamografías normales y cuya enfermedad se manifestó clínicamente antes de la siguiente exploración tuvieron una fase preclínica corta y, dada la elevada letalidad del grupo, también presentaron una fase clínica corta.

En la figura 18.17 se muestran los fallecimientos por causas *distintas al cáncer de mama* en ambos grupos a lo largo de 5 años. La mortalidad fue mucho mayor en las mujeres que no acudieron para realizarse el cribado que en las que sí lo hicieron. Debido a que el cribado solo estaba dirigido al cáncer de mama, ¿por qué las que acudieron para el cribado y las que no lo hicieron presentaban tasas de mortalidad diferentes por causas *distintas* al cáncer de mama? La respuesta es, sin duda, por el sesgo del voluntario (la observación bien documentada de que las personas que participan en los programas de salud difieren en muchos aspectos de quienes no lo hacen: en su estado de salud, actitudes, niveles educativo y socioeconómico y otros factores). Esta es otra demostración de que la comparación entre los participantes y los no participantes no es una estrategia válida a la hora de evaluar un programa de salud.

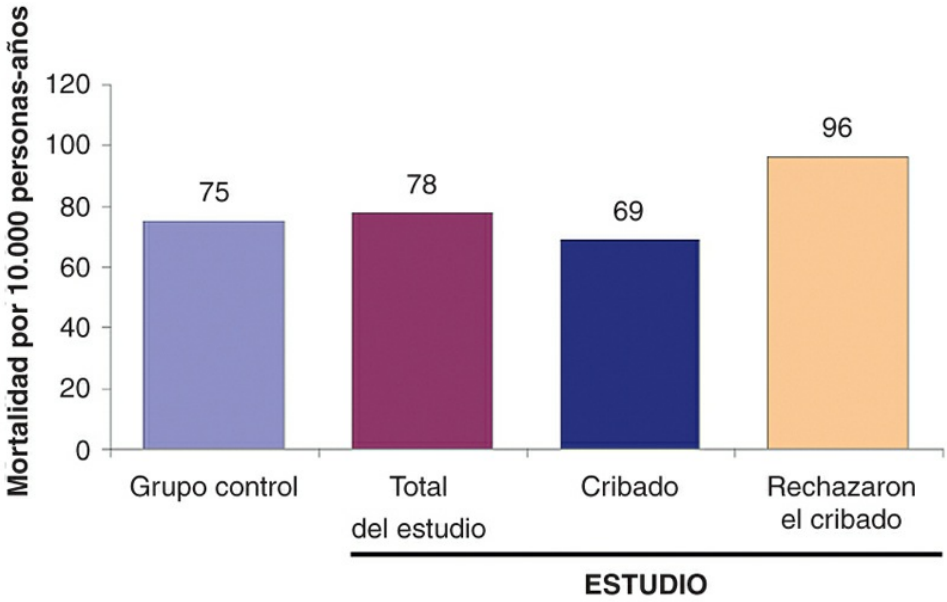


FIG. 18.17 Mortalidad por todas las causas *excepto el cáncer de mama* por 10.000 personas-años, Health Insurance Plan. (Datos de Shapiro S, Venet W, Strax P, et al. Selection, follow-up, and analysis in the Health Insurance Plan Study: A randomized trial with breast cancer screening. *Natl Cancer Inst Monogr.* 1985;67:65–74.)

Antes de abandonar la exposición del estudio HIP, podríamos realizar una digresión y mencionar una aplicación interesante de estos datos llevada a cabo por Shapiro y cols.[7]. En la figura 18.18 se muestra que, en Estados Unidos, la supervivencia relativa a 5 años en pacientes con cáncer de mama es mejor en mujeres de raza blanca que en afroamericanas.

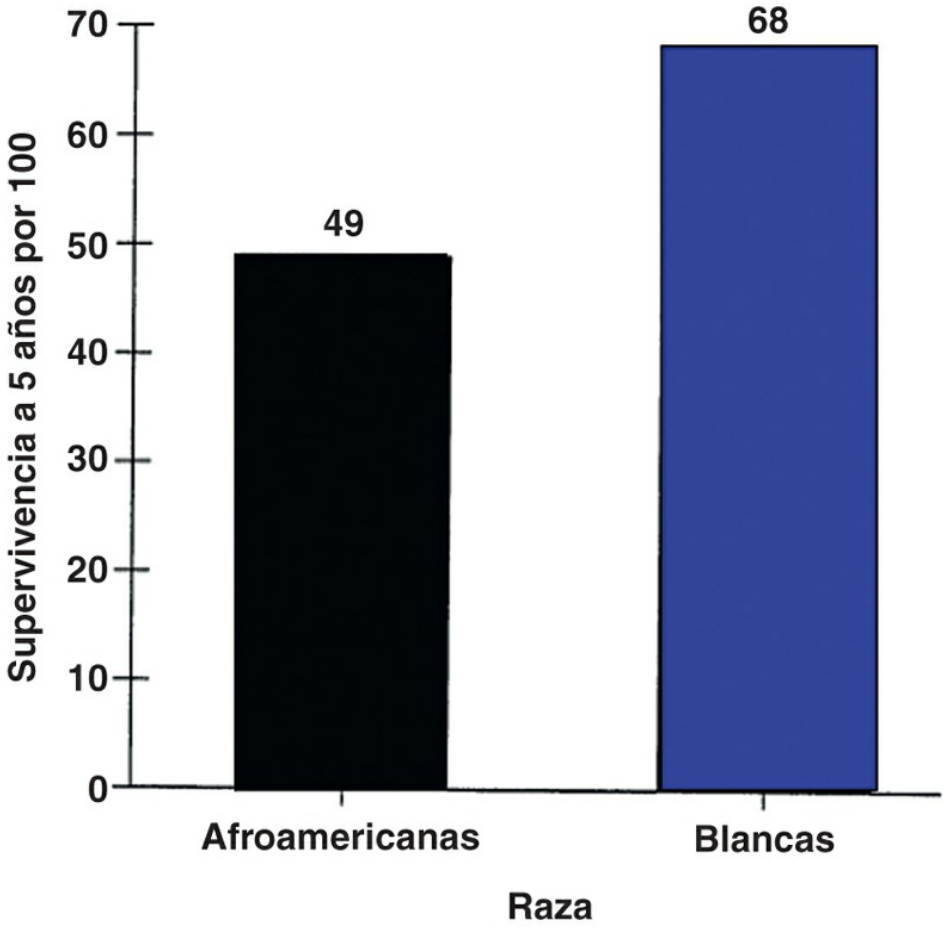


FIG. 18.18 Tasas de supervivencia relativa a cinco años, por raza, en mujeres con cáncer de mama diagnosticadas en 1964-1973 (programa Surveillance, Epidemiology and End Results). (Datos de Shapiro S, Venet W, Strax P, et al. Prospects for eliminating racial differences in breast cancer survival rates. *Am J Public Health.* 1982;72:1142–1145.)

Se ha planteado la cuestión de si esto se debe a una diferencia de la biología de la enfermedad entre ambas razas o a una diferencia del acceso a la asistencia de salud entre las personas de raza blanca y las afroamericanas, lo que puede retrasar el diagnóstico y el tratamiento de la enfermedad en estas últimas. Shapiro y cols. advirtieron que el ensayo aleatorizado sobre la mamografía ofrecía una oportunidad poco común para abordar esta cuestión. Los resultados se muestran en la figura 18.19. Si se analizan primero solo las curvas de supervivencia para las mujeres de raza blanca y afroamericanas del grupo control (v. fig. 18.19A), los datos concuerdan con los de la figura 18.18: las mujeres afroamericanas y las hispanas tenían un peor pronóstico que las de raza blanca. Cuando se analizan las curvas de las mujeres de raza blanca y afroamericanas del grupo de estudio de mujeres que se sometieron a cribado (y que, por tanto, no tenían diferencias de acceso a la asistencia o de utilización de la asistencia, pues el cribado se llevó a cabo según un programa predeterminado) (v. fig. 18.19B), se observa una superposición considerable de las dos curvas: prácticamente no hay diferencia. Esto sugiere con fuerza que el cribado había eliminado la diferencia racial de supervivencia y que la diferencia que suele observarse entre las razas en cuanto al pronóstico del

cáncer de mama se debe en realidad a un menor acceso a la asistencia o a un uso menor de la asistencia en las personas afroamericanas, con el consiguiente retraso en el diagnóstico y el tratamiento, y, por lo tanto, menor supervivencia.

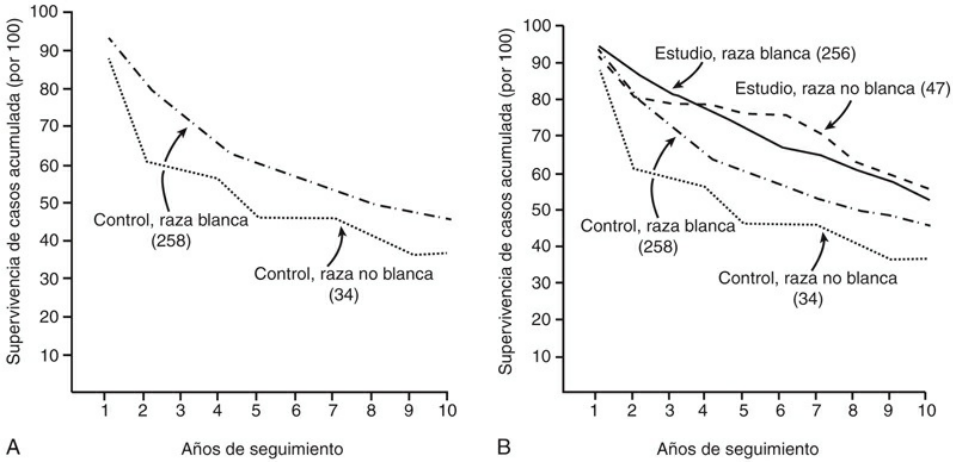


FIG. 18.19 (A) Tasas de supervivencia de los casos acumulada, primeros 10 años después del diagnóstico por raza, grupos control del Health Insurance Plan *(HIP)*. (B) Tasas de supervivencia de los casos acumulada, primeros 10 años después del diagnóstico por raza, grupos de estudio y control del Health Insurance Plan (HIP). (De Shapiro S, Venet W, Strax P, et al. Prospects for eliminating racial differences in breast cancer survival rates. *Am J Public Health.* 1982;72:1142–1145.)

Otros ejemplos de estudios que evalúan el cribado

Mamografía para mujeres de 40-49 años

En la década de 1990 ha habido una gran controversia sobre si la mamografía debe recomendarse universalmente para las mujeres de 40-49 años. Los datos del estudio de Shapiro y cols., así como de otros estudios, establecieron el beneficio de las mamografías periódicas para las mujeres a partir de los 50 años. Sin embargo, los datos son menos claros para las mujeres de 40-49 años. A la hora de interpretar los resultados de los ensayos aleatorizados llevados a cabo en varias poblaciones diferentes surgen muchos problemas. Aunque la reducción de la mortalidad en las mujeres de 40-49 años que se hacen mamografías anuales se ha estimado en un 17%, los datos disponibles suelen proceder de estudios que no fueron diseñados específicamente para evaluar los posibles beneficios en este grupo de edad. Por otra parte, muchos de los ensayos reclutaron a mujeres *de 45-49 años,* lo que sugiere la posibilidad de que, aunque se observaron beneficios, podrían deberse igualmente a mamografías realizadas en mujeres de 50 años o mayores.

Un problema relacionado se ilustra en la figura 18.20. Cuando se compara

la mortalidad a lo largo del tiempo en las mujeres *a partir de 50 años* en las que se realiza el cribado y en las que no (v. fig. 18.20A), las curvas de mortalidad divergen alrededor de unos 4 años después de la inclusión, de modo que el grupo de mamografía muestra una menor mortalidad que persiste en el tiempo. Sin embargo, cuando se comparan las *mujeres de 40- 49 años* en las que se realiza el cribado y en las que no (v. fig. 18.20B), las curvas de mortalidad no sugieren ninguna diferencia en la mortalidad durante al menos 11-12 años después de la inclusión. Se necesitará un mayor seguimiento para determinar si la divergencia observada en las curvas de mortalidad en realidad persistirá y si representa un verdadero beneficio en las mujeres que se han realizado mamografías a los 40-49 años. Sin embargo, la interpretación de estas curvas se complica, porque las mujeres que han sido seguidas durante 10 años o más en estos estudios han superado los 50 años de edad. Por tanto, aunque la mortalidad en las mujeres sometidas a cribado declina después de 11 años, puede que cualquier beneficio observado se deba a las mamografías que se realizaron *después* de los 50 años en lugar de a las mamografías efectuadas a los 40-49 años. El seguimiento posterior de las mujeres que participaron en muchos de estos estudios, y en estudios recién iniciados que están incluyendo a mujeres de 40-49 años, puede ayudar a aclarar estas cuestiones.

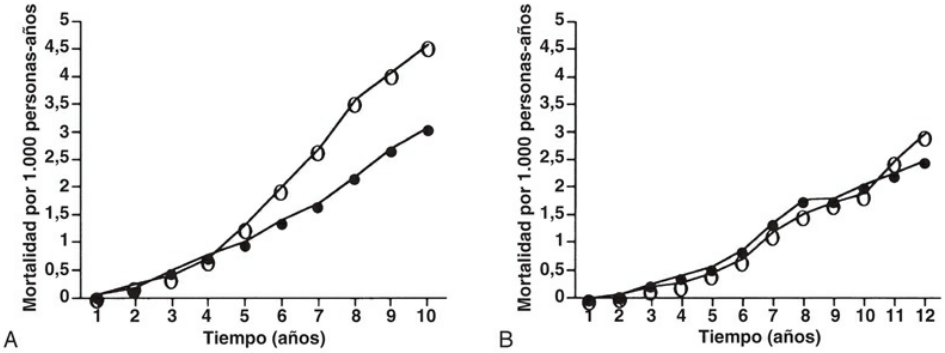


FIG. 18.20 Mortalidad acumulada por cáncer de mama en las mujeres en las que se realizó cribado y en las que no de 50-69 años (A) y de 40-49 años (B). • = con cribado; ^ = sin cribado. (De Kerlikowske K: Efficacy of screening mammography among women aged 40 to 49 years and 50 to 69 years: Comparison of relative and absolute benefit. *Natl Cancer Inst Monogr.* 1997;22:79–86. (A) Modificada de Tabar L, Fagerberg G, Duffy SW, et al. Update of the Swedish two-county program of mammographic screening for breast cancer. *Radiol Clin North Am.* 1992;30:187–210. (B) Modificada de Nystrom L, Rutqvist LE, Wall S, et al. Breast cancer screening with mammography: Overview of Swedish randomized trials. *Lancet.* 1993;341:973–978.)

En 1997, los National Institutes of Health crearon un panel de consenso (liderado por el profesor Gordis) para revisar la evidencia científica de los beneficios de la mamografía en las mujeres de 40-49 años. El panel concluyó que los datos disponibles no justificaban una recomendación universal de la mamografía para todas las mujeres de 40-49 años. El panel recomendó que cada mujer debe decidir por sí misma (tras consultar con su médico) si

realizarse la mamografía[8]. Su decisión puede basarse no solo en un análisis objetivo de la evidencia científica y la evaluación de sus antecedentes médicos individuales, sino también en cómo ella percibe y sopesa cada riesgo y beneficio potencial, el valor que otorga a cada uno y cómo maneja la incertidumbre. Dada tanto la importancia como la complejidad de las cuestiones implicadas en la evaluación de la evidencia, una mujer debería tener acceso a la mejor información pertinente posible sobre los beneficios y los riesgos, presentada de forma comprensible y utilizable.

La mayoría de las mujeres dependerán en gran medida de los conocimientos y la sofisticación de sus médicos en lugar de tomar ellas mismas la decisión sobre cuándo comenzar con las mamografías de cribado. Un problema importante en este sentido es que muchos médicos no tienen conocimientos suficientes de las estadísticas de cribado del cáncer para proporcionar el apoyo que necesitan las mujeres y sus familias a la hora de analizar cuidadosamente los resultados, conclusiones y validez de los estudios sobre las mamografías en mujeres de 40-49 años. Un estudio de Wegwarth y cols. presentó los resultados de una encuesta nacional realizada a médicos de atención primaria en Estados Unidos; se observó que la mayoría de ellos interpretaron por error la mejora de la supervivencia y el aumento de la detección con el cribado como prueba de que dicho cribado salva vidas. Pocos reconocieron correctamente que la reducción de la mortalidad en un estudio aleatorizado constituye una evidencia del beneficio del cribado[9].

El panel de consenso añadió que, para las mujeres de 40-49 años que deciden realizarse una mamografía, los costes de las mamografías deben ser reembolsados por terceros pagadores o cubiertos por las HMO para que los obstáculos financieros no influyan en la decisión de la mujer sobre si debe hacerse o no una mamografía. Las recomendaciones del panel fueron rechazadas por el National Cancer Institute, que era el organismo que había solicitado inicialmente la creación del panel, así como por otros organismos. Había indicios claros de la actuación de fuertes fuerzas políticas en ese momento a favor de la realización de mamografías en mujeres de 40-49 años.

La controversia sobre la mamografía aumentó aún más con la publicación en 2001 de una revisión de Olsen y Gøtzsche sobre la evidencia que respaldaba la mamografía *a cualquier edad*[10]. Entre las cuestiones planteadas por los investigadores estaban las preocupaciones acerca de la posible deficiencia de la asignación aleatoria, la posible falta de fiabilidad de la evaluación de la causa del fallecimiento, su constatación de que, en algunos ensayos, las exclusiones de las mujeres de los estudios se llevaron a cabo después de haber realizado la asignación aleatoria y que las mujeres con cáncer preexistente fueron excluidas únicamente de los grupos de cribado, así como el hecho de que en los dos mejores ensayos no encontraron ningún beneficio.

Un editorial acompañante de la revista *Lancet* concluyó diciendo lo siguiente: «En la actualidad, no existe ninguna evidencia fiable procedente de ensayos aleatorizados a gran escala que apoye los programas de mamografías

de cribado»[11]. Un artículo de 2004 contradijo los argumentos planteados por Olsen y Gøtzsche y llegó a la conclusión de que el consenso previo sobre la mamografía era correcto[12].

Sin embargo, la controversia se mantiene activa. En 2002, el Grupo de Trabajo de los Servicios Preventivos de EE. UU. revisó la evidencia y recomendó la mamografía de cribado cada 1-2 años para las mujeres a partir de 40 años de edad. Utilizando una versión anterior de la metodología que se describe en el capítulo 14, clasificaron la evidencia a favor como «regular» en una escala de «buena», «regular» o «mala»[13]. En 2009, este grupo de trabajo volvió a revisar la cuestión de la mamografía para mujeres de 40- 49 años y recomendó que las mujeres de 50-74 años deberían realizarse una mamografía de cribado cada 2 años. Sin embargo, también llegaron a la siguiente conclusión: «Para la mamografía de cribado bienal en mujeres de 40-49 años, existe una certeza moderada de que el beneficio neto es pequeño». El Grupo de Trabajo dio a su recomendación un grado «C» y señaló que esta calificación es una recomendación en contra del cribado rutinario de las mujeres de 40-49 años. Añadieron lo siguiente: «El Grupo de Trabajo fomenta la toma de decisiones individualizada e informada sobre el momento en el que iniciar el cribado mamográfico (a qué edad)»[14]. El grado «C» se confirmó en una declaración de recomendación más reciente del grupo de trabajo[15].

En 2007, el American College of Physicians publicó nuevas directrices sobre la mamografía para mujeres de 40-49 años basadas en una extensa revisión sistemática que evaluó los beneficios y los perjuicios potenciales[16,17]. El grupo concluyó que la evidencia de beneficio neto es menos clara para las mujeres de 40-49 años que para las de 50-59 años y que la mamografía conlleva riesgos significativos: «Nosotros no creemos que la evidencia respalde una recomendación general». En 2011, el National Health Service de Reino Unido publicó sus directrices, donde se recomienda que las mujeres de 47-73 años se realicen una mamografía cada 3 años[18].

En 2015, la American Cancer Society (ACS) actualizó sus pautas para el cribado del cáncer de mama en mujeres con «riesgo promedio»[19]. La ACS recomendó comenzar el cribado a los 45 años, con un cribado anual hasta los 54 años, después de lo cual debería considerarse el cribado bienal. Como queda claro, ¡este no es un campo donde la ciencia, la epidemiología y las políticas públicas estén totalmente alineadas!

Por tanto, la controversia entre los partidarios y los críticos continúa, y no es probable que las declaraciones de los expertos satisfagan a todo el mundo. Los problemas de metodología y de interpretación son complejos y es probable que no se puedan resolver mediante otros ensayos a gran escala. Tales ensayos son difíciles y costosos de iniciar y de llevar a cabo. Además, debido al tiempo necesario para su realización, estos ensayos también se ven limitados por el hecho de que los resultados a menudo no reflejan los avances más recientes en la tecnología mamográfica. Sin embargo, debido a que gran parte de los datos son equívocos y motivo de controversia, es muy probable

que el progreso provenga de las nuevas tecnologías para la detección del cáncer de mama. Mientras tanto, se deja a las mujeres con la dificultad de la toma de decisiones respecto a sus propias opciones sobre la mamografía, dadas las grandes incertidumbres presentes en la evidencia disponible.

Cribado del cáncer de cuello uterino

Quizás la prueba de cribado para el cáncer que más se ha utilizado históricamente sea el frotis de Papanicolaou. Por tanto, es de suponer que se ha obtenido una evidencia abrumadora de su eficacia en la reducción de la mortalidad por cáncer de cuello uterino invasivo. Por desgracia, nunca se ha realizado un ensayo controlado y aleatorizado diseñado de forma adecuada sobre el cribado del cáncer de cuello uterino y es probable que nunca se realice, porque se ha aceptado, tanto por las autoridades sanitarias como por las mujeres, que este cribado es eficaz para la detección precoz del cáncer de cuello uterino.

En ausencia de ensayos aleatorizados, se han utilizado varias estrategias alternativas. Tal vez el diseño más frecuente para la evaluación haya sido comparar la incidencia y mortalidad en poblaciones con diferentes tasas de cribado. Una segunda estrategia ha consistido en analizar los cambios a lo largo del tiempo de las tasas de diagnóstico de carcinoma in situ. Una tercera estrategia ha sido la de los estudios de casos y controles en los que las mujeres con cáncer invasivo del cuello uterino se comparan con las mujeres del grupo control y se analiza la frecuencia de las pruebas de Papanicolaou previas en ambos grupos. Todos estos estudios suelen verse afectados por los problemas metodológicos planteados anteriormente en este capítulo. Dado el reconocimiento de que el VPH participa en la etiología del cáncer cervical, la prevención actualmente recomienda la prueba del VPH junto con el frotis de Papanicolaou. La ACS recomienda comenzar el cribado a los 21 años con pruebas anuales de Papanicolaou (ya sea con citología convencional o con base líquida) e iniciar las pruebas de VPH a los 30 años o bien únicamente el uso de pruebas de detección de VPH de alto riesgo[20]. Sin embargo, incluso para los tipos de VPH de alto riesgo, la infección puede desaparecer después de la detección aislada del VPH, lo que resulta en un gran número de falsos positivos. En consecuencia, el US Preventive Services Task Force recomienda realizar pruebas de detección selectiva en mujeres de 21 a 65 años con citología (frotis de Papanicolaou) cada 3 años o, en las mujeres de 30 a 65 años que desean un cribado menos frecuente, citología combinada con pruebas para VPH cada 5 años[4].

A pesar de estas reservas, la evidencia indica que muchos de los carcinomas in situ probablemente progresan a cáncer invasivo, por lo que la detección precoz del cáncer de cuello uterino en el estadio in situ se traduciría en un número considerable de vidas salvadas, aunque sea menor que muchas estimaciones optimistas. Gran parte de la incertidumbre a la que nos enfrentamos con respecto a la detección del cáncer de cuello uterino se debe a que inicialmente no se realizó ningún ensayo aleatorizado bien diseñado

antes de incorporarse de modo rutinario a la práctica médica. Esta observación pone de relieve que, en Estados Unidos, se debe cumplir un conjunto de normas antes de que se autoricen nuevos agentes farmacológicos para su uso humano, pero se usa otro conjunto de normas menos estrictas para las nuevas tecnologías o nuevos programas de salud. Ningún fármaco se autorizaría en Estados Unidos sin una evaluación mediante ensayos controlados aleatorizados, pero, por desgracia, no se requiere tal evaluación antes de introducir programas y procedimientos de cribado o de otro tipo. Por supuesto, si se aplicara la prevención universal del VPH a través de la vacunación de adolescentes antes de comenzar a mantener relaciones sexuales, ¡el cáncer cervical desaparecería!

Cribado del neuroblastoma

Algunos de los problemas que acaban de comentarse se encuentran en el cribado del neuroblastoma, un tumor que aparece en niños de corta edad. Tuchman y cols.[21] esbozaron la justificación para el cribado del neuroblastoma: 1) El resultado ha mejorado poco en las últimas décadas. 2) Se sabe que el pronóstico es mejor en los niños que manifiestan la enfermedad antes del año de edad. 3) A cualquier edad, los niños con estadios avanzados de la enfermedad tienen un pronóstico peor que los de los estadios precoces. 4) Más del 90% de los niños que presentan síntomas clínicos de neuroblastoma excretan cantidades mayores de lo normal de catecolaminas en la orina. 5) Estos metabolitos pueden medirse con facilidad en las muestras de orina obtenidas de los pañales.

Estos hechos constituyen una justificación sólida para el cribado del neuroblastoma. En la figura 18.21 se muestran los datos de Japón, donde se había realizado un esfuerzo significativo para el cribado del neuroblastoma. Los porcentajes de niños menores de 1 año en los que se detectó el neuroblastoma se compararon antes y después del inicio del cribado en Sapporo, una ciudad de Hokkaido, y estos datos se compararon con los datos de nacimientos del resto de Hokkaido, donde no se estableció un programa de cribado. Después del inicio del cribado, se detectó un mayor porcentaje de casos de neuroblastoma en niños menores de 1 año en Sapporo que en el resto de Hokkaido.

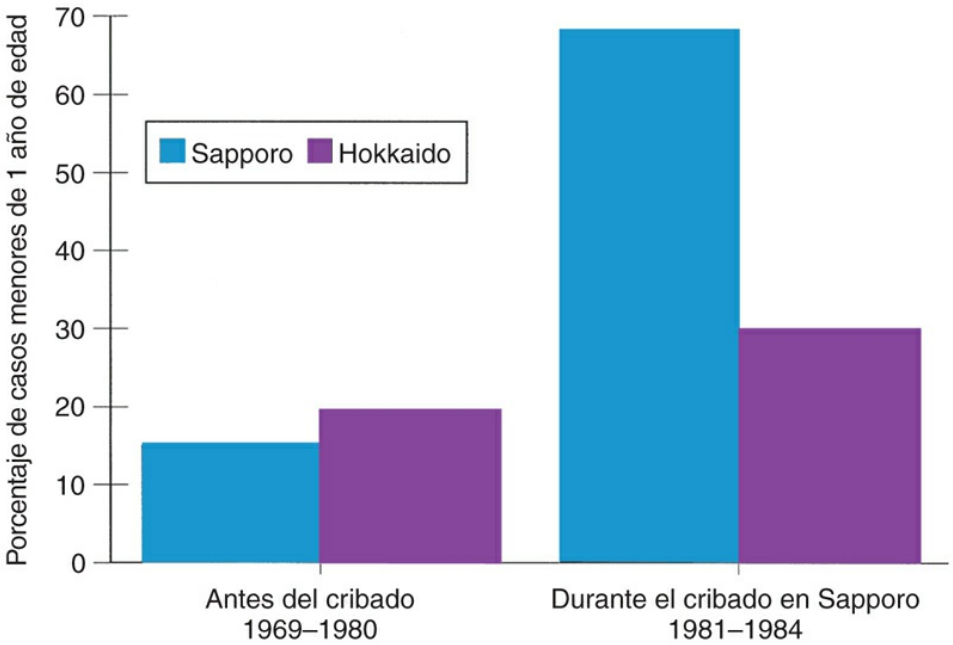


FIG. 18.21 Porcentaje de casos de neuroblastoma menores de 1 año de edad en Sapporo y Hokkaido, Japón, antes y después del cribado. (Modificada de Goodman SN. Neuroblastoma screening data: An epidemiologic analysis. *Am J Dis Child.* 1991;145:1415–1422; basada en datos de Nishi M, Miyake H, Takeda T, et al. Effects of the mass screening of neuroblastoma in Sapporo City. *Cancer.* 1987;60:433–436. Copyright © 1987 American Cancer Society. Reproducida con autorización de Wiley-Liss, Inc., una filial de John Wiley & Sons, Inc.)

Sin embargo, surgen varios problemas graves a la hora de evaluar los beneficios del cribado del neuroblastoma. Ahora está claro que el neuroblastoma es una enfermedad con heterogeneidad biológica y algunos casos tienen un pronóstico mucho mejor desde el principio que otros. Muchos tumores tienen un buen pronóstico porque remiten espontáneamente, incluso sin tratamiento. Además, es más probable que el cribado detecte los tumores de crecimiento lento, menos malignos, y es menos probable que detecte las neoplasias agresivas de crecimiento rápido.

Así pues, es difícil demostrar que el cribado de los neuroblastomas es en realidad beneficioso. De hecho, dos estudios a gran escala sobre el cribado del neuroblastoma se publicaron en 2002. Woods y cols.[22] estudiaron a 476.654 niños en Quebec, Canadá. El cribado se ofreció a todos los niños de entre 3 semanas y 6 meses de edad. La mortalidad por neuroblastoma hasta los 8 años de edad en los niños sometidos a cribado en Quebec no fue menor que en las cuatro cohortes sin cribado (tabla 18.1) ni menor que en el resto de Canadá, excepto en Quebec y en dos cohortes históricas (tabla 18.2). Schilling y cols.[23] estudiaron a 2.581.188 niños en Alemania a los que se ofreció el cribado al año de edad. Estos autores observaron que el cribado del neuroblastoma no redujo la incidencia de la enfermedad diseminada y no pareció reducir la mortalidad por la enfermedad, aunque el seguimiento de la mortalidad aún no se había completado. Los datos disponibles en la actualidad no respaldan, por tanto, el cribado del neuroblastoma. Los resultados de estos estudios demuestran la importancia de la comprensión de

la biología y la historia natural de la enfermedad y la necesidad de obtener pruebas relevantes y rigurosas sobre los posibles beneficios o la falta de beneficios cuando se está planteando el cribado de cualquier enfermedad. La capacidad de detectar una enfermedad mediante el cribado no se puede equiparar con una demostración de que las personas sometidas a cribado obtengan beneficios.

Tabla 18.1

Mortalidad por neuroblastoma a los 8 años de edad en la cohorte de Quebec sometida a cribado, en comparación con las tasas en cuatro cohortes sin cribado

Cohorte control	N.° de fallecimientos esperados en Quebec basándose en la cohorte control	Tasa de mortalidad estandarizada de Quebec (IC del 95%)
Ontario	19,8	1,11 (0,64-1,92)
Minnesota	24,4	0,90 (0,48-1,70)
Florida	15,7	1,40 (0,81-2,41)
Greater Delaware Valley	22,8	0,96 (0,56-1,66)

Hubo 22 fallecimientos por neuroblastoma en la cohorte de Quebec sometida a cribado.

IC, intervalo de confianza.

De Woods WG, Gao R, Shuster JJ, et al. Screening of infants and mortality due to neuroblastoma. *N Engl J Med.* 2002;346:1041–1046.

Tabla 18.2

Mortalidad por neuroblastoma a los 8 años de edad en la cohorte de Quebec sometida a cribado, en comparación con las tasas en cohortes canadienses sin cribado

Cohorte control	N.° de fallecimientos esperados en Quebec basándose en la cohorte control	Tasa de mortalidad estandarizada de Quebec (IC del 95%)
Cohortes históricas		
Quebec	22,5	0,98 (0,54-1,77)
Canadá	21,2	1,04 (0,64-1,69)
Cohorte concurrente		
Canadá, salvo Quebec	15,8	1,39 (0,85-2,30)

Hubo 22 fallecimientos por neuroblastoma en la cohorte sometida a cribado. Todos los datos fueron recogidos por Statistics Canada.

IC, intervalo de confianza.

De Woods WG, Gao R, Shuster JJ, et al. Screening of infants and mortality due to neuroblastoma. *N Engl J Med.* 2002;346:1041–1046.

Problemas en la evaluación de la sensibilidad y la especificidad de las pruebas de cribado

Los nuevos programas de cribado suelen iniciarse después de que una prueba de cribado esté disponible por primera vez. Cuando se desarrolla una prueba de este tipo, los fabricantes de kits de pruebas, los investigadores u otras personas suelen afirmar que la prueba tiene una sensibilidad y una especificidad elevadas. Sin embargo, como se verá, desde un punto de vista práctico, esto a menudo puede ser difícil de demostrar.

En la figura 18.22A se muestra una tabla de 2 × 2, similar a la de capítulos anteriores, donde se representa la realidad (enfermedad presente o ausente) frente a los resultados de la prueba (positiva o negativa).

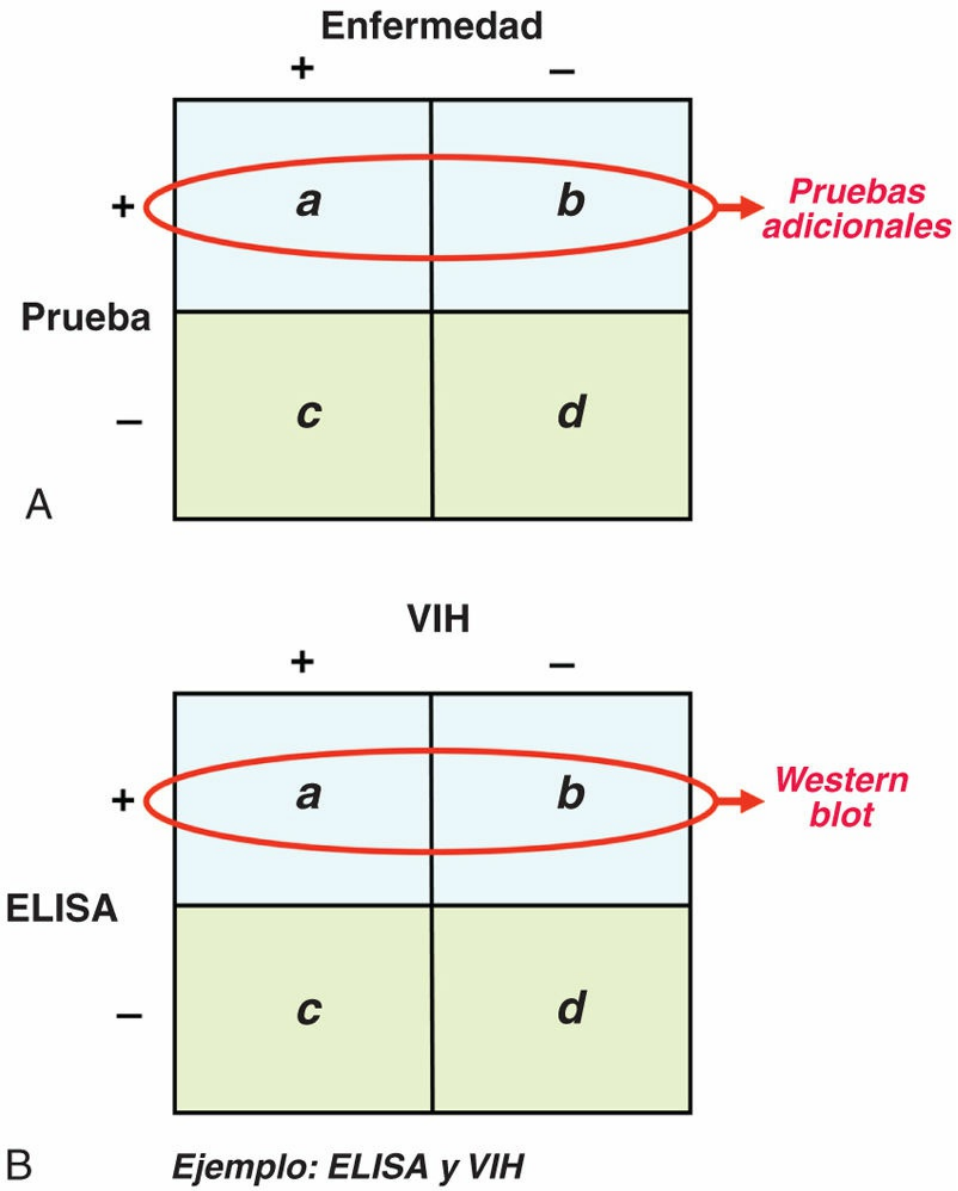


FIG. 18.22 (A) Problema a la hora de determinar la sensibilidad y especificidad debido al seguimiento limitado de las personas con resultados de la prueba negativos. (B) Problema a la hora de determinar la sensibilidad y especificidad debido al seguimiento limitado de las personas con resultados de la prueba negativos para el virus de la inmunodeficiencia humana *(VIH)* mediante el análisis de inmunoabsorción ligada a enzimas *(ELISA)*.

Para calcular la sensibilidad y la especificidad, se necesitan datos en las cuatro celdas. Sin embargo, a menudo solo las personas con resultados positivos de la prueba ($a + b$) (*fila superior* de la figura) son remitidas para pruebas adicionales. Los datos de quienes tuvieron una prueba negativa ($c + d$) a menudo no están disponibles, ya que estos pacientes no se someten a más pruebas. Por ejemplo, como se muestra en la figura 18.22B, la prueba de Western blot sirve como patrón oro para la detección del virus de la inmunodeficiencia humana (VIH), y las personas con un resultado positivo en un análisis de inmunoabsorción ligada a enzimas (ELISA) se remiten para realizar una prueba de Western blot.

Sin embargo, debido a que las personas con resultados negativos en la

prueba ELISA no suelen someterse a más pruebas, los datos necesarios en las celdas inferiores para el cálculo de la sensibilidad y la especificidad de la prueba ELISA a menudo no están disponibles a partir de las pruebas de rutina. Para obtener estos datos, es esencial que se remitan algunas muestras negativas de ELISA para su análisis posterior, junto con las muestras ELISA-positivas.

Interpretación de los resultados de estudios que no muestran beneficios del cribado

En este capítulo, hasta ahora se ha hecho hincapié en la interpretación de los resultados que muestran una diferencia entre los grupos con y sin cribado. Sin embargo, si no se logra demostrar un beneficio de la detección precoz de la enfermedad, cualquiera de las siguientes interpretaciones puede ser posible:

1. La falta aparente de beneficio puede ser inherente a la historia natural de la enfermedad (p. ej., la enfermedad no tiene fase preclínica detectable o dicha fase es muy corta).
2. La intervención terapéutica disponible en la actualidad puede que no sea más eficaz cuando se proporciona antes que cuando se proporciona en el momento del diagnóstico habitual.
3. La historia natural y los tratamientos disponibles en la actualidad pueden ser capaces de aumentar el beneficio, pero las deficiencias de la asistencia prestada a las personas con un resultado positivo en el cribado pueden explicar la falta observada de beneficios (es decir, hay eficacia, pero poca efectividad).

Análisis de rentabilidad del cribado

Algunas personas responden a las cuestiones de rentabilidad centrándose solo en el coste preguntando: si la prueba es barata, ¿por qué no llevarla a cabo? Sin embargo, aunque la prueba de sangre oculta en heces, por ejemplo, para el cribado del cáncer de colon, cuesta muy poco en cuanto al kit de papel de filtro y el procesamiento de laboratorio necesarios, a la hora de calcular el coste total de una prueba de este tipo se debe incluir el coste de las colonoscopias que se realizan después de la prueba inicial en aquellos detectados como «positivos», así como el de las complicaciones infrecuentes de la colonoscopia.

El balance de la rentabilidad no solo incluye los costes económicos, sino también los costes no económicos para el paciente, como la ansiedad, la angustia emocional y las molestias. Aunque la propia prueba no sea invasiva, si el resultado es positivo, ¿se requiere un tratamiento invasivo debido al resultado de la prueba? ¿Cuál es la tasa de falsos positivos de este tipo de pruebas? ¿En qué proporción de personas se realizarán pruebas invasivas o se generará ansiedad a pesar de que realmente no tienen la enfermedad en cuestión? Por tanto, el «coste» de una prueba no solo es el coste de su procedimiento, sino también el de todo el proceso de seguimiento que se pone en marcha por un resultado positivo, incluso si resulta ser un falso positivo. Estas consideraciones se reflejan en las cuatro preocupaciones principales expresadas por la ACS en la revisión de sus directrices para el cribado del cáncer, publicada en 2000 (cuadro 18.3)[24].

Cuadro 18.3 Criterios utilizados por la American Cancer Society para las recomendaciones sobre chequeos relacionados con el cáncer

1. Debe existir una buena evidencia de que cada prueba o procedimiento recomendado es eficaz desde el punto de vista médico para la reducción de la morbilidad o la mortalidad.
2. Los beneficios médicos deben ser mayores que los riesgos.
3. El coste de cada prueba o procedimiento debe ser razonable en comparación con sus beneficios esperados.
4. Las medidas recomendadas deben ser prácticas y factibles.

Elmore y Choe[25] plantearon otro punto de vista sobre la rentabilidad. En un artículo sobre la mamografía de cribado para mujeres de 40-49 años, escribieron lo siguiente:

A continuación, planteamos una manera de explicar la evidencia (con la salvedad de que los números están redondeados y simplificados): de cada 10.000 mujeres en las que se realiza una mamografía de cribado periódica comenzando a los 40 años, 6 de ellas podrían beneficiarse gracias a una disminución del riesgo de mortalidad por cáncer de mama. Sin embargo, incluso este modesto beneficio requiere múltiples exploraciones de cribado y seguimiento de todas las 10.000 mujeres durante más de una década. Dicho de otro modo, 9.994 mujeres no obtienen ningún beneficio de mortalidad en absoluto, porque la mayoría de ellas no desarrollarán cáncer de mama y en algunas el cáncer se detectará cuando ya sea demasiado tarde para su curación[25]*.*

Conclusión

En este capítulo se han revisado algunas de las principales fuentes de sesgo que deben tenerse en cuenta a la hora de evaluar los resultados de los estudios que comparan una población en la que se realiza un cribado con otra en la que este no se lleva a cabo. Los sesgos de selección para el cribado y la selección por pronóstico pueden resolverse, en gran parte, usando un diseño de estudio de tipo ensayo controlado y aleatorizado. Se pueden realizar unas estimaciones razonables sobre el tiempo de adelanto si se dispone de la información adecuada. Pocos de los métodos que se utilizan actualmente para la detección precoz de enfermedades se han evaluado mediante ensayos aleatorizados, y es probable que la mayoría de ellos no se vayan a estudiar de esta manera. Esto se debe a varios factores, como la dificultad y el coste asociado a la realización de tales estudios, así como a los problemas éticos inherentes a la asignación aleatoria de una población para recibir o no modalidades de asistencia que se utilizan ampliamente y que se consideran eficaces, aunque no cuenten con una evidencia sólida que las respalde. Por tanto, estamos obligados a maximizar nuestro uso de la evidencia procedente de estrategias no aleatorizadas y, para ello, se deben tener en cuenta los posibles sesgos y problemas comentados en este capítulo.

Al analizar los programas de detección precoz de enfermedades, hay que ser capaces de identificar los grupos que están en una situación de alto riesgo. Esto incluiría no solo a los que están en riesgo de desarrollar la enfermedad en cuestión, sino también a los que están «en riesgo» de beneficiarse de la intervención. Estos son los grupos en los que los cálculos de rentabilidad resultarán beneficiosos. Se debe tener en cuenta que, aunque la prueba de cribado, como el frotis de Papanicolaou, no sea demasiado invasiva en sí misma, la intervención necesaria ante un resultado positivo de la prueba de cribado puede ser muy invasiva.

La cuestión fundamental es cómo tomar decisiones cuando los datos son no concluyentes, discordantes o incompletos. Este dilema se plantea con frecuencia, tanto en la práctica clínica como en el desarrollo de políticas de salud pública. Estas decisiones deben tener en cuenta en primer lugar el corpus existente de evidencia científica relevante. Sin embargo, en el análisis final, la decisión de realizar o no el cribado de una enfermedad en una población es un juicio de valor que debe tener en cuenta la incidencia y la gravedad de la enfermedad, la viabilidad de su detección precoz, la posibilidad de intervenir con eficacia en las personas con resultados positivos en el cribado y el cálculo de rentabilidad global de un programa de detección precoz.

Para mejorar nuestra capacidad de tomar las decisiones adecuadas, se necesitan conocimientos adicionales sobre la historia natural de la enfermedad y, en concreto, sobre la definición de las características de los individuos que están en riesgo de un mal resultado. Antes de la introducción

de nuevos programas de cribado, debemos solicitar enérgicamente que se lleven a cabo ensayos controlados aleatorizados bien realizados, para no tener que actuar en un entorno de incertidumbre en el futuro, cuando estos ensayos ya sean casi imposibles de efectuar. Sin embargo, debido a que la mayor parte de la práctica médica y de salud pública (incluida la detección temprana de enfermedades) no se ha sometido a los ensayos aleatorizados, y a que las decisiones sobre la detección temprana se deben hacer basándose en datos incompletos y ambiguos, es esencial que nosotros, como profesionales de la salud, advirtamos y conozcamos los problemas metodológicos implicados para que podamos hacer el mejor uso de los conocimientos disponibles en beneficio de nuestros pacientes. Incluso las mejores intenciones y el proselitismo apasionado no pueden sustituir a la evidencia rigurosa que respalda o no respalda el beneficio del cribado.

Bibliografía

1. Whittier JG. from Maud Muller:The Panorama, and Other Poems. Boston: Ticknor and Fields; 1856.
2. Harte B, from *"Mrs. Judge Jenkins: Sequel to Maud Muller" East and West Poems.* Boston, James R. Osgood and Company, 1871.
3. Hutchison GB. Evaluation of preventive services. *J Chronic Dis*. 1960;11:497–508.
4. Moyer VA, on behalf of the U.S. Preventive Services Task Force. Screening for lung cancer: U.S. Preventive Services Task Force Recommendation Statement. Ann Intern Med. 2014;160:330–338.
5. Szklo M, Nieto FJ. Epidemiology: Beyond the Basics. 3rd ed. Burlington, MA: Jones & Bartlett; 2014: 141–145.
6. Shapiro S, Venet W, Strax P, eds. Periodic Screening for Breast Cancer: The Health Insurance Plan Project and Its Sequelae, 1963-1986. Baltimore: Johns Hopkins University Press; 1988.
7. Shapiro S, Venet W, Strax P, et al. Prospects for eliminating racial differences in breast cancer survival rates. *Am J Public Health*. 1982;72:1142–1145.
8. *Breast Cancer Screening for Women Ages 40-49.* NIH Consensus Statement Online, 1997 January 21-23, cited 15:1–35, 2007.
9. Wegwarth O, Schwartz LM, Woloshin S, et al. Do physicians understand cancer screening statistics? A national survey of primary care physicians in the United States. *Ann Intern Med*. 2012;156:340–349.
10. Olsen O, Gøtzsche C. Cochrane review on screening for breast cancer with mammography. *Lancet*. 2001;358:1340–1342.
11. Horton R. Screening mammography: an overview revisited. *Lancet*. 2001;358:1284–1285.
12. Freedman DA, Petitti DB, Robins JM. On the efficacy of screening for breast cancer. *Int J Epidemiol*. 2004;33:43–55.
13. U.S. Preventive Services Task Force Breast cancer screening: a summary of the evidence for the U.S. Preventive Services Task Force. *Ann Intern Med*. 2002;137:347–360.
14. U.S. Preventive Services Task Force. Breast cancer: screening; July 2010 [fecha de última consulta: 14 de junio de 2013]. http://www.uspreventiveservicestaskforce.org/uspstf09/breast

15. Siu AL. on behalf of the U.S. Preventive Services Task Force Screening for breast cancer: US Preventive Services Task Force recommendation statement. *Ann Intern Med*. 2016;164:279–296.
16. Brewer NT, Salz T, Lillie SE. Systematic review: the long-term effects of false-positive mammograms. *Ann Intern Med*. 2007;146:502–510.
17. Qaseem A, Snow V, Sherif K, et al. Screening mammography for women 40 to 49 years of age: a clinical practice guideline from the American College of Physicians. *Ann Intern Med*. 2007;146:511–515.
18. Warner E. Breast-cancer screening. *N Engl J Med*. 2011;365:1025–1032.
19. Oeffinger KC, Fontham ETH, Etzioni R, et al. Breast cancer screening for women at average risk. 2015 Guideline update from the American Cancer Society. *JAMA*. 2015;314(15):1599–1614.
20. American Cancer Society. The American Cancer Society Guidelines for the Prevention and Early Detection of Cervical Cancer [fecha de última consulta: 20 de febrero de 2018. https://www.cancer.org/cancer/cervical-cancer/prevention-and-early-detection/cervical-cancer-screening-guidelines.html.
21. Tuchman M, Lemieux B, Woods WG. Screening for neuroblastoma in infants: investigate or implement? *Pediatrics*. 1990;86:791–793.
22. Woods WG, Gao R, Shuster JJ, et al. Screening of infants and mortality due to neuroblastoma. *N Engl J Med*. 2002;346:1041–1046.
23. Schilling FH, Spix C, Berthold F, et al. Neuroblastoma screening at one year of age. *N Engl J Med*. 2002;346:1047–1053.
24. Smith RA, Mettlin CJ, David KJ, et al. American Cancer Society guidelines for the early detection of cancer. *CA Cancer J Clin*. 2000;50:34–49.
25. Elmore JG, Choe JH. Breast cancer screening for women in their 40s: moving from controversy about data to helping individual women. *Ann Intern Med*. 2007;146:529–531.

Preguntas de repaso del capítulo 18

Las preguntas 1 a 4 se basan en la siguiente información:

Un nuevo programa de cribado se instauró en un determinado país. El programa utilizaba una prueba de cribado que es eficaz para detectar el cáncer Z en un estadio precoz. Supongamos que no existe ningún tratamiento eficaz para este tipo de cáncer y, por tanto, que los resultados del programa no suponen ningún cambio en la evolución normal de la enfermedad. Supongamos también que las tasas observadas se calculan a partir de todos los casos conocidos de cáncer Z y que no hubo cambios en la calidad de la certificación de la muerte de esta enfermedad.

1. *¿Qué pasará con la incidencia aparente de cáncer Z en el país durante el primer año de este programa?*

a. La incidencia aumentará.

b. La incidencia disminuirá.

c. La incidencia se mantendrá constante.

2. *¿Qué pasará con la* prevalencia *aparente de cáncer Z en el país durante el primer año de este programa?*

a. La prevalencia aumentará.

b. La prevalencia disminuirá.

c. La prevalencia se mantendrá constante.

3. *¿Qué pasará con la* letalidad *aparente por cáncer Z en el país durante el primer año de este programa?*

a. La letalidad aumentará.

b. La letalidad disminuirá.

c. La letalidad se mantendrá constante.

4. *¿Qué pasará con la* mortalidad *aparente por cáncer Z en el país como resultado del programa?*

a. La mortalidad aumentará.

b. La mortalidad disminuirá.

c. La mortalidad se mantendrá constante.

5. *El mejor índice (índices) para concluir que un programa de detección precoz de cáncer de mama mejora verdaderamente la historia natural de la*

enfermedad, 15 años después de su inicio, sería:

a. Una mortalidad proporcional menor para el cáncer de mama 15 años después del inicio del programa de detección precoz en comparación con la mortalidad proporcional antes de su inicio.

b. La mejora de la supervivencia a largo plazo en las pacientes con cáncer de mama (ajustada por tiempo de adelanto).

c. Una disminución de la incidencia de cáncer de mama.

d. Una disminución de la prevalencia de cáncer de mama.

e. Ninguna de las anteriores.

6. *En general, el cribado debe llevarse a cabo para enfermedades con la característica o características siguientes:*

a. Enfermedades con una baja prevalencia en los subgrupos identificables de la población.

b. Enfermedades en las que la letalidad es baja.

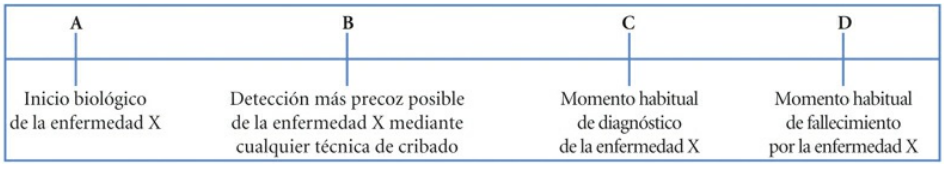
c. Enfermedades con una historia natural que puede modificarse con la intervención médica.

d. Enfermedades que se diagnostican fácilmente y para las que se ha demostrado que la eficacia del tratamiento es equívoca mediante la evidencia procedente de varios ensayos clínicos.

e. Ninguna de las anteriores.

La pregunta 7 se basa en la siguiente información:

El siguiente diagrama muestra la historia natural de la enfermedad X:



7. *Supongamos que la detección precoz de la enfermedad X mediante el cribado mejora el pronóstico. Para que un programa de cribado sea lo más eficaz posible, ¿en qué momento de la historia natural del diagrama debe estar el punto crítico?*

a. Entre A y B.

b. Entre B y C.

c. Entre C y D.

d. En cualquier punto entre A y C.

e. En cualquier punto entre A y D.

8. *¿Cuál de las siguientes no es una posible medida de resultado que podría utilizarse como indicador de los beneficios de los programas de cribado dirigidos a la detección precoz de una enfermedad?*

a. Reducción de la letalidad en las personas sometidas al cribado.

b. Reducción de la mortalidad en la población sometida al cribado.

c. Reducción de la incidencia en la población sometida al cribado.

d. Reducción de las complicaciones.

e. Mejora de la calidad de vida en las personas sometidas al cribado.

CAPÍTULO 19

Epidemiología y políticas de salud pública

Palabras clave

susceptibilidad
evaluación de riesgos
medición de la exposición
número necesario para someterse a la intervención para prevenir un caso/una muerte
revisiones sistemáticas y metaanálisis
sesgo de publicación
incertidumbre

Cualquier trabajo científico es incompleto, tanto si es observacional como experimental.

Cualquier trabajo científico es susceptible de poder ser rebatido o modificado por los nuevos conocimientos.

Esto no nos confiere la libertad de ignorar los conocimientos que ya tenemos ni de posponer la acción que parecen exigir en un momento dado[1].

—Sir Austin Bradford Hill, discurso del presidente, Royal Society of Medicine, 14 de enero de 1965

La experiencia es esa cosa maravillosa que te permite reconocer un error cuando vuelves a cometerlo[2].

—Franklin P. Jones, legendario humorista estadounidense (1908-1980)

Objetivos de aprendizaje

- Revisar el papel de la epidemiología en la prevención de enfermedades y contrastar dos posibles estrategias de prevención: dirigirse a grupos de alto riesgo de

presentar una enfermedad en comparación con centrarse en la población general.

- Definir la evaluación de riesgos y comentar el papel de la epidemiología en la evaluación de riesgos, incluida la medición de posibles exposiciones.
- Describir cómo puede emplearse la epidemiología para modelar las políticas de salud pública en Estados Unidos por vía judicial.
- Presentar la revisión sistemática y el metaanálisis como herramientas para resumir todas las evidencias epidemiológicas disponibles para influir sobre las políticas de salud pública y describir cómo el sesgo de publicación puede afectar a los resultados de las revisiones sistemáticas y los metaanálisis.
- Identificar algunas de las posibles fuentes de incertidumbre cuando se utilizan los resultados de los estudios epidemiológicos como base para elaborar políticas de salud pública.

Un papel importante de la epidemiología es el de servir como base para el desarrollo de políticas que influyan sobre la salud humana, como la prevención primaria y secundaria y el control de las enfermedades. Como se ha visto en capítulos anteriores, los resultados de los estudios epidemiológicos pueden ser relevantes tanto para los problemas de la práctica clínica y de la salud de la comunidad como para las estrategias poblacionales destinadas a la prevención de enfermedades y la promoción de la salud. Como se comentó en el capítulo 1, las aplicaciones prácticas de la epidemiología a menudo se ven como algo tan consustancial de la disciplina que están incorporadas en la propia definición de epidemiología. Desde el punto de vista histórico, las investigaciones epidemiológicas se iniciaron para hacer frente a los nuevos desafíos relacionados con las enfermedades humanas (más frecuente con las enfermedades infecciosas) y la salud pública. De hecho, una de las principales fuentes de motivación en la epidemiología es

la aplicabilidad directa de sus resultados para aliviar los problemas de la salud humana. En este capítulo se presenta una visión general de algunas cuestiones y problemas relacionados con la epidemiología a la hora de aplicarla para la formulación y evaluación de políticas de salud pública.

Epidemiología y prevención

La importancia de la epidemiología en la prevención se ha subrayado en varios de los capítulos anteriores. La identificación de las poblaciones con mayor riesgo, la determinación de la causa de su mayor riesgo y el análisis de los costes y beneficios de eliminar o reducir la exposición al factor o factores causales requieren una comprensión de los conceptos epidemiológicos básicos y de la posible interpretación de los resultados de los estudios epidemiológicos. Además, la evaluación de la solidez de todas las evidencias disponibles y la identificación de los posibles límites de las inferencias derivadas o de la generalización de los resultados tienen una importancia crucial. La epidemiología a menudo se considera la «ciencia básica» de la prevención.

La pregunta de cuántos datos epidemiológicos se necesitan para justificar un esfuerzo de prevención no tiene una respuesta sencilla. Algunos de los temas implicados difieren dependiendo de si se está estudiando la prevención primaria o secundaria. Si se habla de prevención primaria, la respuesta depende de la gravedad de la enfermedad, de los costes (en términos económicos, de sufrimiento humano y de pérdida de calidad de vida), de la fuerza de la evidencia que implica un determinado factor o factores causales en la etiología de la enfermedad en cuestión y de la dificultad de reducir o eliminar la exposición a ese factor.

Con la prevención secundaria, las cuestiones son algo diferentes. Todavía hay que considerar la gravedad de la enfermedad en cuestión. Sin embargo, hay que preguntarse, además, si es posible detectar la enfermedad antes de lo habitual mediante el cribado, qué grado de invasividad y qué coste tendría ese cribado. Otras consideraciones incluyen si una persona que tiene la enfermedad obtiene algún beneficio si el tratamiento se inicia en un estadio previo a lo habitual y si el cribado se asocia a efectos perjudiciales. La epidemiología es, sin duda, un método valioso para resolver muchas de estas cuestiones.

En los últimos años se ha dirigido gran atención a ampliar lo que se ha llamado el modelo epidemiológico tradicional de factor de riesgo, en el que se explora la relación entre un factor independiente (exposición) y un factor dependiente (resultado de la enfermedad) (fig. 19.1). Se ha sugerido que este enfoque debe ampliarse en dos direcciones: en primer lugar, debe incluir la medición no solo del resultado adverso (la propia enfermedad), sino también de los impactos económicos, sociales y psicológicos derivados del resultado de la enfermedad sobre el individuo, su familia y la comunidad en general; en segundo lugar, está claro que la exposición a un supuesto agente causal no suele distribuirse de manera uniforme en una población. Por tanto, se deben analizar los factores que determinan si una persona queda expuesta si la prevención se va a basar en la reducción de la exposición (fig. 19.2). El modelo completo es aún más complejo, como se ve en la figura 19.3: la

relación está influenciada por *determinantes de la susceptibilidad* del individuo a la exposición, como factores genéticos, junto con las influencias ambientales y determinantes sociales. Aunque este enfoque ampliado es intuitivamente atractivo y proporciona un marco de referencia excelente para analizar los problemas de salud pública, aún se debe demostrar si ciertas exposiciones u otras variables independientes se asocian con un mayor riesgo de enfermedades específicas.

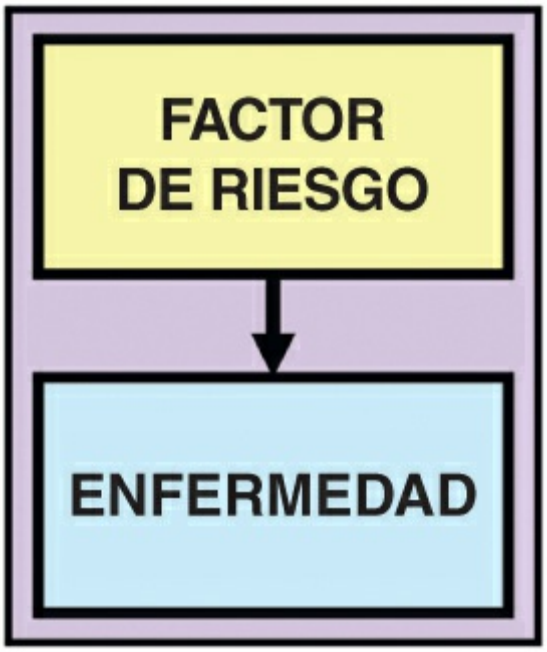


FIG. 19.1 Diagrama de la epidemiología clásica de los factores de riesgo.

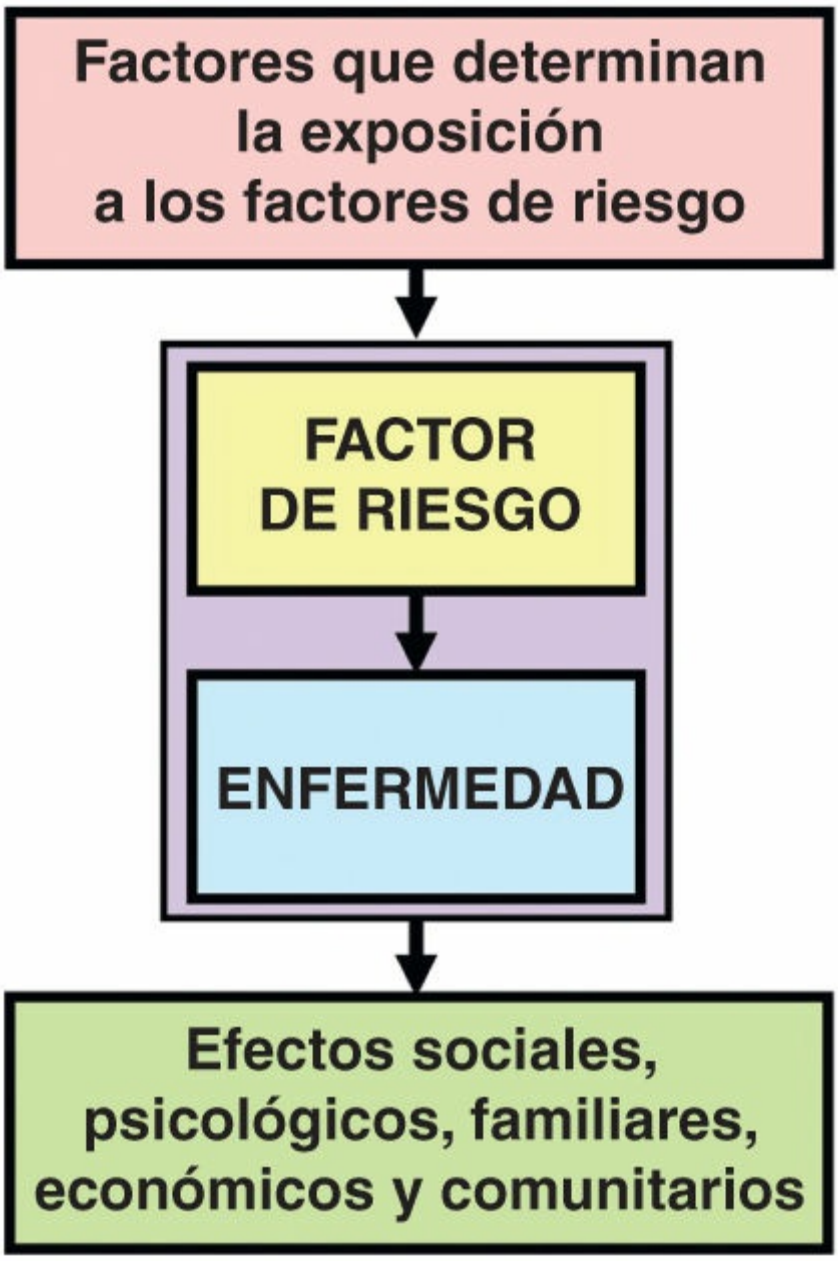


FIG. 19.2 Diagrama del modelo epidemiológico ampliado de los factores de riesgo para incluir determinantes de la exposición, así como los efectos sociales, familiares, psicológicos, económicos y comunitarios de la enfermedad.

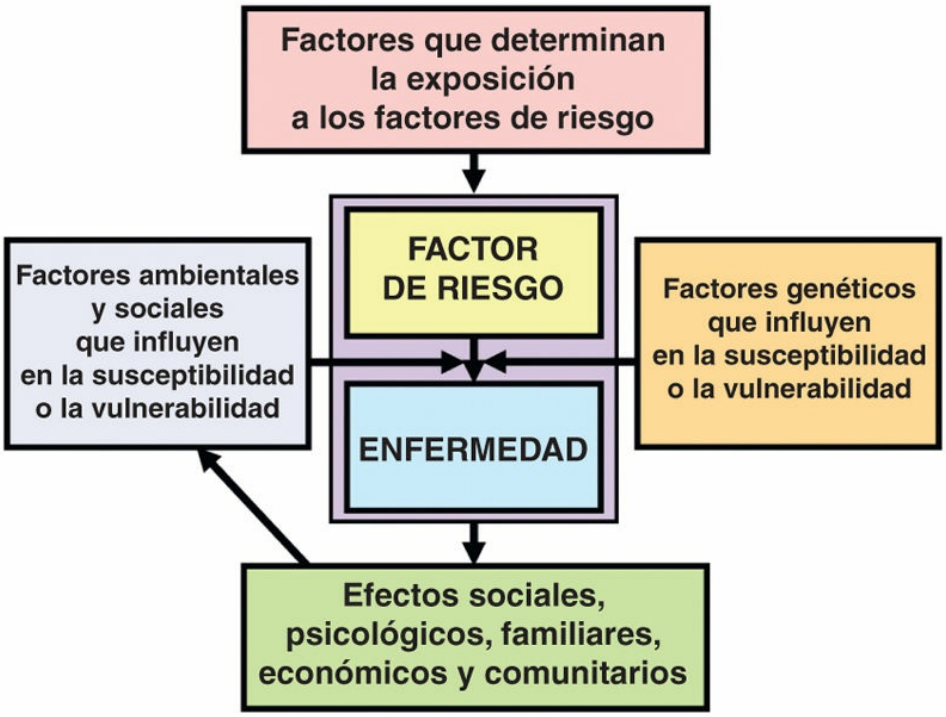


FIG. 19.3 Diagrama del modelo epidemiológico ampliado de los factores de riesgo para incluir las interrelaciones de los factores que determinan la susceptibilidad o la vulnerabilidad.

En cualquier caso, decidir la cantidad y el tipo de datos que se requieren para la prevención estará determinado por la sociedad y reflejará sus valores y prioridades. La epidemiología, junto con otras disciplinas, puede proporcionar muchos de los datos científicos necesarios para abordar cuestiones de riesgos y prevención. Sin embargo, la decisión final en cuanto a si se debe iniciar o mantener un programa de prevención estará determinada en gran medida por consideraciones económicas y políticas, así como por los valores sociales. Al mismo tiempo, es de esperar que tales decisiones también se cimentarán en una sólida base de evidencia científica proporcionada por la epidemiología y otras disciplinas relevantes.

Es importante distinguir entre las exposiciones *macroambientales* y *microambientales.* Las exposiciones macroambientales son las que afectan a poblaciones o comunidades enteras, como la contaminación atmosférica. Las exposiciones microambientales son los factores ambientales que afectan a un individuo específico, como la dieta (y la disponibilidad de alimentos sanos), el hábito tabáquico (del propio individuo o como fumador pasivo) y el consumo de alcohol (individualmente y la disponibilidad de alcohol en la comunidad). Desde el punto de vista de la prevención, los factores macroambientales son más fáciles de controlar y modificar en muchos aspectos, ya que esto se puede lograr mediante la legislación y las normativas (p. ej., estableciendo normas medioambientales para los contaminantes). Por el contrario, la modificación de factores microambientales depende de la modificación de los hábitos y del estilo de vida individuales, y del abordaje de la disponibilidad de alimentos sanos, zonas verdes y vecindarios seguros, lo que a menudo plantea dificultades mucho mayores.

Al abordar los factores microambientales no suele bastar con proporcionar la evidencia científica y las estimaciones de riesgo para inducir a los individuos a modificar sus estilos de vida (p.ej., el abandono del tabaquismo). Las personas a menudo difieren en el grado en que están dispuestas a asumir riesgos en muchos aspectos de sus vidas, incluida la salud. Además, las conductas individuales pueden ser diferentes en función de si se enfrentan con el riesgo de un resultado adverso o con la probabilidad de un evento positivo (fig. 19.4). Además, las personas a menudo pueden señalar a cualquier otro factor a la hora de atribuir la culpa de los problemas de salud provocados por su propio estilo de vida. Por tanto, la comunicación de riesgos, que se ha mencionado anteriormente, debe ir más allá de la comunicación de datos sobre los riesgos a los gestores. También debe ocuparse de la comunicación con la opinión pública de una manera comprensible en el contexto de la percepción que tiene la gente de su riesgo, para que los individuos se sientan motivados a aceptar la responsabilidad y actúen en beneficio de su propia salud en la mayor medida posible. Los epidemiólogos deben trabajar, por tanto, con los divulgadores sanitarios para lograr una educación más apropiada de la opinión pública en lo que respecta a cuestiones de riesgos personales.

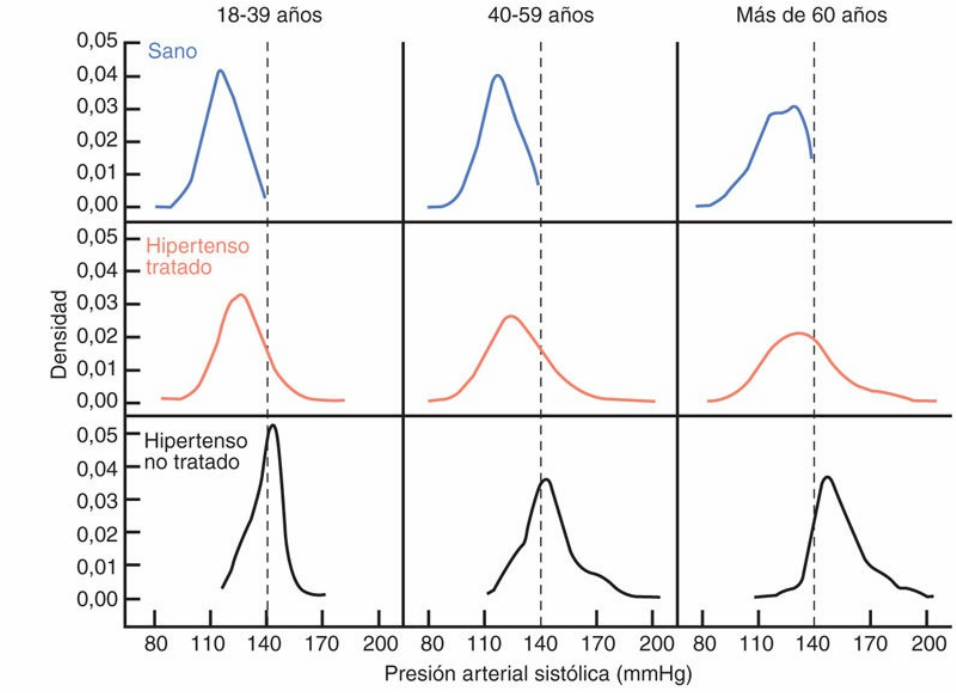


FIG. 19.4 «Eso nunca me pasará a mí.» *versus* «Hoy puede ser mi día de suerte.» Cómo el criterio de valoración puede influir en la percepción del riesgo y la voluntad de actuar de una persona. (Steve Kelley. © 1998 San Diego Union Tribune. Copley News Service.)

Estrategias poblacionales frente a estrategias de alto riesgo para la prevención

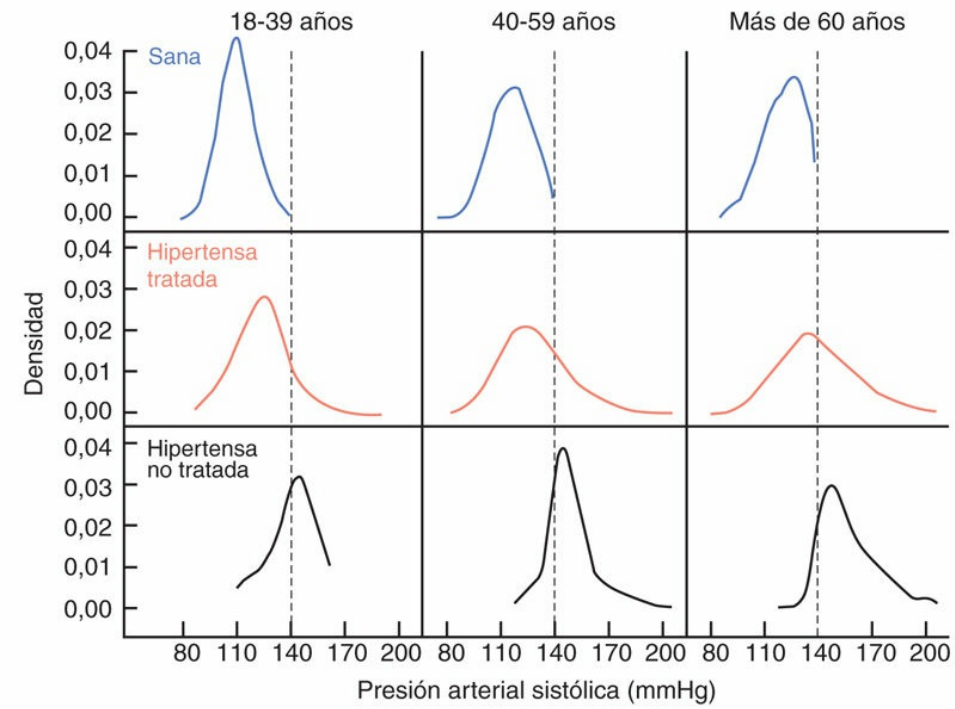
Una cuestión importante en la prevención es si nuestra estrategia debe dirigirse a grupos que se sabe que tienen un alto riesgo o si se deben ampliar los esfuerzos de prevención primaria a la población general en su conjunto. Este tema fue tratado por primera vez por Rose en 1985[3] y ampliado después por Whelton en 1994[4] en una investigación sobre la prevención de la hipertensión y de los fallecimientos por arteriopatía coronaria (AC).

Los estudios epidemiológicos han demostrado que el riesgo de mortalidad por AC aumenta de manera constante con los incrementos tanto de la presión arterial sistólica como diastólica, sin que exista un umbral conocido. En las figuras 19.5A y B se muestra la distribución de la presión arterial sistólica en la población general de hombres y mujeres mayores de 18 años en Estados Unidos (2001-2008), respectivamente.



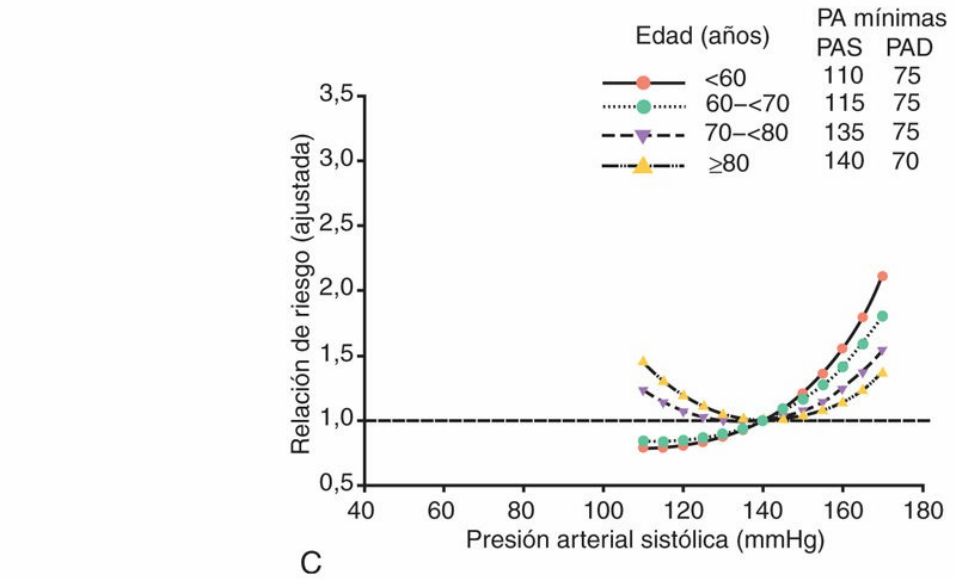
NOTA: La línea discontinua vertical es el punto de corte de la hipertensión sistólica.
FUENTE: CDC/NCHS, National Health and Nutrition Examination Survey, 2001–08.

A



NOTA: La línea discontinua vertical es el punto de corte de la hipertensión sistólica.
FUENTE: CDC/NCHS, National Health and Nutrition Examination Survey, 2001–08.

B



C

FIG. 19.5 (A) Presión arterial sistólica media en hombres mayores de 18 años, por edad y estado de hipertensión. (B) Presión arterial sistólica media en mujeres mayores de 18 años, por edad y estado de hipertensión. (C) Relación de riesgo ajustada de la primera ocurrencia de muerte por todas las causas, infarto de miocardio no fatal o accidente cerebrovascular no fatal en función de la edad (en incrementos de 10 años), presión arterial sistólica *(PAS).* Presión arterial sistólica de referencia para la relación de riesgo: 140 mmHg, respectivamente. Las presiones arteriales sanguíneas *(PA)* son el promedio de todos los registros posteriores a los valores basales, tras instaurar el tratamiento. Los términos cuadráticos para la presión arterial sistólica fueron estadísticamente significativos en todos los grupos de edad (todos *P* <0,001). El ajuste se basó en el sexo, la raza, los antecedentes de infarto de miocardio, insuficiencia cardiaca, enfermedad vascular periférica, diabetes, accidente cerebrovascular o ataque isquémico transitorio, insuficiencia renal y tabaquismo. *PAD,* presión arterial diastólica. (A y B, de Wright JD, Hughes JP, Ostchega Y, et al. Mean systolic and diastolic blood

pressure in adults aged 18 and over in the United States, 2001–2008. *Natl Health Stat Report.* 2011;(35):1–22, 24. C modificada de Denardo SJ, Gong Y, Nichols WW, et al. Blood pressure and outcomes in very old hypertensive coronary artery disease patients: an INVEST substudy. *Am J Med.* 2010;123(8):719–726.)

Si se tiene en cuenta la población general estadounidense mayor de 50 años, en la figura 19.5C se muestra el riesgo de un criterio de valoración compuesto de primera ocurrencia de muerte por cualquier causa, infarto de miocardio no mortal o accidente cerebrovascular no mortal respecto a la presión arterial sistólica; el riesgo aumenta de forma constante con el incremento de la presión arterial sistólica. Las personas menores de 60 años con una presión arterial sistólica de 160 mmHg tenían un riesgo 1,5 veces superior de sufrir AC que aquellas cuya presión arterial sistólica era inferior a 140 mmHg.

Basado en el Joint National Committee on the Prevention, Detection, Evaluation and Treatment of High Blood Pressure (JNC 7), valores tan bajos como los que definen la prehipertensión (presión arterial sistólica y diastólica de 120 a 139 mmHg y de 80 a 99 mmHg, respectivamente) pueden tener un riesgo un 20% superior de sufrir accidentes cerebrovasculares[5].

Así pues, parece razonable combinar una estrategia de alto riesgo con otra de tipo poblacional: un conjunto de medidas preventivas dirigidas a quienes tienen un riesgo especialmente alto y otro diseñado para la prevención primaria de la hipertensión y dirigido a la población general.

Estos análisis pueden tener implicaciones significativas para los programas de prevención. Los tipos de medidas preventivas que podrían utilizarse para las personas de alto riesgo a menudo son diferentes de los que son aplicables a la población general. Las personas en situación de alto riesgo, y que son conscientes de que tienen un alto riesgo, son más propensas a tolerar procedimientos más costosos, incómodos e incluso más invasivos. Sin embargo, a la hora de aplicar una medida preventiva a una población general, la medida debe tener un bajo coste y ser solo mínimamente invasiva; se debe asociar con dolores o molestias relativamente escasos si se quiere que sea aceptable para la población general.

En la figura 19.6 se muestra el objetivo de una estrategia poblacional, que supone un descenso de toda la curva de distribución de la presión arterial cuando se aplica una intervención para reducir la presión arterial en toda una comunidad, como la reducción del contenido de sal de alimentos procesados. Debido a que la presión arterial de la mayoría de los miembros de la población está por encima de los niveles más bajos que se consideran óptimos, incluso un pequeño descenso (desplazamiento a la izquierda) de la curva puede tener unos beneficios significativos para la salud pública, como sugirió Rose hace unas tres décadas[3]. De hecho, este cambio podría prevenir más accidentes cerebrovasculares en la población que un tratamiento satisfactorio limitado a las personas de «alto riesgo». Además, Rose[3] señaló que la estrategia de alto riesgo es fundamental para la protección de las personas susceptibles. En última instancia, sin embargo, nuestro deseo es

entender las causas básicas de la incidencia de la enfermedad (en este caso, hipertensión arterial) y desarrollar e implementar los medios necesarios para su prevención (primaria). Rose llegó a la siguiente conclusión:

Siendo realistas, muchas enfermedades continuarán requiriendo durante mucho tiempo ambas estrategias y, por fortuna, la competencia entre ellas suele ser innecesaria. Sin embargo, la preocupación principal siempre debe ser el descubrimiento y el control de las causas de la incidencia[3].

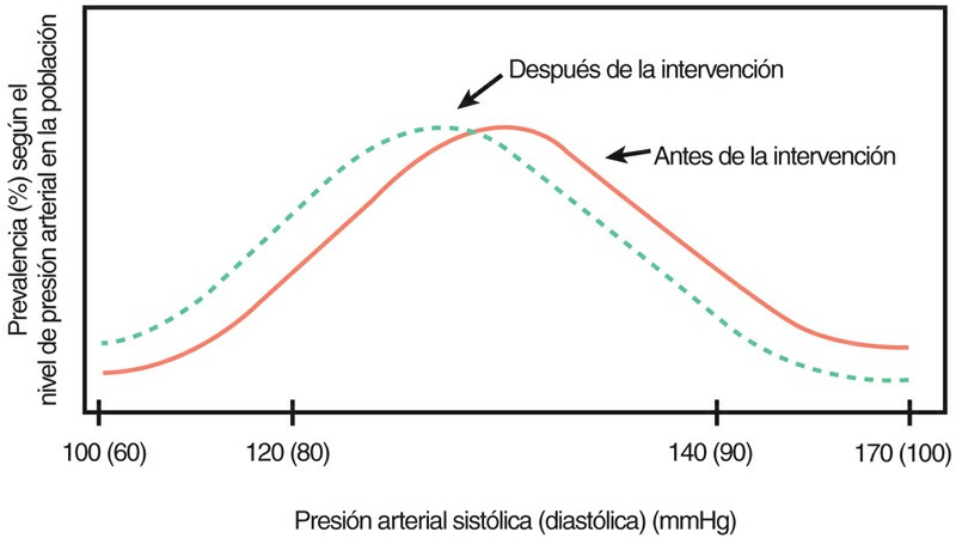


FIG. 19.6 Representación de los efectos de una estrategia de intervención poblacional sobre la distribución de la presión arterial. (De los National Institutes of Health. *Working Group Report on Primary Prevention of Hypertension.* NIH Publication N.° 93– 2669. Washington, DC: National Heart, Lung, and Blood Institute;1993:8.)

Epidemiología y medicina clínica: tratamiento hormonal sustitutivo en mujeres posmenopáusicas

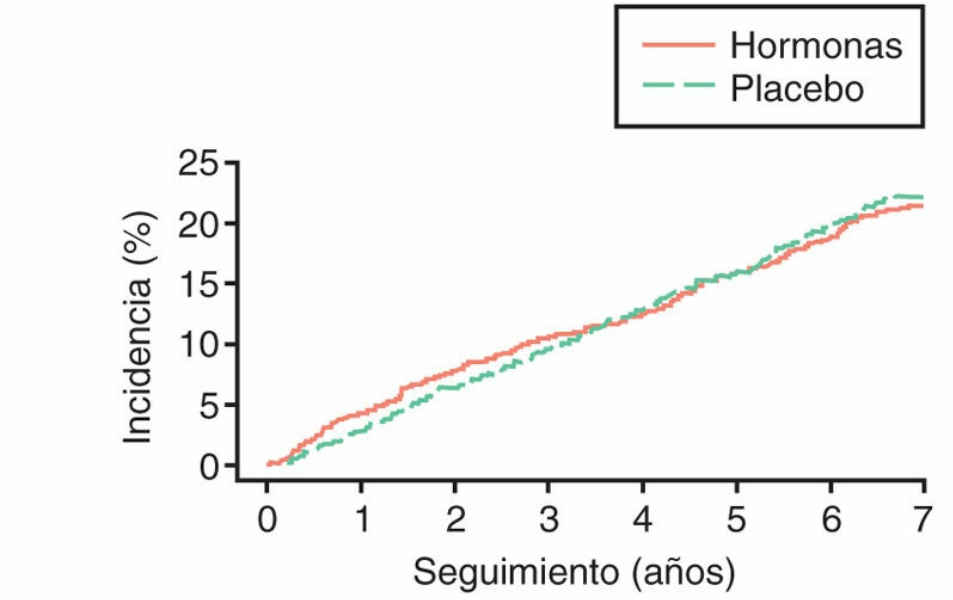
La epidemiología también puede considerarse una ciencia básica de la investigación clínica. Los datos obtenidos a partir de los estudios epidemiológicos son esenciales para la toma de decisiones clínicas en muchas situaciones. La comprensión de la epidemiología es crucial para el proceso de diseño de estudios significativos de la historia natural de las enfermedades, la calidad de los diferentes métodos diagnósticos y la eficacia de las intervenciones clínicas. La epidemiología es muy relevante para abordar muchas incertidumbres y dilemas de las políticas clínicas, algunos de los cuales no siempre pueden resolverse con facilidad.

Un ejemplo claro es el uso del tratamiento hormonal sustitutivo (THS) en mujeres posmenopáusicas. En 1966, el médico Robert Wilson publicó el libro titulado *Feminine Forever,* que recomendaba el THS para las mujeres posmenopáusicas. Después de la publicación de este libro, millones de mujeres posmenopáusicas comenzaron a tomar estrógenos con la esperanza de mantener su juventud y su atractivo y evitar los síntomas desagradables de la menopausia que suelen producirse, como sofocos, diaforesis nocturna y sequedad vaginal. La comunidad médica aceptó en gran parte la recomendación de Wilson del THS con estrógenos, e incluso los libros de texto de ginecología la apoyaban. Sin embargo, en la década de 1970 se describió un aumento del riesgo de cáncer de útero en mujeres que tomaban THS con estrógenos. Como resultado, los estrógenos se combinaron después con progestágenos, que contrarrestan el efecto de los estrógenos sobre el revestimiento endometrial. Esta combinación provoca una hemorragia uterina mensual parecida a un periodo menstrual normal.

Posteriormente se publicaron varios estudios observacionales no aleatorizados que describían otros beneficios para la salud, como un menor número de ataques cardiacos y derrames cerebrales, menos osteoporosis y menos fracturas de cadera asociados con el THS. Considerando todo el corpus de evidencia que se había acumulado, el respaldo a la conclusión de que los estrógenos protegían a las mujeres contra las cardiopatías parecía sólido y, por lo general, concordante. Se aconsejó a las mujeres que, al cumplir 50 años, consultasen con su médico si debían comenzar un THS para protegerse contra las cardiopatías y otras afecciones asociadas con el envejecimiento.

Dado que había poca evidencia favorable procedente de ensayos aleatorizados que utilizasen criterios de valoración objetivos de enfermedad, como el riesgo de infarto de miocardio, se iniciaron dos ensayos aleatorizados: el Heart and Estrogen/Progestin Replacement Study (HERS) y

el Women's Health Initiative (WHI). El estudio HERS[6] incluyó a 2.763 mujeres con AC conocida. Se observó que, a diferencia de las creencias aceptadas, el THS combinado aumentaba el riesgo de las mujeres de sufrir un infarto de miocardio durante los primeros años después de comenzar el tratamiento. El estudio no encontró evidencia de que el THS protegiese durante un periodo de seguimiento de casi 7 años (fig. 19.7).



N.º en situación de riesgo								
Hormonas	1.380	1.303	1.247	1.196	1.133	1.043	984	354
Placebo	1.383	1.334	1.269	1.209	1.122	1.039	976	336

FIG. 19.7 Estimaciones de Kaplan-Meier de la incidencia acumulada de eventos de arteriopatía coronaria (fallecimiento e infartos de miocardio no mortales). (De Grady D, Herrington D, Bittner V, et al, para el grupo de investigación HERS. Cardiovascular disease outcomes during 6.8 years of hormone therapy: heart and estrogen/progestin replacement study follow-up [HERS II]. *JAMA*. 2002;288:49–57.)

El WHI[7] fue un ensayo aleatorizado controlado con placebo de 16.608 mujeres diseñado en 1991 y 1992 para evaluar el THS para la prevención primaria de cardiopatías y otras enfermedades frecuentes en ancianas. La duración prevista del ensayo fue de 8,5 años. Uno de los brazos del ensayo fue una investigación aleatorizada controlada con placebo de estrógeno más progestágeno en mujeres posmenopáusicas que tenían un útero intacto. Este grupo del estudio se interrumpió 3 años antes de tiempo porque, en ese momento, los resultados mostraban un mayor riesgo de ataque cardiaco, accidente cerebrovascular, cáncer de mama y trombosis (fig. 19.8). Aunque el estudio mostró una reducción de la incidencia de osteoporosis, fracturas óseas y cáncer colorrectal, de forma global, los peligros del THS superaban a esos beneficios.

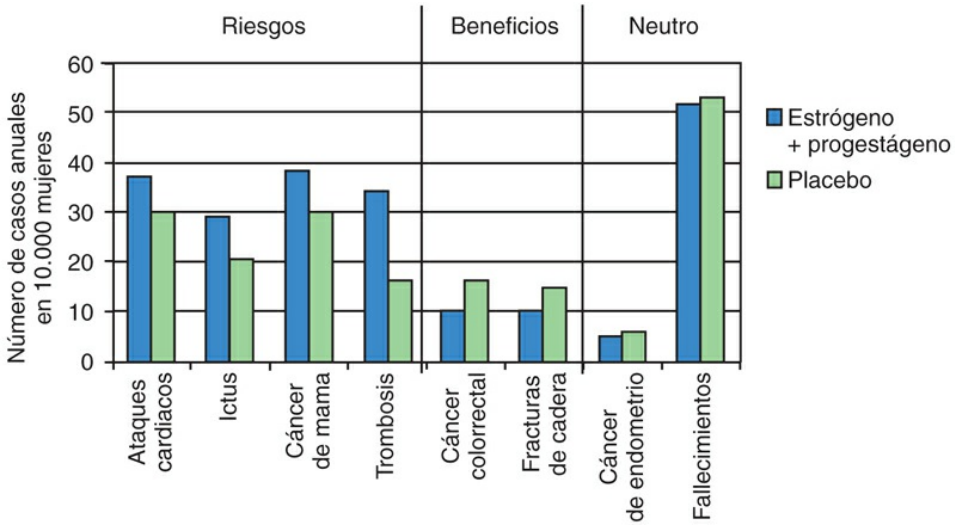


FIG. 19.8 Tasas de enfermedad para las mujeres asignadas al grupo de estrógeno más progestágeno o de placebo en el estudio Women's Health Initiative (WHI). (WHI en línea: disponible en http://www.nhlbi.nih.gov/health/women/upd2002.htm. Consultado el 14 de junio de 2013.)

Solo alrededor del 2,5% de las mujeres incluidas tuvieron eventos adversos. Basándose en los resultados del estudio, se ha estimado que, cada año, por cada 10.000 mujeres que toman estrógeno más progestágeno, sería de esperar que 7 mujeres más tuviesen un ataque cardiaco (37 mujeres que tomasen estrógeno más progestágeno tendrían un ataque cardiaco en comparación con 30 mujeres que tomasen placebo), 8 mujeres más tuviesen un accidente cerebrovascular, 8 mujeres más tuviesen cáncer de mama y 18 mujeres más tuviesen trombosis. Al mismo tiempo, sería de esperar que hubiese 6 casos menos de cáncer colorrectal y 5 fracturas de cadera menos.

Muchas mujeres que habían tomado THS se sorprendieron por los resultados del estudio WHI. Estos indicaban claramente que, en las mujeres que tomaban estrógeno más progestágeno para la protección contra las cardiopatías, los riesgos de criterios de valoración cardiovasculares eran en realidad mayores. Estas mujeres se quedaron con la incertidumbre de si debían continuar con el THS o buscar alternativas. Muchas de ellas también creían que habían sido engañadas por la comunidad médica, ya que, durante muchos años, sus médicos les habían tranquilizado acerca de la eficacia y la seguridad del THS, *a pesar de la ausencia de datos claros* procedentes de ensayos aleatorizados controlados con placebo. Otro elemento que complica el proceso de toma de decisiones de las mujeres es que el WHI no abordó la cuestión a la que se enfrentan muchas de las que a menudo toman THS combinado durante periodos breves para prevenir y aliviar síntomas climatéricos como los sofocos.

Una cuestión metodológica principal es determinar por qué había tal discrepancia entre los resultados del estudio WHI, aleatorizado y controlado con placebo, sobre el riesgo de cardiopatías y los resultados de muchos estudios observacionales no aleatorizados que habían respaldado previamente el beneficio protector del THS combinado. Esta cuestión es

crucial, ya que, en muchas áreas de la medicina y la salud pública, dependemos de los resultados de estudios observacionales no aleatorizados, porque los costes de los ensayos aleatorizados pueden ser prohibitivos y puede que los estudios aleatorizados no sean factibles por otros motivos.

Se han ofrecido varias explicaciones[8-10]. En los estudios observacionales, las mujeres a las que se prescribía THS solían ser mujeres más sanas que tenían un mejor perfil de riesgo cardiovascular. Las mujeres que usan THS suelen tener una mayor educación, son más delgadas, realizan más actividad física, son menos propensas a ser fumadoras, se preocupan más por la salud y tienen un mayor nivel socioeconómico que las mujeres que no lo usan. A menudo, las mujeres a las que se prescribía THS eran consideradas cumplidoras del tratamiento, y este tipo de pacientes suelen tener otros patrones de conducta saludables. Por tanto, se puede haber producido una confusión por el estilo de vida y otros factores en los estudios observacionales. Además, en dichos estudios, cuando se produjeron precozmente los efectos adversos y motivaron la interrupción del THS, puede que estos eventos no siempre se hayan identificado en las mediciones transversales periódicas realizadas. Una explicación adicional relacionada con el riesgo cardiovascular es que los estudios observacionales se realizaron poco después de la menopausia, cuando se sabe que se producen los efectos beneficiosos de la THS, como sus efectos favorables sobre los lípidos y la función endotelial, mientras que el ensayo WHI incluyó mujeres mucho mayores con una aterosclerosis subyacente extensa, entre las cuales existe un predominio de los efectos protrombóticos e inflamatorios de la THS[11].

Claramente, en el futuro será esencial abordar estos problemas cuando se utilicen estudios observacionales no aleatorizados como base para el desarrollo y la difusión de guías clínicas y para establecer nuevas políticas de salud pública.

Evaluación del riesgo

Uno de los usos principales de la epidemiología en relación con las políticas de salud pública es la evaluación del riesgo, que se ha definido como la caracterización de los posibles efectos adversos para la salud de la exposición humana a los peligros ambientales. La evaluación del riesgo se considera, por lo tanto, parte de un proceso global que va de la investigación a la evaluación del riesgo y de aquí a la gestión de dicho riesgo, como se muestra en la figura 19.9. Samet y cols.[12] han revisado la relación entre la epidemiología y la evaluación del riesgo y han descrito la gestión del riesgo como un proceso que implica la evaluación de acciones normativas alternativas y la selección de la estrategia que se debe aplicar. La gestión del riesgo se sigue de la comunicación de este, que es la transmisión de los resultados de la evaluación del riesgo a quienes necesitan saber dichos resultados para participar en la formulación de políticas y tomar las medidas adecuadas de gestión del riesgo, incluyendo la comunicación al público en general.

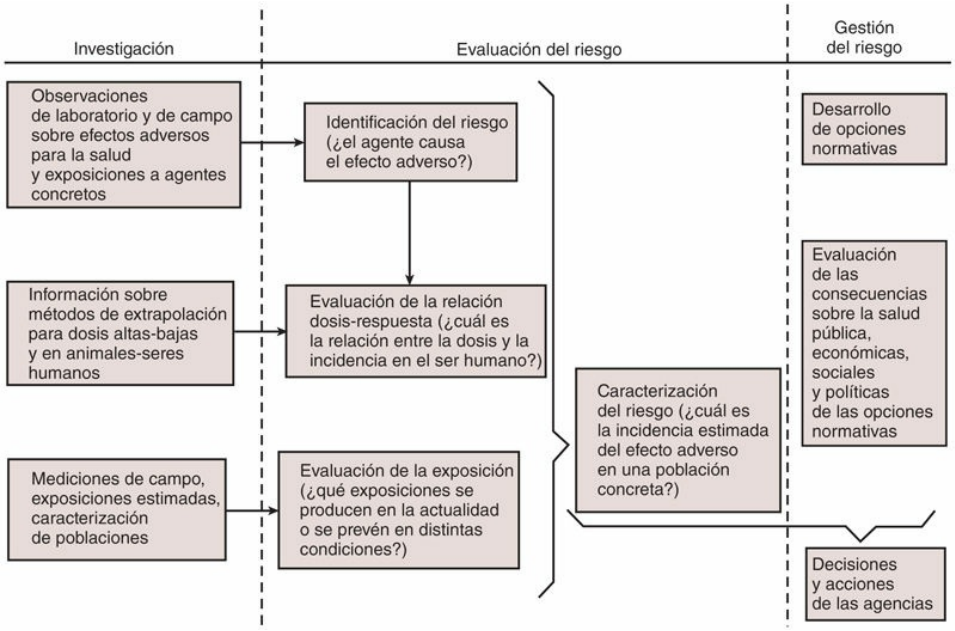


FIG. 19.9 Relaciones entre los cuatro pasos de la evaluación del riesgo y entre la evaluación y la gestión del riesgo. (Modificada del Committee on the Institutional Means for Assessment of Risks to Public Health, Commission on Life Sciences, National Research Council. *Risk Assessment in the Federal Government: Managing the Process.* Washington, DC: National Academy Press; 1983:21.)

El National Research Council (1983) ha descrito cuatro pasos en el proceso de evaluación de riesgos[13]:

1. *Identificación del riesgo:* determinación de si una sustancia química concreta tiene una relación causal con los efectos particulares de

salud.
2. *Evaluación de la relación dosis-respuesta:* determinación de la relación entre la magnitud de la exposición y la probabilidad de aparición de los efectos sobre la salud en cuestión.
3. *Evaluación de la exposición:* determinación del grado de exposición humana antes o después de la aplicación de los controles normativos.
4. *Caracterización del riesgo:* descripción de la naturaleza (y muchas veces la magnitud) del riesgo en el ser humano, incluida la incertidumbre consiguiente.

Es evidente que los datos epidemiológicos son esenciales en cada uno de estos pasos, aunque la epidemiología no es la única disciplina científica relevante en el proceso de evaluación del riesgo. En especial, la toxicología también desempeña un papel destacado, y conciliar los datos epidemiológicos y toxicológicos cuando los resultados de las respectivas disciplinas no concuerdan sigue siendo uno de los principales retos.

Hay varios problemas metodológicos relevantes que afectan al uso de la epidemiología en la evaluación del riesgo. Dado que los estudios epidemiológicos suelen abordar la relación entre una exposición ambiental y el riesgo de una enfermedad, es crucial realizar una evaluación rigurosa de cada variable. Tal vez el problema principal sea la evaluación de la exposición.

Evaluación de la exposición

Los datos sobre la exposición suelen proceder de varios tipos de fuentes (cuadro 19.1) y cada uno tiene sus ventajas e inconvenientes; entre estos últimos, se incluyen la falta de exhaustividad y los sesgos de notificación. Con frecuencia, los investigadores utilizan varias fuentes de información sobre la exposición, pero suelen surgir problemas cuando diferentes fuentes proporcionan información contradictoria.

Cuadro 19.1 Fuentes de los datos de exposición

1. Entrevistas:
    a. Sujeto.
    b. Sustituto.
2. Registros de empleo o de otro tipo.
3. Historiales médicos.
4. Historias clínicas hospitalarias.
5. Registros de enfermedades (p. ej., registros de cáncer).
6. Certificados de mortalidad.

Otro problema a la hora de evaluar la exposición es que los factores

macroambientales suelen afectar a muchas personas de forma simultánea, de modo que las exposiciones individuales pueden ser difíciles de medir. Como resultado, se suelen elegir enfoques ecológicos en los que se utilizan mediciones agregadas en lugar de individuales (descrito en el capítulo 7), y la agregación se lleva a cabo a menudo en grandes áreas y poblaciones. Por tanto, las características de la comunidad se atribuyen a las personas que residen en esa comunidad, pero la validez de la caracterización de una exposición individual mediante este proceso suele ser cuestionable (recuerde la «falacia ecológica»). Además, las historias de exposición personal pueden ser muy difíciles de obtener, tanto de forma retrospectiva como prospectiva y pueden estar sujetas a un error de medición considerable. Además, el largo periodo de latencia o de inducción entre la exposición y el desarrollo de la enfermedad hace que sea necesario determinar las exposiciones pasadas, lo que es particularmente difícil. A veces es posible evaluar la exposición a factores macroambientales a nivel individual, como se hizo en un estudio complementario dentro del Multi-Ethnic Study of Atherosclerosis (MESA Air). En este estudio se calcularon los niveles de contaminación del aire en los hogares considerando la distancia respecto de una carretera principal[14] y mediante el uso de un dispositivo especial para monitorizar en el hogar los niveles de contaminación del aire[15].

Una serie relativamente paralela de problemas se observa cuando se trata de caracterizar las exposiciones laborales de un trabajador individual y vincular una exposición en el trabajo con un resultado adverso para la salud. En primer lugar, debido a que es probable que un trabajador esté expuesto a muchos agentes diferentes en un entorno industrial, suele ser difícil separar el riesgo independiente que se puede atribuir a una sola exposición específica. En segundo lugar, debido a que a menudo suele existir un periodo de latencia prolongado entre la exposición y el desarrollo posterior de la enfermedad, los estudios sobre la relación exposición-enfermedad pueden ser difíciles; por ejemplo, a no ser que pueda realizarse un estudio prospectivo simultáneo (v. capítulo 8), los recuerdos pueden ser deficientes y se pueden haber perdido los registros de exposición. En tercer lugar, el riesgo de enfermedades puede ser mayor en las personas que viven cerca de una planta industrial, por lo que puede ser difícil determinar qué proporción del riesgo de un trabajador se debe a vivir cerca de la planta y cuál se debe a una exposición ocupacional en el propio ámbito laboral.

Tal vez el problema fundamental a la hora de medir la exposición en estudios epidemiológicos sea que las fuentes y las medidas con frecuencia son indirectas. Por ejemplo, en los últimos años ha surgido un interés considerable sobre los posibles efectos para la salud de los campos electromagnéticos (CEM). Este interés surgió tras el artículo de Wertheimer y Leeper, publicado en 1979[16], que describía un aumento de los niveles de leucemia en niños que viven cerca de tendidos eléctricos de alta tensión. Posteriormente, se han planteado muchas cuestiones metodológicas, y la cuestión de si tales campos se asocian con efectos adversos para la salud

sigue sin resolverse. Por ejemplo, las conclusiones fueron discrepantes en una actualización de dos metaanálisis y en un metaanálisis más reciente realizado por el mismo primer autor[17,18].

Al estudiar los CEM, se utilizan varias estrategias para medir la exposición, como la configuración de los circuitos eléctricos en el hogar, la medición puntual o durante 24 horas de los campos, o la autonotificación del uso de aparatos eléctricos. Sin embargo, los resultados de diferentes estudios con respecto al riesgo de enfermedad difieren según el tipo de medición de la exposición utilizado. De hecho, las mediciones reales de campos magnéticos, incluso las de 24 horas, generan asociaciones más débiles con la leucemia infantil que las de los códigos de configuración de los circuitos eléctricos[19]. Esta observación plantea dudas sobre cualquier posible relación causal entre la exposición a los campos magnéticos y la aparición de la enfermedad.

Incluso la mejor medida indirecta de la exposición a menudo deja sin responder las preguntas críticas. En primer lugar, la exposición no suele ser dicotómica; por tanto, se necesitan datos respecto a la *dosis* de exposición para explorar una posible relación dosis-respuesta. En segundo lugar, es importante saber si la exposición fue continua o periódica. Por ejemplo, en la patogenia del cáncer, una exposición periódica en la que se alternan periodos con exposición y sin exposición puede permitir la reparación del ADN durante los periodos sin exposición. En una exposición continua, dicha reparación no podría producirse. Por último, la información acerca de la latencia es crucial: ¿cuánto dura el periodo de latencia y cuál es su rango? Estos datos son esenciales para que se puedan centrar los esfuerzos en la determinación de la exposición en un periodo de tiempo concreto en el que podría haberse producido una exposición causal.

Debido a estos problemas a la hora de medir la exposición utilizando estrategias indirectas, se ha puesto una gran atención en el uso de marcadores biológicos de exposición. (El uso de estos biomarcadores se ha denominado *epidemiología molecular*.)[20] El uso de biomarcadores tiene la ventaja de que permite superar algunos de los problemas de un recuerdo limitado o del desconocimiento de una exposición. Además, los biomarcadores pueden subsanar los errores secundarios a la variación de la absorción o el metabolismo individual al centrarse en un paso posterior de la cadena causal.

Los biomarcadores pueden ser marcadores de exposición, marcadores de cambios biológicos secundarios a exposiciones o marcadores de riesgo o de susceptibilidad. En la figura 19.10 se ilustran esquemáticamente los diferentes tipos de exposición que se pueden escoger para su medición.



FIG. 19.10 ¿Qué exposiciones estamos tratando de medir?

También podríamos querer medir los niveles ambientales de sustancias posiblemente tóxicas en un ambiente general, los niveles a los que está expuesta una persona concreta, la cantidad de sustancia absorbida o la cantidad de sustancia o metabolito de la sustancia absorbida que llega al tejido diana. Los biomarcadores nos acercan a la posibilidad de medir una exposición en una etapa específica en el proceso por el cual una exposición se relaciona con la enfermedad en el ser humano. Por ejemplo, se pueden medir no solo los niveles ambientales de una sustancia, sino también los aductos de ADN que reflejan el efecto de la sustancia sobre los procesos biológicos en el cuerpo después de la absorción.

No obstante, a pesar de estas ventajas, los biomarcadores suelen proporcionar una respuesta dicotómica (una persona ha estado expuesta o no expuesta) y no suelen aclarar cuestiones importantes, como estas:

- ¿Cuál fue la dosis total de exposición?
- ¿Cuál fue la duración de la exposición?
- ¿Hace cuánto tiempo ocurrió la exposición?
- ¿Fue la exposición continua o periódica?

Un ejemplo de algunas de estas deficiencias es la cotinina salival, que es un biomarcador de la absorción de nicotina en los fumadores. Como es un marcador solo del consumo de tabaco reciente, no proporciona información sobre la duración de la exposición o si el hábito era continuo o periódico.

Las respuestas a estas preguntas son cruciales para interpretar de forma adecuada la posible importancia biológica de una determinada exposición. Por ejemplo, al evaluar la plausibilidad biológica de derivar una inferencia causal a partir de las observaciones realizadas sobre la exposición y el resultado, se necesitan datos relevantes que permitan determinar si el intervalo observado entre la exposición y el desarrollo de la enfermedad es compatible (desde el punto de vista biológico) con lo que se sabe a partir de otros estudios sobre el periodo de incubación de la enfermedad.

Además de estas consideraciones, una limitación potencial del uso de biomarcadores de exposición es que, en un estudio de casos y controles tradicional, la obtención de una muestra biológica y la medición de un biomarcador se realizan solo después del inicio de la enfermedad. Por lo tanto, es imposible averiguar si la exposición estuvo presente antes del inicio

de la enfermedad de interés. Sin embargo, esta deficiencia no está presente en los estudios de casos y controles dentro de una cohorte en la que muestras biológicas, como el suero o la orina, se congelan y almacenan al inicio del estudio, es decir, antes de que se desarrollen casos incidentes durante el seguimiento de la cohorte.

Hay que señalar que el uso de biomarcadores no es nuevo en epidemiología. En el Eclesiastés está escrito: «No hay nada nuevo bajo el sol»[21]. Incluso antes de la revolución de la biología molecular, las técnicas de laboratorio eran esenciales en muchos estudios epidemiológicos; se trataba de aislamientos y cultivos bacterianos, fagotipificación de microorganismos, aislamiento viral, estudios serológicos y análisis de fracciones de lipoproteínas-colesterol. Los grandes avances de la biología molecular han dado lugar a una nueva variedad de biomarcadores que son relevantes para áreas como la carcinogénesis. Estos biomarcadores no solo identifican a los individuos expuestos, sino que también arrojan nueva luz sobre el proceso patogénico de la enfermedad en cuestión.

Metaanálisis

Cuando se utilizan datos epidemiológicos para la formulación de políticas de salud pública, surgen varias preguntas de carácter científico:

1. ¿Pueden los métodos epidemiológicos detectar pequeños aumentos del riesgo que sean clínicamente significativos?
2. ¿Cómo se pueden conciliar las discordancias entre los datos en animales y en el ser humano?
3. ¿Cómo se pueden utilizar los datos epidemiológicos incompletos o equívocos?
4. ¿Cómo se pueden interpretar los resultados cuando las observaciones de los estudios epidemiológicos son discordantes?

Muchos de los riesgos de los que nos ocupamos pueden ser muy pequeños, pero es posible que tengan gran importancia para la salud pública debido a la gran cantidad de personas expuestas, con el potencial resultante de efectos adversos para la salud en muchas personas (recuerde la hipótesis propuesta por Rose[3]). Sin embargo, un pequeño aumento observado del riesgo relativo por encima de 1,0 puede deberse fácilmente a sesgos o a otras limitaciones metodológicas, por lo que estos resultados deben interpretarse con mucha cautela a menos que los resultados se hayan replicado y que se haya obtenido otra evidencia que los respalde.

Dado que los resultados de los diferentes estudios epidemiológicos pueden no ser concordantes y a veces pueden entrar en conflicto de forma drástica, se ha intentado sistematizar el proceso de revisión de la literatura epidemiológica sobre un tema determinado. Un proceso, la revisión sistemática, utiliza una metodología estandarizada para seleccionar y evaluar artículos revisados por pares con el fin de sintetizar la literatura sobre un tema de salud específico[22]. Las revisiones sistemáticas pueden acompañarse de un proceso similar, denominado metaanálisis, que se ha definido como «el análisis estadístico de una gran recopilación de resultados de análisis de estudios individuales con el fin de integrar los resultados»[23]. El metaanálisis permite agregar los resultados de un conjunto de estudios incluidos en una revisión sistemática, con una ponderación adecuada de cada estudio en función del número de sujetos analizados y de otras características. Puede ser útil para ofrecer una perspectiva global de un problema cuando los resultados de los estudios son discordantes.

Sin embargo, los metaanálisis plantean una serie de problemas y preguntas. En primer lugar, ¿el análisis debe incluir todos los estudios disponibles o solo los estudios publicados? En segundo lugar, cuando los riesgos relativos o los *odds ratio* de varios estudios son diferentes (es decir, los resultados no son homogéneos), los resultados del metaanálisis pueden ocultar diferencias

significativas entre los estudios individuales. Por tanto, es esencial que una revisión sistemática que resulte en un metaanálisis incluya únicamente estudios con diseños bien establecidos y que reúnan criterios de calidad. En tercer lugar, los resultados de los propios metaanálisis puede que no siempre sean reproducibles por otros autores. Por último, una revisión sistemática, con o sin metaanálisis, está sujeta al problema del sesgo de publicación (se comenta más adelante en este capítulo). En la figura 19.11 se muestra un «diagrama de bosque» hipotético y la definición de sus componentes. El diagrama de bosque es el tipo de presentación que suele utilizarse para mostrar los resultados de los estudios individuales, así como los resultados del metaanálisis. En la figura 19.12 se muestra un diagrama de bosque sobre la relación entre la posición socioeconómica y la depresión. Observe como de los 51 estudios incluidos en este metaanálisis, 5 indican una asociación negativa. Por tanto, los resultados de este metaanálisis no son enteramente homogéneos.

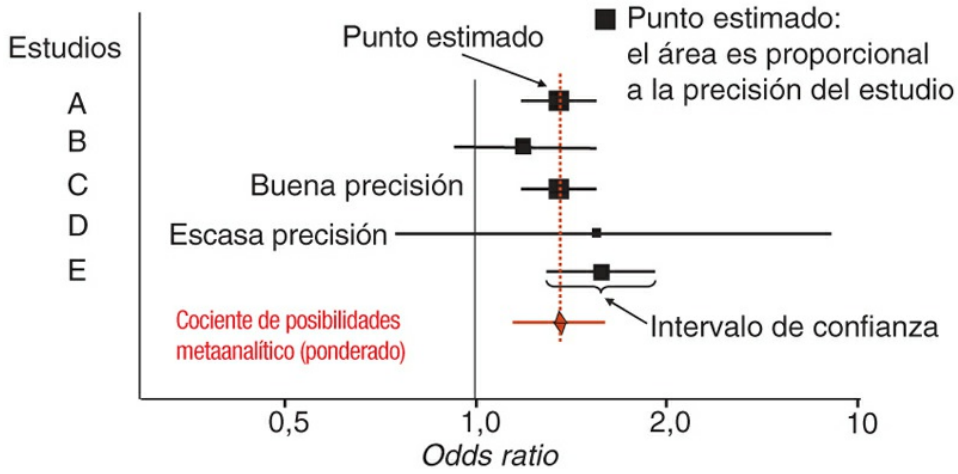


FIG. 19.11 Diagrama de bosque hipotético, con componentes, marcados, que muestran el tipo de presentación esquemática que se usa con frecuencia para mostrar los resultados de estudios individuales *(A–E)*, así como los resultados de metaanálisis.

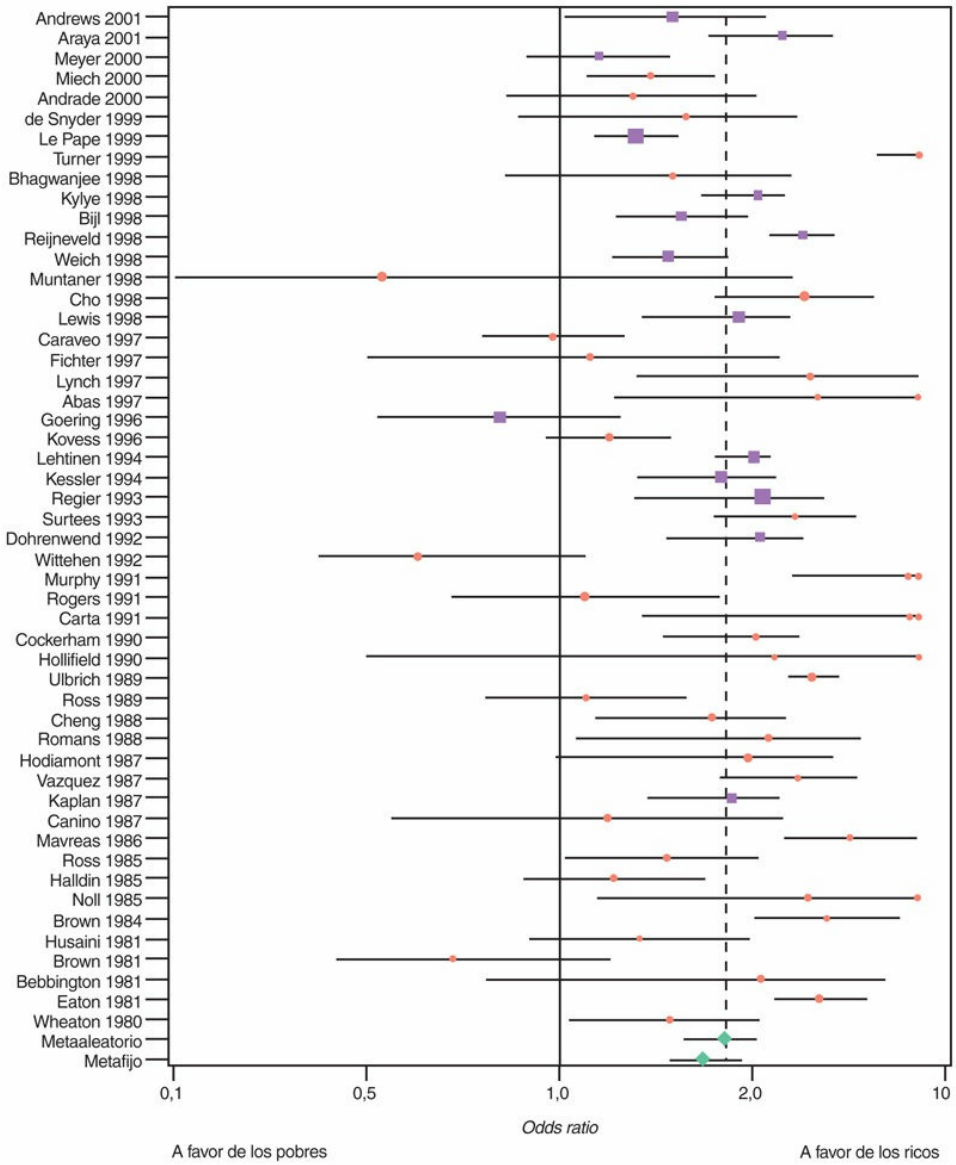


FIG. 19.12 Cocientes de posibilidades *(odds ratios)* de la depresión mayor en el grupo de posición socioeconómica más baja en 51 estudios de prevalencia publicados después de 1979. *Líneas horizontales,* intervalo de confianza del 95%. Los cuadrados muestran estimaciones originales; los rombos muestran resultados metaanalizados. (De Lorant V, Deliège D, Eaton W, et al. Socioeconomic inequalities in depression: a meta-analysis. *Am J Epidemiol.* 2003;157(2):98–112.)

El metaanálisis se aplicaba originariamente a los ensayos aleatorizados, pero esta técnica se está utilizando cada vez más para agregar estudios observacionales no aleatorizados, incluidos los estudios de casos y controles y de cohortes. En estos casos, los estudios no siempre comparten un diseño de investigación común. Por tanto, surge la pregunta de cuál es el grado de similitud que deben tener estos estudios para incluirlos legítimamente en un metaanálisis. Además, el control adecuado de los sesgos (como el sesgo de selección y el sesgo de clasificación errónea) es esencial, pero a menudo resulta ser una dificultad considerable en los metaanálisis. A la vista de las consideraciones que se acaban de mencionar, el tema de los metaanálisis sigue siendo muy controvertido.

Un problema final con los metaanálisis es que, frente a todas las

dificultades discutidas, dotar de carácter cuantitativo a la estimación de un único riesgo relativo u *odds ratio* derivado a partir de todos los estudios puede dar lugar a una falsa sensación de certeza en cuanto a la magnitud del riesgo. La gente a menudo tiende a tener una creencia excesiva en la validez de los resultados cuando se asocian a un número; como consecuencia, en ocasiones se pueden ignorar muchas de las dificultades que surgen en el metaanálisis.

Sesgo de publicación

En el capítulo 16 se describió el uso de los estudios en gemelos como método para distinguir las contribuciones de los factores ambientales y genéticos a la causa fundamental de la enfermedad. En esa exposición se mencionó que el grado de concordancia y discordancia en gemelos es una observación importante para extraer conclusiones sobre el papel de los factores genéticos, pero que las estimaciones de la concordancia publicadas en la literatura pueden estar sobreestimadas por el sesgo de publicación, que es la tendencia a publicar artículos que describan la concordancia para enfermedades raras en los pares de gemelos.

El sesgo de publicación no se limita a los estudios genéticos de gemelos, sino que puede ocurrir en cualquier área de la epidemiología. Es un fenómeno especialmente importante en la publicación de artículos relacionados con los riesgos ambientales y los resultados de los ensayos clínicos. El sesgo de publicación se puede producir porque los investigadores no remiten los resultados de sus estudios cuando dichos resultados no apoyan las asociaciones «positivas» y el aumento de los riesgos (es decir, los «hallazgos nulos»). Además, las revistas pueden seleccionar de modo diferencial para la publicación los estudios que consideran de mayor interés para el lector, y puede que los estudios que no describen ninguna asociación no se incluyan en esa categoría. Como resultado de ello, una revisión de la literatura que se limita a los artículos publicados puede identificar preferentemente estudios que describen un mayor riesgo. Es evidente que tal revisión es altamente selectiva y omite muchos estudios que han obtenido lo que se ha denominado resultados «negativos» (es decir, resultados que no muestran ningún efecto), que tal vez no hayan llegado a ser publicados.

Por tanto, el sesgo de publicación tiene un claro efecto sobre la revisión sistemática y el metaanálisis. Una estrategia para solventar este problema es tratar de identificar los estudios no publicados e incluirlos en el análisis (recogiendo estudios de la literatura «gris», a menudo de presentaciones de conferencias, que con frecuencia comunican resultados nulos que no son publicados). Sin embargo, esto presenta la dificultad de que, en general, es probable que los estudios no publicados no hayan superado la revisión por pares de la revista y, por tanto, su idoneidad para su inclusión en un metaanálisis puede ser cuestionable. Con independencia de si estamos hablando de un tipo tradicional de revisión bibliográfica o de un metaanálisis estructurado, se debe tener en cuenta el problema del posible sesgo de publicación.

Se ha propuesto que para prevenir el sesgo de publicación en las revisiones sistemáticas (y, por lo tanto, en los metaanálisis), se deben implementar registros de estudios, similares a la colaboración Cochrane. También existen estrategias para evaluar el sesgo de publicación en los metaanálisis, como la prueba de Begg y las pruebas de simetría. Estos enfoques se basan en

representar los valores de la medida de la asociación de los estudios (p.ej., riesgo relativo o cociente de posibilidades) frente a sus niveles de precisión (medidos por sus errores estándar, que generalmente son una función de los tamaños muestrales). Utilizando el riesgo relativo como ejemplo, a medida que aumentan los errores estándar, lo que denota una precisión decreciente, los riesgos relativos se vuelven más variables, pero se espera que sigan una distribución simétrica alrededor del riesgo relativo medio entre los estudios. Si la distribución es asimétrica, es probable que exista sesgo de publicación.

Epidemiología en la vía judicial

Como se mencionó anteriormente, la vía judicial se ha convertido en una estrategia principal para la formulación de políticas en Estados Unidos. La epidemiología está asumiendo una importancia creciente en el ámbito legal. En especial en el contexto de los daños producidos por sustancias tóxicas, proporciona uno de los principales tipos de evidencia científica que es relevante para las cuestiones implicadas. Aspectos como los efectos de las dioxinas, los implantes mamarios de silicona, el tabaquismo y los CEM son solo algunos ejemplos.

Sin embargo, el uso de datos procedentes de estudios epidemiológicos no está exento de problemas. La epidemiología responde preguntas sobre *grupos*, mientras que los tribunales suelen requerir información sobre *individuos* (donde es necesario vincular causalmente la exposición individual y su estado de enfermedad). Además, se ha dirigido una atención considerable a la interpretación judicial de la evidencia de causalidad. Mientras que el criterio legal suele ser «más probable» (es decir, que la sustancia o la exposición en cuestión es «más probable» que haya causado la enfermedad de una persona), la epidemiología se basa en gran medida en las directrices del Surgeon General de EE. UU. para las inferencias causales[24]. Se ha sugerido que un riesgo atribuible en los expuestos mayor del 50% podría constituir una prueba de «más probable»[25].

Hasta hace poco, la evidencia procedente de la epidemiología solo se aceptaba con reticencias en los tribunales, pero esto ha cambiado hasta tal punto que los datos epidemiológicos se citan a menudo como la única fuente de evidencia relevante en casos de daños producidos por sustancias tóxicas. Durante muchos años, el principio rector para el uso de la evidencia científica en los tribunales en EE. UU. era la prueba de Frye, que afirma que para que un estudio sea admisible «debe estar suficientemente establecido que haya logrado la aceptación general en el campo al que pertenece»[26]. A pesar de que términos como «aceptación general» y «campo al que pertenece» se dejaban sin definir, se ha traducido en una evaluación de si la opinión científica expresada por un experto estaba aceptada generalmente por otros profesionales de la disciplina.

En 1993, en el caso *Daubert contra Merrell Dow Pharmaceuticals*[27], en el que el demandante alegaba que una deformidad congénita de una extremidad se debió a la ingesta del fármaco Bendectin durante el embarazo, el Tribunal Supremo de EE. UU. estableció un cambio importante de las reglas de evidencia. El tribunal dictaminó que la «aceptación general» no es una condición necesaria para la admisibilidad de la evidencia científica en los tribunales. En su lugar, el juez de primera instancia pasaba a considerarse un «guardián» y se le asignaba la tarea de asegurar que el testimonio de un experto se basase en una base fiable y relevante para la «tarea en cuestión».

Por tanto, el juez «debe hacer una evaluación preliminar de si el razonamiento o la metodología en la que se basa el testimonio tiene validez científica y se puede aplicar adecuadamente a los hechos tratados». Entre las consideraciones citadas por el tribunal están si la teoría o técnica en cuestión puede ser y ha sido probada y si la metodología ha sido sometida a revisión por pares y se ha publicado.

Debido a sus nuevas responsabilidades, los jueces que presiden juicios en los que la epidemiología es una fuente importante de evidencia deben tener un conocimiento básico de los conceptos epidemiológicos (como, por ejemplo, el diseño del estudio, los sesgos y factores de confusión, así como las inferencias causales) si tienen que ser capaces de dictaminar con fundamento si la estrategia utilizada por los expertos sigue el «método científico» aceptado. Consciente de esta necesidad, el Federal Judicial Center de EE. UU. ha publicado un manual de investigación sobre la evidencia científica para los jueces que incluye una sección sobre epidemiología[28]. Aunque es prematuro saber cuál será el efecto final de la sentencia Daubert, dado el enorme aumento del uso de la epidemiología en los tribunales, es evidente que el dictado de sentencias requerirá un mayor conocimiento de la epidemiología por las numerosas partes involucradas en los procesos judiciales que utilizan la evidencia derivada de estudios epidemiológicos.

Fuentes y efecto de la incertidumbre

En 1983, el National Research Council de Estados Unidos escribió:

La dificultad analítica dominante [al realizar evaluaciones de los riesgos para la toma de decisiones políticas] es la incertidumbre generalizada… Los datos pueden ser incompletos y a menudo hay una gran incertidumbre en las estimaciones de los tipos, la probabilidad y la magnitud de los efectos para la salud asociados con una sustancia química, de los efectos económicos de una medida normativa propuesta y de la magnitud de las exposiciones humanas actuales y posibles en el futuro[29].

Esta visión sigue siendo tan relevante hoy como cuando fue escrita. La incertidumbre es una realidad que debe aceptarse y afrontarse. La incertidumbre es una parte integral de la ciencia. Lo que hoy se acepta como la «verdad» a menudo resulta ser transitorio. Mañana puede aparecer un estudio que contradiga o invalide la mejor información científica disponible en la actualidad.

La incertidumbre es relevante no solo para la evaluación de riesgos, sino también para cuestiones de tratamiento, cuestiones de prevención como el cribado y cuestiones de economía sanitaria. Es evidente que es un aspecto de gran relevancia en el ámbito legal comentado anteriormente (fig. 19.13).



"Your Honor, we, the jury, find this one too close to call."

FIG. 19.13 «Señoría, nosotros, el jurado, observamos un empate en este caso.» Una interpretación de la incertidumbre. (Arnie Levin/The New Yorker

Collection/The Cartoon Bank.)

Algunas de las posibles fuentes de incertidumbre se enumeran en el cuadro 19.2. Como se ve allí, las fuentes de incertidumbre pueden estar en el diseño del estudio o en su desarrollo e implementación, o pueden deberse a la presentación e interpretación de las observaciones del estudio. Muchas de estas fuentes se han comentado en los capítulos anteriores.

Cuadro 19.2 Ejemplos de posibles fuentes de incertidumbre en epidemiología

1. Incertidumbre secundaria al diseño del estudio:
    a. Puede que el estudio no se haya diseñado para proporcionar una respuesta pertinente a la cuestión de interés.
    b. Sesgos que no se han identificado o que no se han abordado adecuadamente:
        (1) Sesgo de selección.
        (2) Sesgo de información.
    c. Errores de medición que pueden dar lugar a errores de clasificación.
    d. Tamaño de la muestra inadecuado.
    e. Elección inapropiada de los métodos analíticos.
    f. No se han tenido en cuenta posibles factores de confusión.
    g. Uso de medidas indirectas que pueden no medir correctamente los resultados que son las principales variables dependientes de interés.
    h. Problemas de validez externa (generalización a la población de interés): las conclusiones referentes a las posibles intervenciones puede que no sean generalizables a la población diana.
2. Incertidumbre secundaria a las deficiencias a la hora de realizar e implementar el estudio:
    a. Las observaciones pueden estar sesgadas si no se aplicó el enmascaramiento de los observadores.
    b. Mala calidad de los métodos de laboratorio o de estudio.
    c. Gran proporción de no participantes y/o de personas que no responden.
    d. Incapacidad de identificar las razones de la falta de respuesta y las características de las personas que no responden.
3. Incertidumbre secundaria a la presentación e interpretación de los resultados del estudio:
    a. ¿Cómo se expresaron los resultados?
    b. Si el estudio evaluó el riesgo y la posible etiología, ¿los

factores implicados se describieron como factores de riesgo o como factores causales?

c. Si el estudio evaluó la eficacia de una medida preventiva propuesta, ¿el beneficio de la medida se expresó como reducción del riesgo relativo o como reducción del riesgo absoluto? ¿Por qué se escogió para ser expresado de esa manera, y cómo se interpretó el hallazgo?

Una de las cuestiones que figuran en el cuadro 19.2 es la de si, en un estudio sobre la eficacia de una medida preventiva, los resultados se describen como una reducción del riesgo relativo o una reducción del riesgo absoluto. A menudo, se escoge el *porcentaje* de reducción de la mortalidad, ya que ofrece una visión más optimista de la eficacia de una medida preventiva. Sin embargo, si se utiliza la reducción del riesgo absoluto, como el *número* de individuos por 1.000 cuyas vidas se salvarían, el resultado parece menos impresionante (recuerde los riesgos de enfermedad asociados con el THS presentados anteriormente en este capítulo). Si la tasa de eventos adversos, como la mortalidad por la enfermedad que se observa sin cribado, es baja, un porcentaje de reducción siempre parecerá más impresionante que una reducción del riesgo absoluto, porque el *número* de eventos que se podrían prevenir es pequeño, aunque el *porcentaje* de reducción sea mayor. Si, por ejemplo, la mortalidad en los sometidos a cribado es de 1 por 100.000 y en aquellos no cribados es de 2 por 100.000, la reducción resultante del examen es de 50%, pero la diferencia absoluta es simplemente de 1 por 100.000.

Una medida más relevante de la efectividad (y la eficiencia) de una medida preventiva o curativa es el número necesario para someterse a la intervención para prevenir un caso o una muerte por la enfermedad. Esta medida se basa en la diferencia absoluta. Por ejemplo, si la diferencia entre una nueva estrategia preventiva y la estrategia actual (tratamiento habitual) es del 20%, el número necesario a intervenir con el fin de prevenir la aparición de un caso incidente es ([100 × 1] ÷ 20) = 5. Sin embargo, si la diferencia es solo del 2%, este número se convierte en ([100 × 1] ÷ 5) = 20. Observe como la efectividad es la misma si las tasas de mortalidad son, por ejemplo, 60% y 40% o 6% y 4% en dos estudios que evaluaron diferentes intervenciones novedosas para prevenir la misma enfermedad: ([60% - 40%] ÷ 60%) = 33.3% en el primer estudio, y ([6% - 4%] ÷ 6 %) = 33.3% en el segundo estudio. Sin embargo, está claro que el primer estudio aborda un problema de salud pública más importante para el cual la prevención sería más eficiente, ya que se puede prevenir un caso sometiendo menos individuos a la nueva intervención.

Otra cuestión que contribuye a la incertidumbre a la hora de elaborar políticas y que no se suele relacionar con estudios epidemiológicos específicos, es cómo se encaja la *evidencia anecdótica,* como la proporcionada por una persona que afirma que fue sometida a cribado para la detección de cáncer de mama hace 10 años, recibió un tratamiento precoz y está viva y aparentemente bien 10 años después del cribado. A menudo hay una

tendencia a aceptar que esta evidencia respalda la eficacia del cribado a la hora de reducir la mortalidad por la enfermedad. Sin embargo, la evidencia anecdótica presenta dos problemas principales. En primer lugar, no tiene en cuenta los tumores de crecimiento lento que podrían haber sido detectados por el cribado, pero que tal vez no hubiesen afectado a la supervivencia, incluso si el paciente no hubiese sido sometido a cribado. En segundo lugar, no tiene en cuenta los tumores de crecimiento muy rápido que el cribado no habría detectado y para los que la persona no habría recibido tratamiento precoz. Es decir, para esa evidencia anecdótica de supervivencia después del cribado no hay grupo de comparación de personas que fueron sometidas a cribado, pero que *no* sobrevivieron. Como dijo un sabio, «el plural de *anécdota* no es "datos"». Sin embargo, a pesar de estas importantes limitaciones, la evidencia anecdótica que proporcionan los pacientes que han sobrevivido a una enfermedad grave puede tener un fuerte impacto emocional que puede influir significativamente en los gestores.

En última instancia, el impacto de la incertidumbre científica en la formulación de políticas de salud pública dependerá de cómo las principales partes implicadas consideren la incertidumbre. Entre los diferentes grupos de implicados están los científicos (incluidos los epidemiólogos), los gestores, los políticos y la opinión pública (o las poblaciones diana). Cada uno de estos grupos puede tener diferentes niveles de complejidad y distintos niveles y tipos de interés propio y pueden ver los datos de manera diferente y estar influenciados en grados diversos por los colegas, amigos y diversos grupos de la sociedad. Además, las personas tienen diferentes personalidades con distintos niveles de tolerancia al riesgo y diferentes formas de manejar la incertidumbre. Un mediador importante es el conjunto de valores que cada individuo tiene respecto a cuestiones como el valor de una vida humana y los principios que deben guiar la asignación de recursos limitados en una sociedad. El resultado es una compleja interacción de incertidumbre derivada de las características de un estudio, que interactúa con una red de relaciones relativas a los elementos que se acaban de describir. En la figura 19.14 se esquematizan algunas de las interrelaciones que influyen en el efecto de la incertidumbre sobre las políticas de salud pública. Estos factores son algunas de las principales preocupaciones a la hora de formular políticas clínicas y de salud pública adecuadas. Es importante que se tengan en cuenta si se quiere desarrollar e implementar con éxito un plan de acción para hacer frente a los problemas de salud en la población.

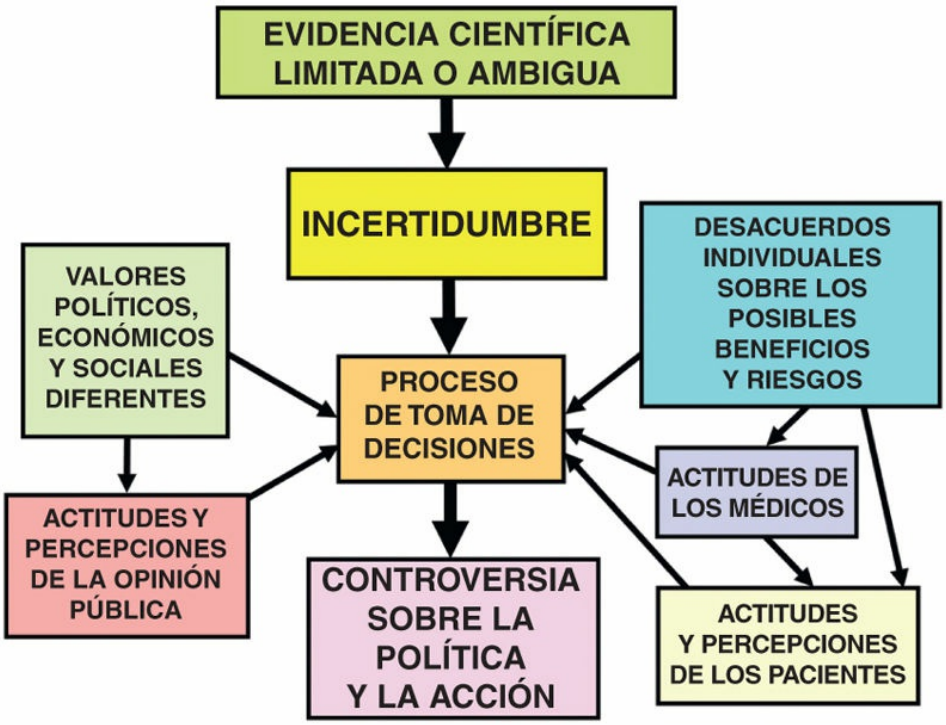


FIG. 19.14 Presentación esquemática de algunos de los factores implicados en el impacto de la incertidumbre sobre el proceso de toma de decisiones para la política sanitaria.

Aspectos políticos referentes al riesgo: ¿cuáles deberían ser los objetivos?

Se suele aceptar que las políticas de salud se elaboran en gran medida a través de los procesos de legislación y regulación. Como se señaló anteriormente, en EE. UU. la vía judicial se ha convertido en un instrumento importante para el desarrollo e implementación de políticas de salud públicas. Lo ideal sería que cada uno de estos procesos reflejase los valores y las aspiraciones sociales.

Hay ciertas cuestiones sociales destacadas que se deben considerar a la hora de tomar decisiones acerca de los riesgos. Estas son algunas de las preguntas que se deben plantear:

1. ¿Qué porcentaje de la población debe ser protegido por la política?
2. ¿Qué nivel de riesgo está dispuesta a tolerar la sociedad?
3. ¿Qué nivel de control del riesgo está dispuesta a asumir la sociedad?
4. ¿Quién debe tomar las decisiones sobre el riesgo?

A primera vista, podría parecer atractivo proteger a toda la población de cualquier grado de riesgo, pero, siendo realistas, esto es difícil de lograr, cuando no imposible. Con independencia de lo que muestren los datos de riesgo sobre las poblaciones, hay un número muy pequeño de individuos que presentan una sensibilidad extraordinaria a concentraciones mínimas de ciertas sustancias químicas. Si la cantidad admisible de una sustancia química se va a establecer en un nivel que proteja a *todos* los trabajadores, puede suceder que se detengan todos los procesos de fabricación. Del mismo modo, si se exige un riesgo cero para los trabajadores o para otras personas que puedan estar expuestas, la base económica de muchas comunidades podría destruirse. Por tanto, la elaboración de políticas requiere un equilibrio entre lo que *se puede* hacer y lo que *se debería* hacer. El grado de prioridad relacionado con la eliminación de todos los riesgos y la decisión acerca de qué porcentaje de riesgo debe ser eliminado no son decisiones científicas, sino que dependen de los valores sociales. Es de esperar que este tipo de decisiones sociales se basen en los conocimientos epidemiológicos y en otros datos científicos disponibles en el contexto de consideraciones políticas, económicas, éticas y sociales.

Conclusión

Los objetivos de la epidemiología son mejorar nuestra comprensión de la biología, la patogenia y otros determinantes de las enfermedades para mejorar la salud humana y prevenir y tratar mejor las enfermedades. Es necesario un conocimiento profundo de los problemas metodológicos que surgen con el fin de interpretar mejor los resultados epidemiológicos correctamente como base para la formulación de políticas tanto clínicas como de salud pública. El uso adecuado y prudente de los resultados de los estudios epidemiológicos es fundamental para la evaluación de los riesgos para la salud humana y para el control de estos riesgos. Por tanto, este uso es importante para la prevención tanto primaria como secundaria. Los gestores a menudo se ven obligados a desarrollar políticas en presencia de datos científicos incompletos o equívocos. En medicina clínica, en el proceso tanto diagnóstico como terapéutico, las decisiones se hacen a menudo con datos incompletos o ambiguos; es posible que esto haya sido un serio impedimento en el ámbito de la salud pública y la medicina comunitaria. Ningún conjunto simple de reglas puede eliminar esta dificultad. Como escribió H. L. Mencken: «Siempre hay una solución fácil para cualquier problema humano: nítida, plausible y equivocada[30]». Desarrollar el mejor proceso para la formulación de políticas racionales en estas circunstancias, un proceso que es relevante tanto en el ámbito de la medicina clínica como de la salud pública, sigue siendo un reto considerable.

Bibliografía

1. Hill AB. The environment and disease: association or causation? *Proc R Soc Med*. 1965;58:295–300.
2. Jones FB. Saturday Evening Post, November 29; 1953.
3. Rose G. Sick individuals and sick populations. *Int J Epidemiol*. 1985;14:22–38.
4. Whelton PK. Epidemiology of hypertension. *Lancet*. 1994;344:101–106.
5. Chobanian A, et al. The seventh report of the Joint National Committee on prevention, detection, evaluation and treatment of high blood pressure: the JNC 7 report. *JAMA*. 2003;289:2560–2572.
6. Grady D, Herrington D, Bittner V, for the HERS Research Groupet al. Cardiovascular disease outcomes during 68 years of hormone therapy: heart and estrogen/progestin replacement study follow-up (HERS II). *JAMA*. 2002;288:49–57.
7. The Women's Health Initiative Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results The Women's Health Initiative randomized controlled trial. *JAMA*. 2002;288:321–333.
8. Grodstein F, Clarkson TB, Manson JE. Understanding the divergent data on postmenopausal hormone therapy. *N Engl J Med*. 2003;348:645–650.
9. Michels KB. Hormone replacement therapy in epidemiologic studies and randomized clinical trials—are we checkmate? *Epidemiology*. 2003;14:3–5.
10. Whittemore AS, McGuire V. Observational studies and randomized trials of hormone replacement therapy: what can we learn from them? *Epidemiology*. 2003;14:8–10.
11. Manson JE, Bassuk SS, Harman SM, et al. Postmenopausal hormone therapy: new questions and the case for new clinical trials. *Menopause*. 2006;13:139–147.
12. Samet JM, Schnatter R, Gibb H. Epidemiology and risk assessment. *Am J Epidemiol*. 1998;148:929–936.
13. National Research Council Committee on the Institutional Means for Assessment of Risks to Public Health. *Risk Assessment in the Federal Government: Managing the Process*.

Washington, DC: National Academy Press; 1983:21.
14. Auchincloss AH, Diez Roux AV, Dvonch JT, et al. Associations between recent exposure to ambient fine particular matter and blood pressure in the Multi-Ethnic Study of Atherosclerosis (MESA). *Environ Health Perspect*. 2008;116:486–491.
15. Cohen MA, Adar SD, Allen RW, et al. Approach to estimating participant pollutant exposures in the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). *Environ Sci Technol*. 2009;43(13):4687–4693.
16. Wertheimer N, Leeper E. Electrical wiring configurations and childhood cancer. *Am J Epidemiol*. 1979;109:273–284.
17. Kheifets L, Monroe J, Vergara X, et al. Occupational electromagnetic fields and leukemia and brain câncer: an update of two meta-analyses. *J Occup Environ Med*. 2008;50:677–688.
18. Kheifets L, Ahlbom A, Crespi CM, et al. Pooled analysis of recente studies on magnetic fields and childhood leukaemia. *Br J Cancer*. 2010;103:1128–1135.
19. Calvente I, Fernandez MF, Villalba J, et al. Exposure to electromagnetic fields (non-ionizing radiation) and its relationship with childhood leukemia: a systematic review. *Sci Total Environ*. 2010;408(16):3062–3069.
20. Bonassi S, Taioli E, Vermeulen R. Omics in population studies: a molecular epidemiology perspective. *Environ Mol Mutagen*. 2013;54(7):455–460.
21. Ecclesiastes 1:9.
22. Porta M. A Dictionary of Epidemiology. 5th ed. New York: Oxford University Press; 2008.
23. Glass GV. Primary, secondary and meta-analysis of research. *Educ Res*. 1976;5:3–8.
24. U.S. Department of Health, Education and Welfare. *Smoking and Health: Report of the Advisory Committee to the Surgeon General.* Washington, DC, Public Health Service; 1964.
25. Black B, Lilienfeld DE. Epidemiology proof in toxic tort litigation. *Fordham Law Rev*. 1984;52:732–785.
26. Frye v. United States, 293 F 1013.(D.C. Cir 1923).
27. Daubert v. Merrell Dow Pharmaceuticals, Inc, 113 S Ct 2786; 1993.
28. Green M, Freedman M, Gordis L. Reference guide on

epidemiology In:. Reference Manual on Scientific Evidence. Washington, DC: The National Academies Press; 2011:549.
29. National Research Council Committee on the Institutional Means for Assessment of Risks to Public Health. *Risk Assessment in the Federal Government: Managing the Process.* Washington, DC: National Academy Press; 1983:11.
30. Mencken HL. The divine afflatus. New York Evening Mail. November 16, 1917.(Essay reprinted in Mencken HL: *Prejudices,* series 2 New York, Alfred A Knopf; 1920.).

CAPÍTULO 20

Aspectos éticos y profesionales en epidemiología

Palabras clave

ética
consentimiento informado
autonomía individual
privacidad y confidencialidad
principio de beneficencia
principio de utilidad
acceso a los datos
raza y etnia
conflicto de intereses
riesgo voluntario e involuntario

Ningún hombre es una isla, completo en sí mismo;

todo hombre es un fragmento del continente, una parte de tierra firme…

La muerte de cualquier hombre me afecta, porque soy parte de la humanidad;

por eso nunca procures saber por quién doblan las campanas;

doblan por ti.

—John Donne, clérigo y poeta inglés (1572-1631), Meditación XVII

Objetivos de aprendizaje

- Describir las obligaciones éticas que los investigadores tienen con las personas que se ofrecen como voluntarios para participar en estudios epidemiológicos.

- Evaluar cómo la privacidad y confidencialidad de las historias clínicas están protegidas en los estudios epidemiológicos y cómo se rige el acceso a los datos epidemiológicos.
- Describir las implicaciones científicas y éticas a la hora de clasificar la raza y la etnicidad en los estudios epidemiológicos.
- Comentar los aspectos relacionados con los conflictos de intereses.
- Revisar cómo se interpretan y se comunican a la opinión pública los resultados de los estudios epidemiológicos.

En los versos citados al comienzo de este capítulo, John Donne subraya la interconexión de todas las personas. La epidemiología también nos enseña lecciones importantes acerca de las conexiones y relaciones. En los capítulos precedentes se ha demostrado que las enfermedades no surgen en el vacío. Muchas enfermedades contagiosas dependen claramente de los contactos humanos (y frecuentemente con animales) para la transmisión y para la propagación de epidemias. Por otra parte, en los últimos años, un número creciente de enfermedades, que durante mucho tiempo se pensaba que no tenían una etiología infecciosa, se están identificando como de origen infeccioso en diversos grados. Por ejemplo, el microorganismo *Helicobacter pylori* se ha implicado en la etiología de la úlcera péptica y el cáncer gástrico (v. cap. 14). Muchos casos de cáncer de cuello uterino están relacionados con el virus del papiloma humano (VPH), sobre todo los tipos 16 y 18, por lo que existe una justificación para fomentar programas de prevención mediante la vacunación contra el VPH entre las jóvenes para la prevención de por vida.

Por otra parte, uno de los principales objetivos de la epidemiología es el impacto del ambiente sobre el riesgo de las enfermedades humanas. Esto refleja una combinación de factores: en primer lugar, estamos expuestos al riesgo derivado de los efectos de la naturaleza, incluidas las inundaciones y otros desastres naturales, como el tsunami que afectó a Japón en 2011, causando daños masivos en la planta de energía nuclear de Fukushima Daiichi; el huracán María en 2017, que devastó Puerto Rico y las Islas Vírgenes de los Estados Unidos; o el terremoto en la Ciudad de México, que

mató a unos 2.000 habitantes en 2017. En segundo lugar, también somos vulnerables al daño ambiental y ecológico secundario a ciertas actitudes, estilos de vida y conductas humanas. Los efectos negativos que las actividades humanas tienen sobre nuestro planeta a menudo no se tienen en cuenta adecuadamente, y algunos legisladores influyentes lamentablemente niegan su existencia. Estas actividades y efectos son, entre otros, la contaminación del aire, la desaparición de la capa de ozono, el calentamiento global, el cambio climático, la contaminación de los suministros naturales de agua, la deforestación y el desarrollo desenfrenado, entre muchos otros. Los efectos negativos de muchos de estos problemas solo ahora están comenzando a apreciarse y comprenderse por completo, así como el legado de los daños ambientales que se deja a las generaciones futuras. A medida que se estudian estos efectos, también es necesaria una mayor comprensión de las variaciones individuales de la vulnerabilidad humana a los agentes ambientales determinada genéticamente.

Otro aspecto de la interdependencia con relevancia para los epidemiólogos es su necesidad de desarrollar relaciones de colaboración con otros epidemiólogos y con profesionales de otros campos. Hemos aprendido que muchas investigaciones epidemiológicas requieren enfoques multidisciplinares, por lo que, desde el punto de vista profesional, los epidemiólogos no pueden ser más productivos y eficaces actuando como «islas». Por tanto, la lección de la «conectividad» expresada en los versos de John Donne parece un elemento integral tanto para la dinámica de las enfermedades y las situaciones investigadas por los epidemiólogos como para la práctica diaria de la epidemiología. También se aplica a la participación de los epidemiólogos a la hora de formular e implementar políticas relacionadas con la salud, como lo demuestra la historia de Semmelweis descrita en el capítulo 1.

Hoy en día, vivimos en una época despersonalizada en la que los individuos a menudo consideran que su propio progreso es el principal objetivo en la vida. A menudo se pierde el sentido de comunidad y la preocupación por los demás. La visión del mundo de John Donne que subrayaba la interdependencia de las personas a veces parece ajena a algunos puntos de vista actuales sobre el mundo, uno de los cuales se ve en clave de humor en la figura 20.1. Una de las mejores expresiones de la necesidad de equilibrar simultáneamente los intereses y necesidades contrapuestos del individuo y de la comunidad fue escrita por Hillel, un sabio talmúdico que vivió hace unos 2.000 años. Él dijo: «Si yo no soy para mí mismo [si yo no me cuido a mí mismo], ¿quién será para mí?, pero si solo soy para mí [es decir, si solo me ocupo de mí mismo], ¿qué valgo? Y si no es ahora, ¿cuándo?»

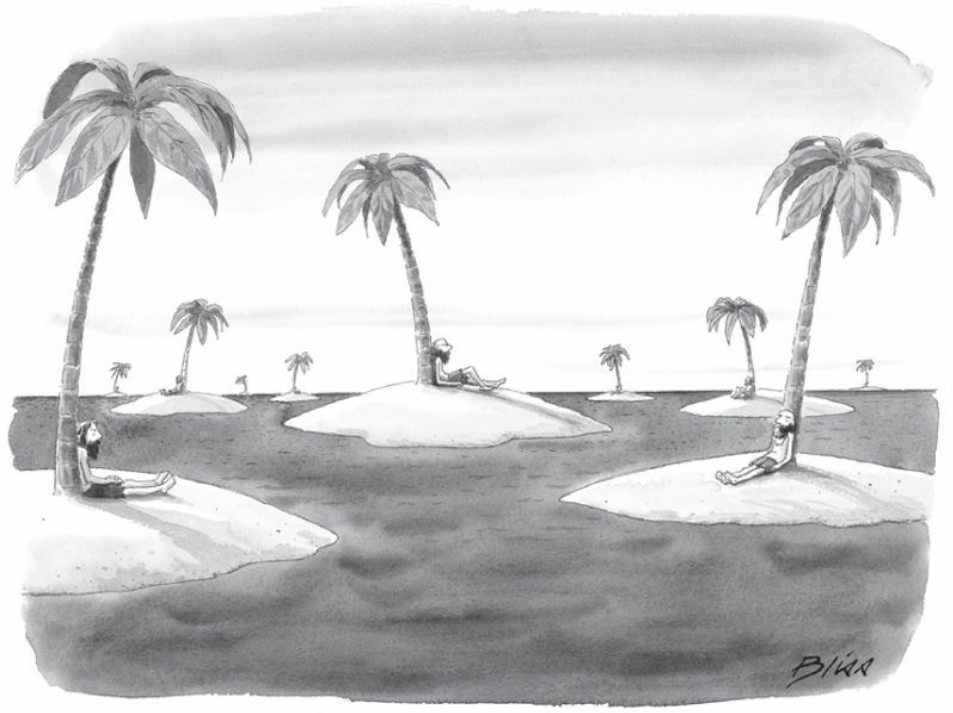


FIG. 20.1 «Ningún hombre es una isla»: un punto de vista diferente. (Harry Bliss/The New Yorker Collection/The Cartoon Bank.)

Otro factor que incide en la epidemiología y los epidemiólogos es el rápido ritmo del cambio social y del progreso tecnológico. En una de las historias que se cuentan de Adán y Eva en el Jardín del Edén, después de ser expulsado del paraíso, Adán se vuelve a Eva y le dice: «Eva, querida, estamos viviendo en una época de cambios[1]». En el siglo XXI, nosotros también vivimos en un periodo de cambios dramáticos. El contexto social y científico en rápida evolución en el que se lleva a cabo la investigación epidemiológica ha dado lugar a nuevos retos para quienes trabajan en epidemiología, para los que utilizan los resultados de los estudios epidemiológicos y para la ciudadanía. Además, los principales avances tecnológicos, incluidos los enormes incrementos de la capacidad de computación y los avances espectaculares en la tecnología de laboratorio (por ejemplo, recuerde el capítulo 16 sobre los avances significativos en genética en los últimos años), han permitido analizar rápidamente grandes cantidades de muestras biológicas y mantener conjuntos de datos muy grandes. Estos avances han hecho posibles muchos estudios poblacionales que no hubieran sido concebibles hace una década o dos. El registro médico electrónico (RME) está sustituyendo a las historias clínicas de papel en los hospitales y en él se incluyen consultas ambulatorias y consultas hospitalarias, resultados de exámenes de laboratorio, electrocardiogramas (ECG) y tomografías computarizadas/resonancias magnéticas (TC/RM). Cómo usar estos datos (que no se han recopilado con fines de investigación) para estudios epidemiológicos es un desafío importante para el análisis de «macrodatos». Al mismo tiempo, estos avances tecnológicos han introducido problemas nuevos y diferentes relacionados con la privacidad, la confidencialidad y el

individuo.

A la luz de las consideraciones anteriores, en este capítulo se revisan brevemente algunos problemas éticos y profesionales que son cruciales para la investigación epidemiológica y para la aplicación de los resultados de esta investigación con el fin de mejorar la salud humana. Los temas que se describirán incluyen varios referentes a la realización real de los estudios epidemiológicos y otros que están relacionados con aspectos sociales más amplios y que van más allá de la propia investigación epidemiológica real.

Problemas éticos de la epidemiología

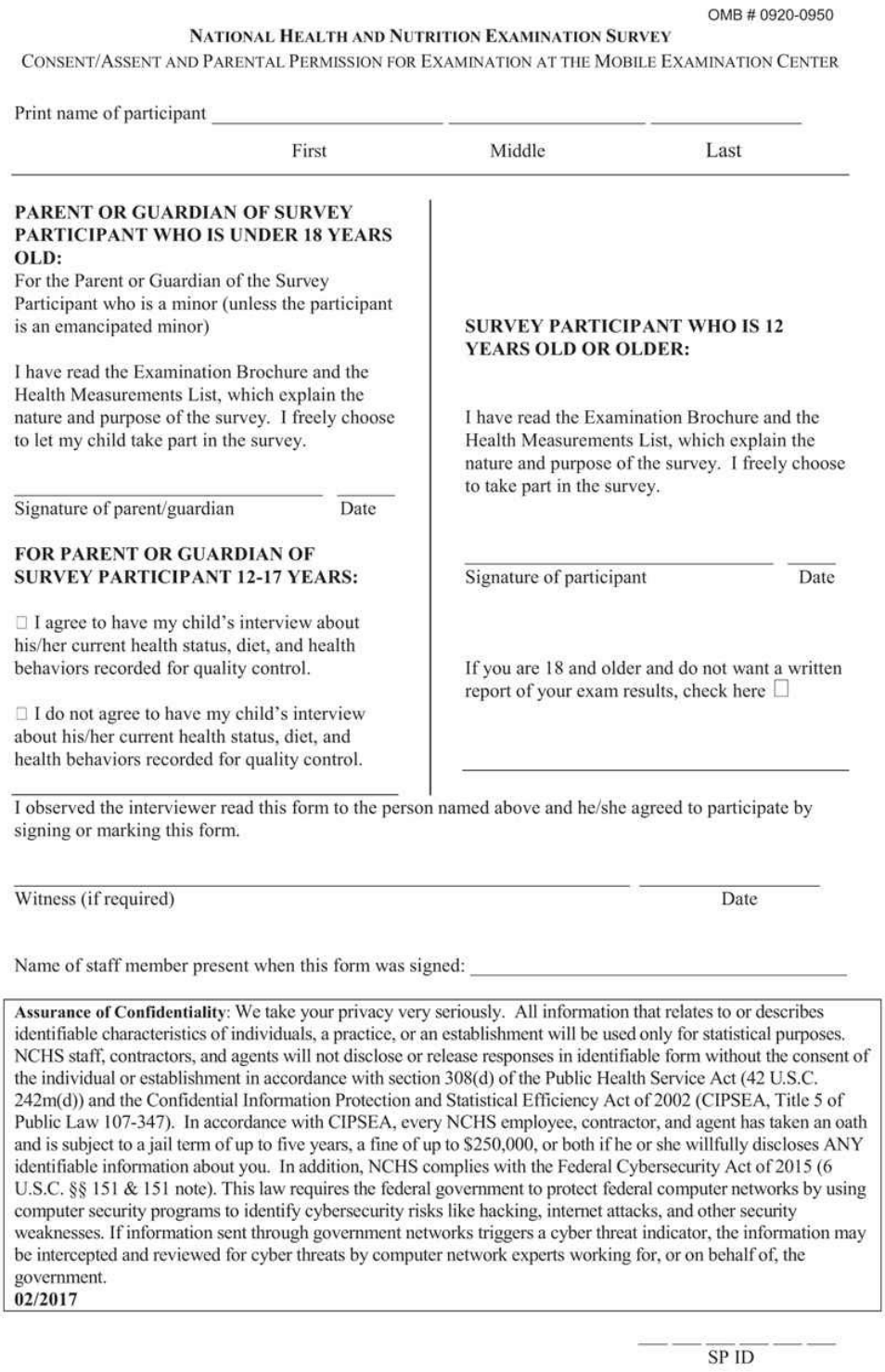
Es evidente que en cualquier actividad científica el fraude, el engaño o la tergiversación suscitan una desaprobación universal y condena por parte de los miembros de la disciplina, por otros profesionales y por la ciudadanía. Estas cuestiones no se tratan en este capítulo. Hoy en día, es probable que algunos de los dilemas éticos más difíciles de la epidemiología sean más sutiles, que impliquen juicios, filosofías, actitudes y opiniones, donde puede que el consenso sea más difícil de obtener.

¿Difiere la epidemiología de otras disciplinas científicas y médicas con respecto a las cuestiones éticas? Aunque la epidemiología comparte muchas características con otras disciplinas científicas, difiere en algunos aspectos importantes. Es una disciplina que surgió en gran medida a partir de la medicina y la salud pública, e, incluso en sus primeros años, sus conclusiones tuvieron implicaciones políticas inmediatas para la asistencia clínica o las acciones en salud pública. Los estudios de John Snow sobre el cólera en Londres (v. cap. 1) y su eliminación de la palanca de la bomba de agua de Broad Street, que sus estudios habían implicado en el brote (si la palanca de la bomba se retiró en realidad antes o después del pico del brote), reflejaban las claras implicaciones políticas de sus hallazgos.

El objetivo último de la epidemiología es mejorar la salud humana; la epidemiología es la ciencia básica de la prevención de enfermedades. Por tanto, la relación de la epidemiología con el desarrollo de las políticas públicas es parte integral de la disciplina. Como resultado, los problemas éticos y profesionales van más allá de los que podrían aplicarse a una disciplina científica, como la biofísica o la fisiología, y deben verse en un contexto más amplio. En primer lugar, los hallazgos epidemiológicos tienen relevancia social directa y a menudo inmediata. En segundo lugar, los estudios epidemiológicos suelen financiarse con recursos públicos y a menudo tienen implicaciones importantes para la asignación de recursos sociales limitados. En tercer lugar, la investigación epidemiológica con frecuencia implica a seres humanos de alguna manera, y las personas que participan en los estudios epidemiológicos no suelen obtener ningún beneficio personal por participar en estos estudios o de sus resultados.

Obligaciones de los investigadores con los sujetos de estudio

¿Cuáles son las obligaciones de los investigadores con los sujetos en los estudios observacionales no aleatorizados, que son los que llevan a cabo la mayoría de los epidemiólogos? En primer lugar, en la medida de lo posible, se debe obtener un consentimiento verdaderamente informado de cada persona, concordante con el principio de autonomía individual. Aquí surge la duda de si se puede obtener un consentimiento verdaderamente informado de un sujeto en un estudio epidemiológico. Si se piensa que la divulgación completa de los objetivos y las hipótesis del estudio a los sujetos puede introducir un sesgo de respuesta o de otro tipo, es evidente que el consentimiento no puede ser completamente «informado». Otro problema en el consentimiento hace referencia a la intimidad y a la confidencialidad. Durante muchos años, de buena fe, los epidemiólogos aseguraban a los sujetos que sus datos se mantendrían en secreto y que este compromiso era incondicional. Sin embargo, los datos de las investigaciones han pasado a estar sometidos a citación judicial en los últimos años, con solo algunas excepciones. Por tanto, la garantía de confidencialidad que figura en las declaraciones de consentimiento informado ahora debe incluir restricciones que permitan un incumplimiento de la confidencialidad que podría estar regulado por ley y, por tanto, escaparía al control del investigador. En 2003 entraron en vigor en EE. UU. nuevas normativas de privacidad que afectan significativamente a los derechos de los pacientes sobre la información de la salud (fig. 20.2). Demasiado a menudo recibimos noticias sobre violaciones de datos y la divulgación de información confidencial que se puede emplear para dañar a los participantes de una investigación o causar daño a sus reputaciones y cuentas bancarias o que involucran otros tipos de fraude. Volveremos al tema de la privacidad y la confidencialidad más adelante en este capítulo.



OMB # 0920-0950

NATIONAL HEALTH AND NUTRITION EXAMINATION SURVEY

CONSENT/ASSENT AND PARENTAL PERMISSION FOR EXAMINATION AT THE MOBILE EXAMINATION CENTER

Print name of participant _____ _____ _____

First Middle Last

PARENT OR GUARDIAN OF SURVEY PARTICIPANT WHO IS UNDER 18 YEARS OLD:

For the Parent or Guardian of the Survey Participant who is a minor (unless the participant is an emancipated minor)

I have read the Examination Brochure and the Health Measurements List, which explain the nature and purpose of the survey. I freely choose to let my child take part in the survey.

_____ _____

Signature of parent/guardian Date

FOR PARENT OR GUARDIAN OF SURVEY PARTICIPANT 12-17 YEARS:

☐ I agree to have my child's interview about his/her current health status, diet, and health behaviors recorded for quality control.

☐ I do not agree to have my child's interview about his/her current health status, diet, and health behaviors recorded for quality control.

SURVEY PARTICIPANT WHO IS 12 YEARS OLD OR OLDER:

I have read the Examination Brochure and the Health Measurements List, which explain the nature and purpose of the survey. I freely choose to take part in the survey.

_____ _____

Signature of participant Date

If you are 18 and older and do not want a written report of your exam results, check here ☐

I observed the interviewer read this form to the person named above and he/she agreed to participate by signing or marking this form.

_____ _____

Witness (if required) Date

Name of staff member present when this form was signed: _____

Assurance of Confidentiality: We take your privacy very seriously. All information that relates to or describes identifiable characteristics of individuals, a practice, or an establishment will be used only for statistical purposes. NCHS staff, contractors, and agents will not disclose or release responses in identifiable form without the consent of the individual or establishment in accordance with section 308(d) of the Public Health Service Act (42 U.S.C. 242m(d)) and the Confidential Information Protection and Statistical Efficiency Act of 2002 (CIPSEA, Title 5 of Public Law 107-347). In accordance with CIPSEA, every NCHS employee, contractor, and agent has taken an oath and is subject to a jail term of up to five years, a fine of up to $250,000, or both if he or she willfully discloses ANY identifiable information about you. In addition, NCHS complies with the Federal Cybersecurity Act of 2015 (6 U.S.C. §§ 151 & 151 note). This law requires the federal government to protect federal computer networks by using computer security programs to identify cybersecurity risks like hacking, internet attacks, and other security weaknesses. If information sent through government networks triggers a cyber threat indicator, the information may be intercepted and reviewed for cyber threats by computer network experts working for, or on behalf of, the government.

02/2017

_ _ _ _ _ _

SP ID

FIG. 20.2 Ejemplo de documento de consentimiento (National Health and Nutrition Examination Survey [NHANES] Consent/Assent and Parental Permission for Examination at the Mobile Examination Center. De Centers for Disease Control and Prevention https://wwwn.cdc.gov/nchs/data/nhanes/2017-2018/documents/2017_adult_consent_form.pdf. Consultado el 2 de noviembre de 2018.)

Otra cuestión es la de establecer un equilibrio entre los derechos del individuo y el bienestar de la sociedad. En uno de los primeros estudios sobre varones con alto riesgo de infección por el virus de la inmunodeficiencia humana (VIH) se aseguró la confidencialidad a los participantes. En la

primera entrevista que se realizó a continuación se preguntó a los sujetos si habían donado sangre durante los 2 años anteriores. Varios sujetos que resultaron ser VIH- positivos refirieron haber donado sangre en los 2 años anteriores a la prueba del VIH, lo que suscitó el temor de que la sangre donada podría haber sido utilizado en una transfusión. Aunque la sangre pudo haber sido descartada por el banco de sangre, no había ninguna manera de comprobarlo sin violar la confidencialidad y el compromiso inicial con los sujetos. Tal vez los investigadores deberían haber previsto este problema cuando se realizó la entrevista antes de obtener el consentimiento informado de los sujetos. Pero, incluso cuando se prevén estas cuestiones, surgen problemas de este tipo. En este caso, ¿cómo se puede establecer un equilibrio entre el compromiso original con los sujetos y la necesidad de determinar si alguien había recibido sangre de estos donantes, de modo que pudiese prevenirse la transmisión del VIH?

Una tercera obligación con los sujetos es la de comunicar los resultados del estudio a dichos sujetos al finalizar el estudio. Nuestra estrategia respecto a esta cuestión puede ser diferente dependiendo de si se ha observado que el sujeto ha desarrollado un problema de salud relacionado con una exposición que se está estudiando o si solo se ha observado que tiene un mayor riesgo de desarrollar en el futuro la enfermedad debido a la exposición. En cualquier caso, la comunicación clara y concisa a los sujetos de los resultados referentes a los riesgos puede considerarse una posible expresión del principio ético de beneficencia (la obligación del investigador de ayudar a los sujetos a promover sus intereses legítimos importantes, como la prevención y control de enfermedades, para ellos mismos y para sus familias y amigos). Sin embargo, de acuerdo con este principio, no solo tenemos que proporcionar beneficios como la prevención de la enfermedad, sino que también hay que equilibrar los beneficios y los costes o perjuicios (principio de utilidad).

Si, por ejemplo, un sujeto ha estado expuesto a un factor del que se demuestra en un estudio que es un marcador significativo de la fase subclínica de una neoplasia maligna, ¿se debe ofrecer esta información al sujeto? Por otra parte, dado que no se dispone de ningún tratamiento eficaz para dicha neoplasia y que no hay una evidencia sólida de que la detección temprana de la enfermedad sea beneficiosa, ¿podríamos estar aumentando la ansiedad de una persona al ofrecer esta información sin proporcionarle ningún beneficio?

Por otra parte, se podría argumentar que un participante en cualquier estudio tiene derecho a recibir las conclusiones obtenidas de este, aunque los resultados no tengan relación directa con la salud de la persona, o incluso si puede causar una mayor ansiedad. ¿Por qué tendríamos los investigadores que tomar esta decisión en nombre de los participantes? De hecho, muchos epidemiólogos ofrecen ahora a todos los sujetos participantes la opción de solicitar un informe de los resultados del estudio cuando este se ha completado.

Protección de la privacidad y la confidencialidad

Las preocupaciones acerca de la privacidad y la confidencialidad en nuestra sociedad han aumentado con la creciente erosión de la privacidad individual debido a los registros informatizados. La protección de la privacidad y la confidencialidad en el marco de la investigación médica, incluidas las investigaciones epidemiológicas, se ha convertido en una cuestión importante. Los orígenes de estas preocupaciones son bastante antiguos. Hipócrates escribió en el Juramento Médico que actualmente es de uso general lo siguiente:

Guardaré secreto sobre lo que oiga y vea en la sociedad por razón de mi ejercicio y que no sea indispensable divulgar, sea o no del dominio de mi profesión, considerando como un deber el ser discreto en tales casos.

Dado que Hipócrates calificó «lo que oiga y vea» con la frase «que no sea indispensable divulgar», aparentemente consideraba que algunos tipos de información *sí eran* «apropiados para divulgar». Es de suponer entonces que, bajo ciertas circunstancias, Hipócrates habría defendido el intercambio cuidadosamente controlado de la información personal en interés del beneficio social. Por ejemplo, si se diagnosticara un caso de viruela en una ciudad americana, Hipócrates probablemente habría apoyado la notificación de este caso a las autoridades sanitarias. Por tanto, la autonomía individual respecto a la privacidad y la confidencialidad es un principio importante, pero no es ilimitado.

En lo que se refiere a la privacidad y la confidencialidad en los estudios epidemiológicos, la atención se ha centrado en el uso de las historias clínicas (tanto las antiguas en papel como los más comunes RME en la actualidad). En primer lugar, hay que aclarar que las historias clínicas se necesitan en los estudios epidemiológicos para dos propósitos principales: 1) generar datos agregados o validar la información obtenida por otros medios sin tener que contactar a los pacientes y/o 2) identificar a pacientes individuales para seguimiento posterior utilizando métodos como entrevistas o pruebas de laboratorio.

Debido a que los objetivos de la epidemiología de mejorar la salud humana son evidentemente loables, podría surgir la tentación a primera vista de descartar cualquier preocupación sobre el mal uso de los datos de las historias clínicas y sobre las intrusiones en la privacidad individual por los epidemiólogos. Sin embargo, las palabras del magistrado del Tribunal Supremo de EE. UU. Louis D. Brandeis son tan ciertas hoy como cuando

fueron escritas por primera vez en 1928:

La experiencia nos debería enseñar a estar más en guardia que nunca para proteger la libertad cuando las intenciones del gobierno son benéficas. Los hombres nacidos en libertad están naturalmente alerta para repeler la invasión de su libertad por gobernantes perversos. Los mayores peligros para la libertad se ocultan en las solapadas transgresiones de los hombres de celo, bien intencionados, pero sin entendimiento[2].

El principio ético de autonomía es un argumento de peso para ofrecer un consentimiento informado significativo en muchas áreas relacionadas con la investigación, incluidas la privacidad y la confidencialidad. Las preocupaciones por la protección de la confidencialidad en el ámbito de la investigación son válidas. A lo largo de los años, estas preocupaciones han dado lugar a dos propuestas legislativas importantes que parecen razonables al principio, pero que en realidad podrían perjudicar seriamente la investigación epidemiológica e impedir el progreso tanto en el campo de la salud pública como en la práctica clínica. Estas dos propuestas son las siguientes:

1. Se debe exigir el consentimiento del paciente antes de permitir a los investigadores el acceso a las historias clínicas.
2. Los datos de las historias clínicas deberían ser puestos a disposición de los investigadores sin ningún tipo de información que pueda identificar a un individuo.

Ambas propuestas están en consonancia con el principio ético de *no maleficencia* (no hacer daño) a los sujetos que participan en un estudio de investigación. Sin embargo, si la sociedad tiene un interés personal en los resultados de estudios epidemiológicos y de otros campos biomédicos, es necesario encontrar un equilibrio entre los intereses del individuo y los de la comunidad en general.

Consideremos cada una de estas dos propuestas por separado. ¿Por qué la primera propuesta, que requiere el consentimiento del paciente antes de que a los investigadores se les permita el acceso a la historia clínica, impide la realización de muchos estudios?

- En un primer paso de un estudio, a menudo se deben revisar las historias para identificar qué pacientes cumplen los criterios de reclutamiento (p. ej., qué pacientes tienen la enfermedad en cuestión y, por tanto, son elegibles para su inclusión en un estudio de casos y controles).

- Muchos estudios epidemiológicos no se diseñan hasta muchos años después de que un paciente fuera hospitalizado (p.ej., puede disponerse de una nueva prueba que no se utilizaba cuando el paciente se encontraba enfermo), por lo que no podría haberse obtenido el consentimiento informado del paciente en ese momento. Cuando, posteriormente, el estudio se lleva a cabo, lo que puede ser años o décadas después, muchos pacientes pueden haber fallecido o no pueden ser localizados.
- Algunos pacientes rechazan ser entrevistados en los estudios epidemiológicos, pero los no participantes pueden caracterizarse a partir de datos de sus historias clínicas para que se puedan evaluar los sesgos derivados de su falta de participación. Si las historias clínicas no estuviesen disponibles debido a la negativa del paciente, se introduciría un posible sesgo de selección, sin que se pudiese evaluar su magnitud y su dirección.

En cuanto a la segunda propuesta, ¿por qué la información de las historias clínicas que identifica a pacientes individuales es esencial para la mayoría de los estudios epidemiológicos?

- La revisión de las historias clínicas es a menudo el primer paso en la identificación de un grupo suficientemente grande de personas con una enfermedad a quienes se realizará un seguimiento posterior.
- La identificación de la información es esencial para vincular los registros de determinados individuos obtenidos de diferentes fuentes (como los registros hospitalarios, los registros médicos, los registros laborales y los certificados de defunción en los estudios de cáncer profesional).

Como se observa en la figura 20.3, la vinculación de los registros es crucial para la elaboración de información no sesgada y completa sobre cada sujeto, no solo en estudios ocupacionales (como se muestra aquí), sino también en muchos tipos de investigaciones epidemiológicas. Un ejemplo es un estudio de cohortes no concurrente (v. cap. 8) para evaluar la relación entre el estado del receptor de estrógeno en el cáncer de mama con la mortalidad mediante el uso de datos de historias clínicas y la vinculación con el Índice Nacional de Muerte.

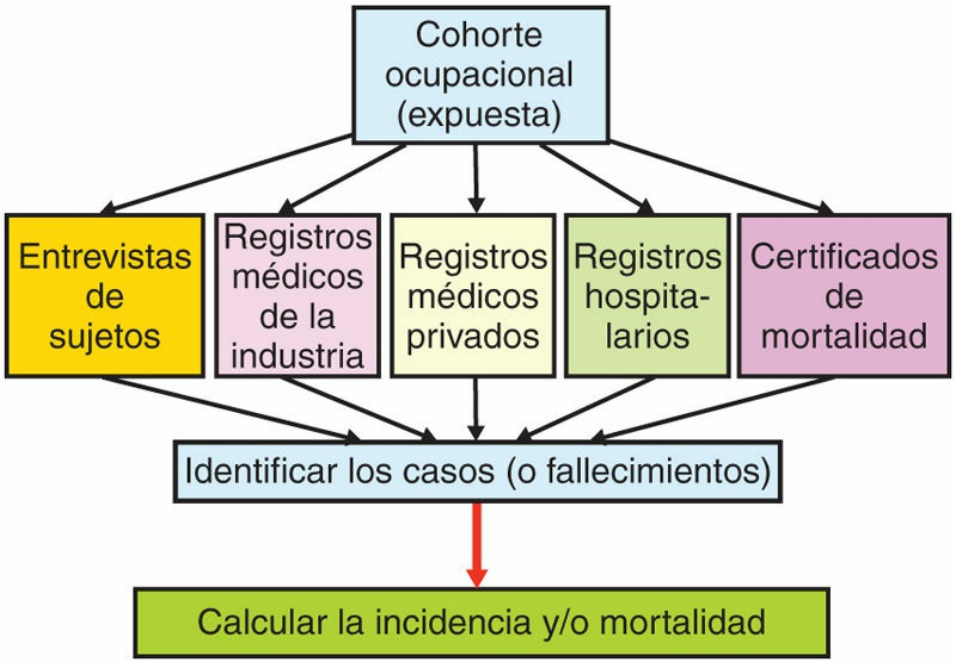


FIG. 20.3 Uso de los vínculos entre registros en los estudios ocupacionales.

Por tanto, se observa que el uso de las historias clínicas frecuentemente es esencial para los estudios epidemiológicos. De hecho, muchos de los avances significativos en la protección de la salud humana resultantes de la investigación epidemiológica no podrían haberse realizado si el acceso a las historias clínicas se hubiese restringido[3]. Al mismo tiempo, sin embargo, hay que mantener la vigilancia sobre la protección de la privacidad individual y la confidencialidad. Durante muchos años, los estudios epidemiológicos han utilizado los siguientes procedimientos diseñados para proteger la confidencialidad de los sujetos:

- Se requiere el consentimiento informado de los participantes del estudio para todas las fases de la investigación, salvo la revisión de las historias clínicas. El lenguaje del consentimiento informado debe tener en cuenta el nivel educativo de los participantes (generalmente establecido en un nivel de comprensión de octavo grado).
- Todos los datos obtenidos se guardan bajo llave.
- Solo se utilizan los números de estudio en los formularios de datos. No aparecen identificadores personales en los formularios de datos ni en los archivos informáticos. A los analistas solo se les proporcionan datos no identificados para el análisis. La clave para vincular los números de estudio con los nombres individuales se guarda por separado bajo llave por el investigador principal o su representante (coordinador del estudio).
- La información de identificación individual se destruye al final del estudio a menos que exista una justificación específica para conservarla. Dicha conservación debe ser aprobada por la junta de revisión institucional o por el comité de investigación humana.

- Todos los resultados se publican solo en forma agregada o de grupo para que los individuos nunca sean identificados.
- A menos que sea esencial para el estudio, la información de identificación individual no se introduce en los archivos informáticos, y los identificadores individuales (como las ubicaciones geográficas de área pequeña) no se incluyen en las tabulaciones rutinarias generadas a partir de los datos informatizados.
- Se insiste de forma periódica al personal de investigación sobre la importancia de mantener la privacidad y la confidencialidad.

Cuando las personas consienten en participar en los estudios epidemiológicos, aceptan voluntariamente una cierta invasión de su privacidad por el bien común de la sociedad, con la esperanza de que se logren avances en la promoción de la salud y la prevención de enfermedades como resultado de los estudios que están haciendo posibles. Por tanto, los investigadores tienen la obligación ética de proteger la privacidad y confidencialidad de los sujetos en estos estudios en la medida de lo posible. Las políticas descritas anteriormente que se encuentran en vigor en la actualidad han tenido un gran éxito en la consecución de este objetivo.

Consciente de la importancia del uso de las historias clínicas en la investigación epidemiológica y de la eficacia de las medidas actuales para proteger la privacidad y la confidencialidad, la Privacy Protection Study Commission estadounidense recomienda que no sea necesario el consentimiento del paciente para el uso de historias clínicas en la investigación epidemiológica[4]. Sin embargo, el 14 de abril de 2003, el panorama cambió drásticamente en EE. UU., cuando entraron en vigor las nuevas normativas federales sobre privacidad en virtud de la Ley de Transferibilidad y Responsabilidad de Seguros de Salud *(Health Insurance Portability and Accountability Act, HIPAA)*[5] de 1996. La ley fue introducida en respuesta a la creciente preocupación de la opinión pública por la falta de control individual sobre la información médica y la erosión general de la privacidad individual en EE. UU. La transferencia electrónica de información médica y los temores sobre el posible uso indebido de la información genética disponible gracias a los nuevos métodos de laboratorio también dieron lugar al desarrollo de estas nuevas normativas.

Las normativas de la HIPAA constituyen la primera protección sistemática de la privacidad a nivel nacional de la información médica en Estados Unidos. Las normativas otorgan un mayor control a los pacientes sobre su información médica y establecen límites para el uso y divulgación de las historias clínicas. Con algunas excepciones, en la actualidad se requiere la autorización firmada de cada individuo para la divulgación de su información médica protegida. Dicha información médica protegida puede ser revelada a las autoridades de salud pública sin autorización individual con fines de salud pública, incluidas, entre otras, la vigilancia, las investigaciones y las intervenciones de salud pública. La información médica

protegida también puede divulgarse para la investigación médica sin autorización individual bajo ciertas condiciones, incluidas las siguientes: 1) si una junta de revisión institucional ha proporcionado una exención, 2) para actividades preparatorias del inicio de la investigación y 3) para la investigación sobre la información de un fallecido[6]. Las normativas son muy complejas y se requerirá tiempo antes de que se pueda valorar todo el impacto de las nuevas normativas sobre las investigaciones y actividades clínicas y de salud pública, así como sobre la investigación epidemiológica. Se han publicado descripciones extensas de estas normativas[7-9].

Acceso a los datos

Cuando un estudio se ha completado, ¿a quién «pertenecen» los datos? ¿Quién debe tener acceso a los datos, ya sea «en bruto» o parcialmente «elaborados», y en qué condiciones? Vivimos en una era en la que podemos estar seguros de que prácticamente todos los datos de investigación generados que tratan un tema polémico se volverán a analizar por expertos reales o supuestos que apoyen diferentes posiciones. Algunas de las preguntas pertinentes en relación con el intercambio de datos son las siguientes:

- ¿En qué momento se ha completado realmente un estudio?
- ¿La política en materia de intercambio de los datos de investigación debería depender de quién ha financiado el estudio?
- ¿La política debería depender de quién solicita los datos y de las posibles motivaciones de esa persona para hacer la solicitud?
- ¿Bajo qué condiciones deben incluirse elementos identificadores de los participantes individuales con los datos? En general, el intercambio de datos requiere eliminar de ellos todos los identificadores de los participantes individuales.
- ¿Cómo pueden protegerse los intereses del investigador?
- ¿Se puede acceder a todos los datos o solo están disponibles los datos resumidos?
- ¿Cómo de extenso debe ser el intercambio de datos? ¿Debe limitarse a las variables solicitadas o debe compartirse el conjunto de datos completo?
- ¿Puede solicitar alguien los datos o son de acceso restringido?
- ¿Quién pagará los gastos involucrados en compartir los datos?

La dificultad consiste en encontrar un equilibrio adecuado entre los intereses del investigador, por un lado, y los de la sociedad, por otro, porque no siempre coinciden.

Raza y etnicidad en los estudios epidemiológicos

Una cuestión destacada que ha recibido una atención creciente en los últimos años es el uso de las denominaciones de raza y etnicidad en los estudios epidemiológicos. Estas variables se utilizan tanto para describir las poblaciones como para probar las hipótesis en las que la raza puede actuar como una variable independiente, como en los numerosos estudios epidemiológicos sobre las disparidades raciales en el estado de salud. Una revisión de PubMed en octubre de 2017 de los términos de búsqueda *raza, etnicidad, salud* generó 17.603 publicaciones y de *disparidades por etnicidad racial* dio lugar a 12.001 registros. Claramente, la raza y la etnicidad son el foco de muchas investigaciones médicas y epidemiológicas. Sin embargo, existe preocupación por el uso de estos términos en la investigación epidemiológica.

En un importante artículo publicado en 2008, Jay Kaufman y Richard Cooper estudian el concepto de raza como se usa en epidemiología y proponen una serie de precauciones en su uso. La raza y la etnicidad no se pueden tratar como variables dicotómicas, ya que «la variación humana es continua, no discreta» y, en última instancia, «los esquemas de clasificación racial existentes son el resultado de procesos históricos y políticos y no tienen nada de natural, objetivo o científico»[10].

La raza se utiliza a menudo como descriptor para caracterizar a los individuos que se estudian en ensayos clínicos o para describir las inclusiones y exclusiones de las poblaciones en los diferentes tipos de estudios epidemiológicos. La raza y el origen étnico, empleados como este tipo de variable, pueden ser útiles para este propósito y pueden ser importantes para evaluar la posibilidad de generalizar los resultados más allá de la población estudiada.

Cuando las variables que designan la raza o la etnicidad se incluyen en estudios diseñados para probar las hipótesis, la atención se centra a menudo en las posibles asociaciones de la raza con ciertos resultados de salud. Sin embargo, tal y como han señalado Bhopal y Donaldson[11], desde el punto de vista biológico, la raza está mal definida y mal comprendida, y puede tener una validez dudosa. La investigación del ADN indica que la diversidad genética es un continuum sin interrupciones claras que puedan delimitar los grupos raciales[12]. La raza se ha descrito como «un sistema arbitrario de clasificación visual» que no demarca subgrupos distintos de la población humana[13]. A partir del censo estadounidense de 2000, unas directrices nuevas permiten a los encuestados que se identifiquen con más de un grupo racial. En el futuro, esta política puede complicar el uso de los datos censales sobre la raza en los estudios epidemiológicos.

Una estrategia alternativa es usar la etnicidad en lugar de la raza. Sin embargo, la clasificación de las personas por la etnicidad tampoco es sencilla

o directa. La etnicidad es una variable compleja que implica orígenes o trasfondos sociales compartidos, una cultura y tradiciones compartidas que son distintivas y que se mantienen entre generaciones, dando lugar a un sentimiento de identidad y de grupo, o bien una lengua o tradición religiosa comunes[14]. ¿Cuáles han sido los resultados de utilizar designaciones raciales en la investigación epidemiológica? Muchos autores creen que, dadas las ambigüedades implicadas en la definición de la raza, las investigaciones que utilizan las tasas de enfermedades en función de la raza no han aumentado significativamente nuestra comprensión básica de las causas y la patogenia de las enfermedades humanas[15]. Sin embargo, hay quien ha argumentado que, aunque dichas designaciones no han mejorado nuestra comprensión de los mecanismos biológicos de la enfermedad, el uso de variables raciales en la investigación ha ayudado a identificar subgrupos (sobre todo minorías e inmigrantes) a los que se deben dirigir más recursos de asistencia sanitaria. Por ejemplo, los datos de mortalidad específica de cada raza en EE. UU. han puesto de manifiesto que[16]:

- Un niño afroamericano tiene más del doble de probabilidades que un niño de raza blanca de fallecer durante el primer año de vida.
- Las personas de raza negra tienen más probabilidades de padecer una enfermedad renal terminal, pero menos probabilidades de recibir trasplantes de riñón que las personas de raza blanca.
- Las personas de raza negra son más propensas a sufrir hipertensión.
- Las tasas de mortalidad para la mayoría de las causas de muerte son mucho mayores para personas afroamericanas que para personas de raza blanca.

En los estudios relacionados con las necesidades sanitarias y las prioridades de la asistencia sanitaria de diversas poblaciones, puede que se describa la raza de un grupo de población, que se realice una comparación explícita con otros grupos raciales y étnicos, o que exista una comparación implícita, pero que no se indique explícitamente. Las tasas de mortalidad en función de la raza se utilizan con frecuencia para establecer objetivos nacionales y estatales de salud. Los Centros para el Control y la Prevención de Enfermedades han señalado que «las tasas de mortalidad en función de la raza y el origen hispano son importantes para el seguimiento del estado de salud de estos grupos de población y para orientar las políticas y programas dirigidos a la reducción de las disparidades»[16].

Uno de los problemas de utilizar variables raciales es que, al hacerlo, incluso los investigadores bienintencionados pueden estigmatizar inadvertidamente a ciertos subgrupos de la población. Como resultado, algunas designaciones raciales pueden, en realidad, convertirse en sinónimos de características de estilo de vida no deseadas, como el comportamiento delictivo y el consumo de drogas. Como ha señalado Bhopal, «al hacer

hincapié en los aspectos negativos de la salud de los grupos étnicos minoritarios, la investigación puede haber dañado su posición social y haber desviado la atención de sus prioridades de salud»[15].

¿Qué conclusiones pueden extraerse? Ninguna variable, incluida la raza, debería utilizarse de manera acrítica de forma rutinaria en ningún estudio epidemiológico. Tal vez la mejor estrategia al planificar un estudio epidemiológico en el que se vaya a tener en cuenta la raza sea hacer una serie de preguntas, entre las que se encuentran las siguientes:

- ¿Por qué se estudia la raza?
- ¿Sobre qué base se clasifican los participantes del estudio en función de la raza?
- ¿Qué validez tendrán las denominaciones de raza y cómo van a contribuir a aumentar nuestro conocimiento biológico de la enfermedad en cuestión o a mejorar las actividades preventivas en ciertos grupos desfavorecidos?
- Si la raza se usa como un sustituto de ciertos factores del estilo de vida, como la dieta, ¿podría obtenerse directamente la información sobre la dieta y otros factores del estilo de vida, sin necesidad de utilizar la raza como sustituto?
- Al mismo tiempo, también deberíamos preguntarnos si puede causarse algún perjuicio por el uso de designaciones raciales en un estudio concreto y si tales designaciones pueden servir involuntariamente como sinónimos virtuales de estilos de vida o características indeseables.
- ¿La validez de construcción de ciertas variables es la misma en todas las razas o etnias? Por ejemplo, la educación se utiliza a menudo como indicador indirecto de la posición socioeconómica (PSE). ¿La educación es indicativa de la PSE en la misma medida en negros y blancos en los Estados Unidos?

En cualquier estudio, las variables raciales que se utilicen deben tener un propósito definido que se pueda expresar con precisión y deben cumplir con los mismos estándares de fiabilidad y validez que serían de esperar en cualquier otra variable que se mida en el estudio. El beneficio potencial de la utilización de tales variables en un estudio debe superar claramente cualquier posible perjuicio que pudiese producirse. La raza puede ser una variable apropiada y potencialmente útil a tener en cuenta en estudios epidemiológicos siempre que los temas anteriores se hayan considerado y abordado adecuadamente.

Jean-Claude Moubarac realizó una revisión exhaustiva del uso de la raza y el origen étnico en las publicaciones sobre disparidades de salud en epidemiología y salud pública[17]. Presentó una revisión de 280 artículos publicados entre 2009 y 2011 e identificó cuatro problemas principales pendientes. En primer lugar, los investigadores generalmente no

diferenciaban la raza de la etnicidad. En segundo lugar, los autores a menudo atribuían la etnicidad a las categorías raciales. En tercer lugar, los procedimientos de medición comunes que subyacen a cada concepto rara vez se especificaban claramente. Y en cuarto, hubo un fracaso general en atribuir límites a las taxonomías de raza y etnicidad comúnmente utilizadas. Por lo tanto, resulta claro que se necesita una precisión mucho mayor si en el futuro se quiere investigar adecuadamente la raza y la etnicidad en epidemiología.

Conflicto de intereses

Se pueden producir sesgos tanto reales como percibidos debido a un conflicto de intereses. Este conflicto puede surgir en cualquier etapa de un estudio, desde la decisión inicial de si un estudio específico debe realizarse en primer lugar hasta el análisis y la interpretación de los datos y la difusión de los resultados. La investigación epidemiológica en Estados Unidos en la actualidad se lleva a cabo por epidemiólogos que trabajan en el mundo académico, la industria o el gobierno. Estos tres entornos difieren en varios aspectos. Los fondos para la investigación epidemiológica en el gobierno y la industria suelen ser internos, mientras que los epidemiólogos académicos deben buscar apoyo financiero externo del gobierno, la industria o instituciones. Como resultado, la investigación realizada por epidemiólogos académicos suele someterse a una revisión por pares más rigurosa como parte del proceso de solicitud de financiación. Sin embargo, aún más importante es el hecho de que el empleador del epidemiólogo académico no suele tener ningún interés en cuáles serán los resultados del estudio. Esto contrasta con otros contextos en los que el empleador puede verse afectado significativamente (desde los puntos de vista político, económico o legal) por las características de los resultados de la investigación. En consecuencia, la presión franca o sutil por parte del empleador para no iniciar un estudio o para prolongar el proceso que culmina en la notificación de los resultados (o en su ocultación) puede introducir un sesgo grave en las revisiones de la literatura referente a cuestiones como los riesgos laborales. Además, estos sesgos pueden ser imposibles de evaluar.

El sesgo que puede producirse a partir de este tipo de estudios que no se han llevado a cabo y que podrían haber revelado asociaciones de exposiciones específicas con resultados adversos no ha recibido un nombre concreto. En este contexto, se podría recordar un diálogo del relato «Estrella de plata» del personaje Sherlock Holmes de Sir Arthur Conan Doyle en el que Holmes investiga la desaparición de un caballo de carreras con ese nombre y el asesinato de su entrenador. Cuando Holmes está a punto de abandonar el pueblo durante la investigación, el inspector local se vuelve hacia él y le pregunta:

– ¿Existe algún detalle sobre el que desee llamar mi atención?

– Sobre el curioso incidente del perro aquella noche. [contesta Holmes]

– El perro no hizo nada por la noche.

– Ese es precisamente el curioso incidente –comentó Sherlock Holmes[18]*–.*

(Holmes describió más tarde cómo logró identificar al villano. Explicó que, cuando el intruso entró en el establo, «el perro no hizo nada durante la noche» y ni siquiera ladró mucho, lo que indicaba que «obviamente, el visitante nocturno era alguien a quien el perro conocía bien».)

Teniendo presente la conversación anterior, el posible sesgo introducido por los estudios que no se hacen podría denominarse *sesgo de Estrella de plata.* Holmes entendió por qué el perro no actuó y fue capaz de aplicar este conocimiento para resolver el problema que tenía entre manos. Del mismo modo, puede haber mucho que deducir cuando un fabricante no realiza lo que parece ser un estudio claramente necesario sobre los posibles efectos adversos de un producto. Pero, cuando tal asociación se ha sugerido, a menudo es difícil determinar si ciertos estudios epidemiológicos no se iniciaron debido a los intereses creados y las preocupaciones sobre sus posibles resultados. A falta de evidencia que documente una decisión explícita de no realizar un determinado estudio, este tipo de sesgo suele ser difícil o imposible de cuantificar e incluso de detectar.

Aunque el ámbito académico no es inmune a sus propios problemas y presiones, los problemas relacionados con la investigación epidemiológica que se plantean en este contexto tienen menos probabilidades de estar relacionados con el impacto potencial de los hallazgos específicos del estudio. Sin embargo, se debe considerar la posibilidad de un conflicto de intereses en relación con cualquier estudio epidemiológico, con independencia del contexto específico en el que se realizó la investigación. De hecho, tal conflicto puede estar más relacionado con las fuentes de financiación que con el propio contexto de la investigación. No obstante, se debe reconocer la posibilidad de que, en contadas ocasiones, ciertos conflictos de intereses tanto individuales como institucionales pueden influir en la publicación y difusión de los resultados. Para evitar sesgos de publicación, hay que intentar asegurar que los resultados del estudio (sean cuales sean) se publiquen en una revista revisada por pares en el momento oportuno. Los requisitos para el registro de los ensayos clínicos son un paso importante en esa dirección (v. descripción en cap. 11) El patrocinio del estudio debe quedar claramente indicado en el artículo que presenta sus resultados, al igual que cualquier interés económico o de otra índole de los investigadores o sus familias que pueda verse afectado por los resultados del estudio.

Interpretación de los resultados

Muchos de los temas cruciales con respecto al modo en el que se llevan a cabo los estudios epidemiológicos se relacionan con la idoneidad del diseño del estudio y con la interpretación y la comunicación de los resultados. A menudo se ha acusado a los epidemiólogos de comunicar nuevos riesgos sin cesar, muchos de los cuales no son importantes y no se han confirmado en estudios posteriores. El resultado es que los medios de comunicación presentan a la opinión pública muchos riesgos, a menudo no confirmados, lo que genera un escepticismo sobre los riesgos descritos recientemente, porque se es incapaz de distinguir los riesgos verdaderos e importantes de los no confirmados o triviales (fig. 20.4); por tanto, si la información es dudosa, a menudo, las personas no están dispuestas a asumir la responsabilidad de su propio cuidado de la salud[19]. Aquí vuelve a surgir la pregunta: ¿cómo puede valorarse la importancia de un único estudio que muestra un aumento del riesgo? ¿Cuántos estudios confirmatorios son necesarios?

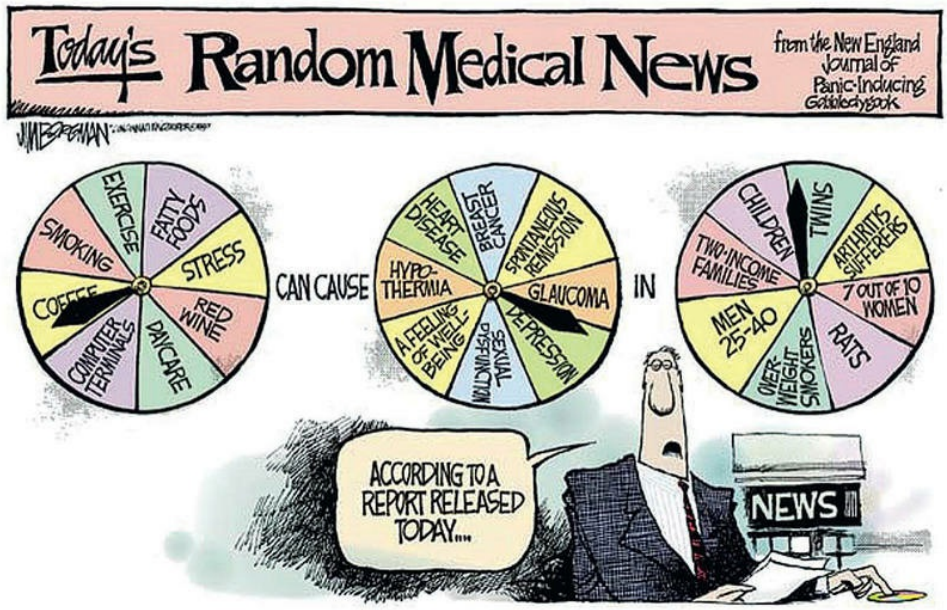


FIG. 20.4 Aproximación al flujo aparentemente interminable de los riesgos descritos que se presentan a la opinión pública. (JIM BORGMAN © Cincinnati Enquirer. Reproducida con autorización de ANDREWS MCMEEL SYNDICATION. Todos los derechos reservados.)

Un problema adicional es que, en años precedentes, los hallazgos epidemiológicos aislados o las controversias científicas solían abordarse y resolverse en el seno de la comunidad científica antes de que los hallazgos se difundiesen a la opinión pública. En la actualidad, tanto los hallazgos iniciales sin confirmar como las controversias científicas a menudo se publican en la prensa o en los medios populares (y cada vez más en las redes

sociales), incluso antes de que hayan aparecido los estudios en las revistas revisadas por pares (fig. 20.5). El dilema es que, aunque la mayor educación de la opinión pública y una mayor conciencia pública de los problemas científicos son loables, a menudo se incrementan de forma injustificada los niveles de ansiedad por los estudios aislados que se difunden ampliamente y que con frecuencia son refutados más tarde. El problema se ve agravado porque se ha descrito un sesgo en los periódicos consistente en que no se publican los resultados de estudios que no muestran ningún efecto[20].



FIG. 20.5 «La comunidad científica está dividida. Algunos dicen que esta mercancía es peligrosa; otros dicen que no.» Convivir con la incertidumbre científica. (Mischa Richte/The New Yorker Collection/The Cartoon Bank.)

Existe una incertidumbre significativa que se asocia con las conclusiones relativas a ciertas preguntas, tales como si la mamografía es beneficiosa para las mujeres de 40-50 años y si la prueba del antígeno prostático específico es beneficiosa para los varones con cáncer de próstata localizado. Afrontar la incertidumbre es difícil, y a menudo doloroso, para las personas que se están esforzando para tomar una decisión personal acerca de si seguir las intervenciones propuestas. Los epidemiólogos deben ayudar a la opinión pública a comprender la incertidumbre y a afrontar el reto de tomar decisiones frente a una información equívoca e incompleta.

Una pregunta adicional es: ¿En qué punto un aumento trivial notificado del riesgo relativo, aunque sea estadísticamente significativo, se convierte en un riesgo significativo desde el punto de vista biológico que merece la preocupación de la opinión pública? Esta pregunta se relaciona con la cuestión general de la percepción que tiene la opinión pública del riesgo. Estas percepciones se reflejan en las tablas 20.1 y 20.2. Para muchos de los riesgos enumerados, el grado de preocupación de la opinión pública y el cambio de las conductas no parecen acordes con la magnitud del riesgo.

Tabla 20.1

Riesgos involuntarios

Riesgo involuntario	Riesgo de mortalidad por persona por año
Atropello (Estados Unidos)	1/20.000
Atropello (Reino Unido)	1/16.600
Riada (Estados Unidos)	1/455.000
Terremoto (California)	1/588.000
Tornados (Medio Oeste)	1/455.000
Caída de un rayo (Reino Unido)	1/10 millones
Accidente de aviación (Estados Unidos)	1/10 millones
Accidente de aviación (Reino Unido)	1/50 millones
Fuga de una central nuclear	
En los límites del lugar (Estados Unidos)	1/10 millones
A 1 km (Reino Unido)	1/10 millones
Rotura de un dique (Países Bajos)	1/10 millones
Mordedura de animal venenoso (Reino Unido)	1/5 millones
Leucemia	1/12.500
Gripe	1/5.000
Meteorito	1/100.000 millones

De Dinman BD. The reality and acceptance of risk. *JAMA.* 1980;244:1226. Copyright 1980, American Medical Association.

Tabla 20.2

Riesgos voluntarios

Riesgo voluntario	Riesgo de mortalidad por persona por año
Tabaquismo: 20 cigarrillos/día	1/200
Alcoholismo: 1 botella de vino/día	1/13.300
Fútbol/fútbol americano	1/25.500
Carreras de automovilismo	1/1.000
Conducción de vehículos (Reino Unido)	1/5.900
Motociclismo	1/50
Escalada	1/7.150
Toma de anticonceptivos orales	1/5.000
Motonáutica	1/5.900
Piragüismo	1/100.000
Carreras de caballos	1/740
Boxeo amateur	1/2 millones
Boxeo profesional	1/14.300
Esquí	1/430.000
Embarazo (Reino Unido)	1/4.350
Aborto: legal <12 semanas	1/50.000
Aborto: legal >14 semanas	1/5.900

De Dinman BD. The reality and acceptance of risk. *JAMA.* 1980;244:1226. Copyright 1980, American Medical Association.

Si el riesgo absoluto es bajo, aunque el riesgo relativo en los individuos expuestos esté aumentado de manera significativa, el riesgo real para los

individuos expuestos seguirá siendo muy bajo. Es interesante que la opinión pública a menudo prefiera abordar los temas «calientes» (como el riesgo descrito de la daminozida presente en las manzanas), para los que la evidencia puede ser escasa, mientras se ignoran los factores de riesgo bien establecidos, como el tabaquismo, el consumo de alcohol y la exposición al sol, para los que la evidencia disponible justifica con claridad unos cambios del estilo de vida que dependen de la iniciativa individual.

Los epidemiólogos tienen un papel destacado a la hora de comunicar los riesgos para la salud e interpretar los datos epidemiológicos para las personas sin conocimientos de epidemiología; si los epidemiólogos no participan en esta actividad, la asumirán otras personas con mucha menos formación y experiencia. Esta es una parte esencial del proceso de elaboración de políticas. Los estudios de las poblaciones humanas a menudo proporcionan resultados divergentes, y los epidemiólogos con frecuencia dudan a la hora de sacar conclusiones basadas en los datos existentes. En el ámbito académico, los epidemiólogos pueden criticar el diseño de los estudios y sus hallazgos, y la conclusión típica es recomendar investigaciones adicionales para resolver un problema. Sin embargo, los responsables políticos que trabajan en primera línea no disfrutan del lujo de poder retrasar sus acciones, sino que deben tomar decisiones inmediatas (regular o no regular). Incluso la decisión de no regular en este momento representa una decisión política. Lo ideal es que estas decisiones se basen en los hallazgos epidemiológicos existentes. Sin embargo, los responsables políticos no pueden actuar de manera racional esperando a que los resultados de estudios futuros dirijan sus acciones con respecto a problemas urgentes de salud actuales. Por tanto, los epidemiólogos deben sacar las mejores conclusiones posibles a partir de los datos disponibles en la actualidad, siendo plenamente conscientes de que mañana puede aparecer un estudio mejor, o incluso un estudio perfecto, que tal vez contradiga las conclusiones actuales.

Los epidemiólogos tienen varios papeles en el proceso de elaboración de políticas, incluida la generación e interpretación de los datos, la presentación de opciones políticas específicas que sean consistentes con los datos, la proyección del impacto de cada opción, la elaboración de propuestas políticas concretas y la evaluación de los efectos de las políticas después de que se hayan aplicado. ¿Debería un epidemiólogo ser a la vez investigador y defensor de una política específica? ¿La defensa de una posición implica una pérdida de objetividad y de credibilidad científica? Estas son preguntas difíciles, pero muchas cuestiones claras, como los peligros para la salud debidos al consumo de cigarrillos, necesitan urgentemente la participación de los epidemiólogos en el esfuerzo para eliminar esta fuente del peligro para la salud pública. Así pues, la pregunta no es solo si es ético que un epidemiólogo sea defensor de una política, sino si es ético que un epidemiólogo *no* sea defensor cuando la evidencia del riesgo es tan convincente. Por tanto, el epidemiólogo debe actuar como educador y como investigador. Los esfuerzos educativos del epidemiólogo se dirigen a muchas

poblaciones diana, como otros científicos, otros profesionales de la salud, legisladores, responsables políticos, abogados, jueces y la opinión pública. Cada grupo debe ser tratado de forma diferente, en función de sus necesidades específicas y de los objetivos hacia los que se dirige el esfuerzo educativo. Los epidemiólogos deben aprender a trabajar con los medios de comunicación, como la radio, televisión, revistas, periódicos y, cada vez más, con las redes sociales, con el fin de avanzar en sus esfuerzos educativos. Los epidemiólogos también deben familiarizarse con lo que se sabe sobre cómo se perciben los riesgos por parte de los pacientes, los profesionales sanitarios, y la opinión pública general, para que puedan ayudar a estos grupos a lidiar con los resultados de los estudios epidemiológicos y con sus implicaciones para las medidas preventivas, incluidos los cambios del estilo de vida[21].

Conclusión

Los aspectos éticos y profesionales que se plantean a la epidemiología reflejan principalmente las obligaciones de los epidemiólogos con los participantes en los estudios epidemiológicos y clínicos, así como los problemas globales de una disciplina que se sitúa en la interfase entre la ciencia y la política de salud pública. Los problemas son complejos, a menudo sutiles, y no tienen respuestas simples. Dada la posición central de la epidemiología en el desarrollo de políticas tanto clínicas como de salud pública y sus implicaciones para las normativas ambientales, los cambios del estilo de vida individuales y los cambios rápidos en la práctica clínica, los resultados de los estudios epidemiológicos atraen la atención general y están muy expuestos a la opinión pública. A medida que la epidemiología aborde nuevas preguntas en el futuro, los dilemas éticos y profesionales a los que se enfrenta la disciplina también seguirán evolucionando. Existe una necesidad crucial de un diálogo continuo entre los epidemiólogos y quienes utilizan los resultados de los estudios epidemiológicos, incluidos los médicos y los responsables políticos, así como la población que se verá afectada por las nuevas políticas de salud y de prevención.

Bibliografía

1. Cited in Strong WS. Copyright in a time of change. J Electronic Pub. 1999; 4(3) [fecha de última consulta: 20 de agosto de 2013]: http://quod.lib.umich.edu/j/jep/3336451.0004.302/--copyright-in-a-time-of-change?rgn=main;view=fulltext.
2. Brandeis L. Dissenting opinion in *Olmstead v. United States*, 277 U.S. 438 (1928).
3. Gordis L, Gold E. Privacy, confidentiality, and the use of medical records in research. *Science*. 1980;207:153–156.
4. *The Report of the Privacy Protection Study Commission:*Personal Privacy in an Information Society. Washington, DC: US Government Printing Office; 1977.
5. *Health Insurance Portability and Accountability Act of 1996*. Pub. L. No. 104-191, 110 Stat. 1936 (1996).
6. Centers for Disease Control and Prevention.HIPAA Privacy Rule and public health: guidance from CDC and the U. S. Department of Health and Human Services. *MMWR*. 2003;52(suppl):1–20.
7. Gostin LO. National health information privacy: regulations under the Health Insurance Portability and Accountability Act. *JAMA*. 2001;285:3015–3021.
8. Gostin LO, Hodge Jr JG. Personal privacy and common goods: a framework. *Minn Law Rev*. 2002;86:1439–1480.
9. Kulynych J, Korn D. The new federal medical-privacy rule. *N Engl J Med*. 2002;347:1133–1134.
10. Kaufman JS, Cooper RS. Race in epidemiology: new tools, old problems. *Ann Epidemiol*. 2008;18(2):119–123.
11. Bhopal R, Donaldson L. White, European, Western, Caucasian, or what? Inappropriate labeling in research on race, ethnicity and health. *Am J Public Health*. 1998;88:1303–1307.
12. Marshall E. DNA studies challenge the meaning of race. *Science*. 1998;282:654–655.
13. Fullilove MT. Abandoning "race" as a variable in public health research—an idea whose time has come. *Am J Public Health*. 1998;88:1297–1298.
14. Senior PA, Bhopal R. Ethnicity as a variable in epidemiological research. *BMJ*. 1994;309:327–330.

15. Bhopal R. Is research into ethnicity and health, racist, unsound or important science? *BMJ*. 1997;314:1751–1756.
16. Rosenberg HM, Maurer KD, Sorlie PD, et al. Quality of death rates by race and Hispanic origin: a summary of current research, 1999. National Center for Health Statistics. *Vital Health Stat*. 1999;2(128):1–13.
17. Moubarac JC. Persisting problems related to race and ethnicity in public health and epidemiology research. *Rev Saude Publica*. 2013;47(1):104–115.
18. Doyle AC. Silver Blaze. The Complete Sherlock Holmes. New York: Doubleday; 1930.
19. Taubes G. Epidemiology faces its limits. *Science*. 1995;269:164–169.
20. Koren G, Klein N. Bias against negative studies in newspaper reports of medical research. *JAMA*. 1991;13:1824–1826.
21. Klein MP, Stefanek ME. Cancer risk elicitation and communication: lessons from the psychology of risk perception. *CA Cancer J Clin*. 2007;57:147–167.

Respuestas a las preguntas de repaso

Nota para el lector: para consultar la respuesta razonada de todas las opciones de respuesta, acceda a http://www.studentconsult.com y active/acceda a la versión íntegra del libro y el contenido extra (en inglés).

Capítulo 1

No contiene preguntas de repaso.

Capítulo 2

1. b
2. a
3. b
4. d
5. c
6. c

Capítulo 3

1. e
2. 10%
3. c
4. c
5. d
6. b
7. c
8. c
9. d
10. d

Capítulo 4

1. 5/1.000
2. 30%
3. e
4. b
5. b
6. a
7. 2,5 o 250
8. d
9. c
10. d
11. 9,6/1.000
12. e
13. d
14. a
15. b

Capítulo 5

1. 72,0%
2. 84,0%
3. 69,2%
4. d
5. d
6. b
7. 3,3%
8. b
9. 70,0%
10. 57,1%
11. 0,4
12. b

Capítulo 6

Las respuestas a las preguntas 6 a 8 se basan en el cálculo y la cumplimentación de la tabla que se proporciona (como se muestra más adelante en esta sección).

1. 54,8%
2. c
3. c
4. b
5. c
6. 0,982 o 98,2%
7. 0,006 o 0,6%
8. c
9. a
10. b

Capítulo 7

1. c
2. a
3. c
4. b
5. c
6. d
7. e
8. d
9. c
10. c
11. c

Capítulo 8

1. d
2. a
3. c
4. a
5. c
6. d
7. b

Capítulo 9

No contiene preguntas de repaso.

Capítulos 10 y 11

1. e
2. e
3. c
4. b
5. b
6. a
7. c
8. 57
9.
    9a. b
    9b. c
    9c. e
    9d. d
    9e. a
10. 0,67
11. 43

Capítulo 12

1. 15,3
2. d
3. e
4. e
5. 4,5
6. 6,3
7. 1:7 (0,143)
8. e
9. e
10. 1,94
11. 1,50
12. Las probabilidades de cáncer de próstata son un 50% más altas entre los pacientes que nunca recibieron aspirina en comparación con los que siempre la toman.
13. b

Capítulo 13

1. b
2. 27,5/1.000
3. 84,6%
4. 3,6/1.000
5. 78,3%

Capítulo 14

1. c
2. a
3. e
4. b
5. d

Capítulo 15

1. e
2. c
3. c
4. 12
5. 18,7
6. 9
7. 6,2
8. d
9. b

Capítulo 16

1. c
2. c
3. b
4. b
5. c

Capítulo 17

1. b
2. b
3. a
4. d
5. d

Capítulo 18

1. a
2. a
3. b
4. c
5. b
6. c
7. b
8. c

Capítulos 19 y 20

No contienen preguntas de repaso.

Para las preguntas 6 a 8 del capítulo 6:

Supervivencia de pacientes con SIDA tras el diagnóstico

(1) Intervalo desde el comienzo del tratamiento (meses)	(2) Vivos al comienzo del intervalo	(3) Muertos durante el intervalo	(4) Perdidos durante el intervalo	(5) Número efectivo expuesto al riesgo de morir durante el intervalo: Col (2) – ½ [Col (4)]	(6) Proporción que murió durante el intervalo: Col (3)/Col (5)	(7) Proporción que no murió durante el intervalo: 1 – Col (6)	(8) Proporción acumulada que sobrevivió desde la incorporación al final del intervalo: supervivencia acumulada
x	l_x	d_x	w_x	l'_x	q_x	p_x	P_x
1-12	248	96	27	234,5	0,4094	0,5906	0,5906
13-24	125	55	13	118,5	0,4641	0,5359	0,3165
25-36	57	55	2	56,0	0,9821	0,0179	0,0057

Índice alfabético

Los números de página seguidos de la letra *f* indican figuras, los seguidos de la letra *t* indican tablas y los seguidos de la letra *c* indican cuadros.

A

Abandonos, 213, 226

Aborto, cáncer de mama y, 294, 294*t*

Accidente cerebrovascular, 342-343, 343*f*

Acciones preventivas, 9-16

ADN (ácido desoxirribonucleico), 328*t*-331*t*

  secuenciación, 328*t*-331*t*

Adopción, estudios, 312-313, 314*t*

Aflatoxina, 304, 304*t*

Agentes estimulantes de la eritropoyesis (AEE), 24, 24*f*-25*f*

Agua, fluoración, 14-15, 15*f*

Aire, contaminación, infarto de miocardio y, 172, 173*f*

Ajuste, 296-297

  directo, por edad, 82-84, 82*t*

  indirecto por edad, 84-85, 85*t*

Alcohol, abuso, 281-282, 282*f*

Alelo, 328*t*-331*t*

Análisis

  de intención de tratar, 211, 365-366

  de la relación coste-beneficio, pruebas de cribado, 373-374

Anencefalia, 109, 109*f*

Años

  de vida ajustados por discapacidad (AVAD), 87, 88*t*

  potenciales de vida perdidos (AVPP), 71-73, 72*f*, 73*t*

APVP, *Véase* Años, potenciales de vida perdidos (APVP)
Árbol genealógico, 328*t*-331*t*
Artritis, 55
Artrosis, cirugía artroscópica sobre la rodilla, 231-232, 232*f*
Asignación aleatoria
  beneficios, 205
  consideraciones éticas, 235
  definición, 201
  estratificada, 205, 206*f*
  métodos, 201-202
  objetivo, 205
Asistencia
  escalonada, 227
  prenatal, 339
Asma, 53-54
Asociaciones
  directrices, 276-280, 276*c*
  especificidad, 279
  falsas, 271
  reales, 271, 271*f*-272*f*
  tipos, 271-274
Asociada a la asistencia sanitaria (AAS), 55, 56*f*, 56*t*
Aspectos
  de información geográfica, 61-62, 61*f*
  éticos
    descripción, 396-397
    en la epidemiología, 395-408
    ensayos clínicos con asignación aleatoria, 235
  profesionales, en la epidemiología, 395-408
Autismo, 328*t*-331*t*
Autonomía, 399-400
Autosómico dominante, 328*t*-331*t*

AVAD, *Véase* Años, de vida ajustados por discapacidad (AVAD)

B

Beneficencia, 397-399

Borrelia burgdorferi, 35

BRCA1, 313-314

BRCA2, 313-314

Brote, de la enfermedad, 26, 27*f*

  investigación, 36-38

  periodicidad, 33-34, 34*f*

C

Café, consumo, cáncer pancreático y, 163, 163*t*-164*t*, 164*f*-165*f*, 271, 272*f*, 295*f*

Calidad de vida, 86-87

Calmette-Guérin, bacilo, vacunación, 201, 201*t*

Cambio

  porcentual anual (APC), 58*f*

  social, 396

Campos electromagnéticos, 386

Cáncer

  cervical uterino, 81*f*, 165-166, 357*f*, 370-371

  de colon, 50*f*, 361, 361*f*

  de estómago, 280-281, 313, 314*t*, 315*f*

  de la cavidad bucal, 302*t*-303*t*

  de mama

    abortos provocados y, 294, 294*t*

    cribado, 367-370, 368*f*

    deficiencia de progesterona y, 182-183, 183*f*

    letalidad los 5 años, 366, 366*f*

    metastásico, efectos del apoyo psicosocial sobre la supervivencia de las pacientes con, 233-234, 233*f*-234*f*

    prevención, tamoxifeno, 230-231, 230*f*-231*f*

tasas
de incidencia, 50*f*, 75*f*-76*f*
de mortalidad, 16, 16*f*, 66, 75*f*-76*f*
de supervivencia, 367*f*
de próstata, 50*f*
de pulmón, 50*f*
exposición a radiación y, 303*t*
incidencia, 49-50
tabaquismo y, 16, 158-159, 159*t*, 162-163, 265*t*, 277, 279*f*, 297*f*, 297*t*, 302*t*-303*t*
tasas de mortalidad, 16, 16*f*, 66, 68, 278*f*, 297-298, 298*f*
esofágico, tabaquismo y, 298*f*, 299
hepático, 304*t*
pancreático, consumo de café y, 163, 163*t*-164*t*, 164*f*-165*f*, 271, 272*f*, 295*f*
papilar, 75-76, 78*f*
tiroideo, 50*f*, 75-76, 77*f*-78*f*
uterino, tasas
de incidencia, 50*f*
de mortalidad, 66
Caracterización del riesgo, 385
Carcinoma
epidermoide, tabaquismo y, 321, 322*f*
hepatocelular, 322
Caso(s)
de intervalo, 360
incidentes, estudios con diseño de casos y controles, 159-160
prevalentes, estudios con diseño de casos y controles, 160
primario, 32
Causas, *Véase también* Etiología, de la enfermedad
asociaciones observadas como causales, 276-280, 276*c*
congruencia con otros datos, 278
consideración de las explicaciones alternativas, 277-278

directas, 274, 274*f*
directrices para la evaluación, 276-280, 276*c*
especificidad de la asociación, 279
fuerza de la asociación, 277
indirectas, 274, 274*f*
inferencias, 269-288
interrupción de la exposición, 278
postulados de Koch, 275-276
principales de muerte, 3, 4*f*, 4*t*, 265*f*
en los niños, 68*f*
tendencias de mortalidad en Estados Unidos, 73-74, 73*f*-74*f*
relación
dosis-respuesta, 277
temporal, 276, 276*f*
replicación de los hallazgos, 277
verosimilitud biológica, 277
Centers for Disease Control and Prevention (CDC), 403
Certificado de defunción, 76, 78*c*, 79*f*-80*f*
Chapin, Charles V, 334
Chernóbil, accidente, 45, 45*f*
Childs, B., 323-324
Circuncisión, 166, 166*t*
Cirugía artroscópica sobre la rodilla, tratamiento de la artrosis, 231-232, 232*f*
Clasificación errónea
diferencial, 292
no diferencial, 292
Cociente de posibilidades
de mortalidad estandarizado, 84, 85*t*
de productos cruzados, 248
definición, 245-248
diseño de estudios
de casos y controles, 243*t*, 246-248, 247*f*

de cohortes, 243*t*, 248
estimaciones del riesgo relativo, 249-250, 249*f*-250*f*
interpretación, 248
sin emparejamiento, 295
Cólera, 3, 3*f*, 14-16, 14*t*-15*t*, 397
Collaborative Perinatal Study, 184
Comparaciones de casos y controles, 251
Confidencialidad, 399-401, 400*f*
Conflicto de intereses, 404-405
Confusión, sesgo, 271-272, 294-299, 295*f*-296*f*, 295*t*-296*t*, 296*c*
Consejo genético, 328*t*-331*t*
Consentimiento informado, 235, 397, 398*f*, 400
Consumo de cigarrillos, 16
cáncer
de pulmón y, 158-159, 159*t*, 162-163, 265*t*, 277, 279*f*, 297*f*, 297*t*, 302*t*-303*t*
esofágico y, 298*f*, 299
carcinoma epidermoide y, 321, 322*f*
coronariopatía y, 158, 158*t*, 243, 243*t*, 262, 263*t*, 264, 265*t*
diseño de un estudio de cohortes, 178, 179*t*
materno, 272*f*-273*f*
Contaminación
continuada, 26
periódica, 26
Control(es)
correspondiente
de mejor amigo, 162
de vecindario, 161
definición, 157-174
del riesgo, 384, 385*f*
históricos, 200
hospitalizados, estudios con diseño de casos y controles, 162-163, 162*f*, 169
mediante enmascaramiento doble, 207-208

múltiples, 169-171
no hospitalizados, estudios con diseño de casos y controles, 161-162
orígenes, 161-165
problemas, 163-165
selección, 160-165
simultáneos no aleatorizados, 200-201
Convulsiones
febriles, 142
no febriles, 143, 143*f*
Coronariopatía, 229
estudio con diseño transversal, 154
hipertensión y, 380, 380*f*-381*f*
incidencia acumulada, 383*f*
infarto de miocardio, 245*f*
niveles del colesterol y, 244*t*, 272, 272*f*
tabaquismo y
diseño de un estudio
de casos y controles, 158, 158*t*
de cohortes, 178, 179*t*, 243, 243*t*, 263*t*, 264, 265*t*
Coronary Drug Project, 213, 213*t*-214*t*
Crenezumab, 18-19
Cribado
genético, 328*t*-331*t*
neonatal, 328*t*-331*t*
Criterio(s), 286
de valoración de los resultados, 336, 336*c*
transversal, 85-86
Cromosoma, 328*t*-331*t*
Cuantificación, pronóstico de la enfermedad, 123, 124*f*
Cuarentena, 29
Curva
bimodal, 94

de supervivencia, 131, 132*f*
unimodal, 94-95

D

Datos
acceso a, 401-402
de grupo, 336-341, 337*f*
asistencia prenatal, 339
desventajas, 338
resultados de la investigación, 337-339
sesgos, 339-340
hospitalarios, limitaciones, 58, 58*c*
individuales, 341-350
Daubert contraMerrell Dow Pharmaceuticals, 391
Defecto del tubo neural, 109
Deficiencia de progesterona, cáncer de mama y, 182-183, 183*f*
Deleción, 328*t*-331*t*
Demencia, proporción de la prevalencia, 55-56, 57*f*
Denominadores, problemas con, 56-58
Densidad de incidencia, 46
Determinantes de la susceptibilidad, 378
Diabetes mellitus
pruebas de cribado, 99, 100*f*
tasas de mortalidad, 80*f*
Diagnóstico
confirmación tisular, 124
enfoque basado en la población a, 6-7
métodos, 124
Diseño(s)
antes-después, evaluaciones de los servicios asistenciales, 344-345
combinados, evaluaciones de los servicios asistenciales, 348-350, 348*f*
con y sin programa, evaluación de los servicios asistenciales, 345-346, 345*t*,

346*f*
factorial, 211, 211*f*-212*f*
transversal
de casos, 172-174, 173*f*
no planificado, 208-211, 210*f*
Disfunción eréctil, prevalencia, 156, 156*f*
Distribución, bimodal, 94
Doll, Richard, 158-159
Dominante, definición, 328*t*-331*t*
Donne, John, 395
Down, síndrome, 109, 109*f*

E

Edad, mortalidad y, 80
Efectividad
de los programas de cribado, 354*c*
en las evaluaciones de los servicios asistenciales, 335
Efecto(s)
cohorte, 85-86, 85*t*-86*t*
de apoyo psicosocial sobre la supervivencia de las pacientes con cáncer de mama metastásico, 233-234, 233*f*-234*f*
Eficacia, 335
cálculo, 222
comparativa, investigación, 226
efectividad, comparación con, 222
Eficiencia, en las evaluaciones de los servicios asistenciales, 335
Embarazo, exposición a enfermedades durante el, 184, 184*f*
Emparejamiento, 167-169
definición, 167
excesivo, 168-169
grupal, 167-168
individual, 168-169

Encuestas mediante entrevista, error, 57*c*
Endémico, 23-26, 24*f*
Enfermedad(es)
aparición, 32-36, 65-93, 89*f*
brotes, 26
clínica, 21-23, 355-356
complejas, 308-313, 328*t*-331*t*
concepto del «iceberg», 21-22, 22*f*
detección temprana, 353-354, 374
diarreica, 33-34
distribución, 35
evolución natural, 123-146, 355-357, 355*f*
exposición y, asociación entre, 178, 179*f*, 241-242, 270*f*, 299
factores causales, 20, 21*t*
fase(s)
en el individuo y en la población, 43-45, 43*f*
preclínica detectable, 356
gravedad, 22-23, 23*f*
hereditarias mendelianas, 307-308
impacto futuro, 87-89, 88*f*
incidencia, *Véase* Tasa(s), de incidencia
inicio biológico, 123
latente, 23
no clínica, 23
parámetros para la evaluación del impacto, 86-89
periodicidad, 33-34
persistente (crónica), 23
preclínica, 23, 355-356, 356*f*, 359
prevalencia, 51-55, *Véase también* Prevalencia
prevención de, *Véase* Prevención
progresión, 43-45, 44*f*, 357-358
renal poliquística, 318*f*

subclínica, 21-23
teoría miasmática, 14
transmisión, 20-40
aspectos generales, 20, 41
directa, 20, 21*c*
gotículas, 21*f*
indirecta, 20, 21*c*
superficies corporales, 21, 22*f*
transmitida por los alimentos
causas principales, 26
evaluaciones de riesgos, 241, 241*t*
tasa de ataque, 50-51
Enfoque
basado en la población
a la prevención, 380-382, 382*f*
al diagnóstico, 6-7
descripción, 6-7
de alto riesgo, 6-7
epidemiológico, evaluación de cribado, 353-376
Enmascaramiento, control, 207-208, 207*t*
Ensayos clínicos
con asignación aleatoria, 197-239
aplicación de metaanálisis a, 388
cirugía artroscópica en el tratamiento de la artrosis de la rodilla, 231-232, 232*f*-233*f*
consentimiento informado, 235
consideraciones éticas, 235
control(es)
históricos, 200
mediante enmascaramiento, 206
simultáneos no aleatorizados, 200-201
definición, 174*t*, 197-198

diseño, 198, 198*f*, 241*f*, 343*f*
de estudios de cohortes, comparación con, 186
factorial, 211, 211*f*-212*f*
transversal, 208-211, 209*f*-210*f*
efectos del apoyo psicosocial sobre la supervivencia de las pacientes con cáncer de mama metastásico, 233-234, 233*f*-234*f*
evaluaciones de los servicios asistenciales, 342-344, 343*f*
factores limitantes, 226
generalización, 223-224, 224*f*
medicamentos nuevos, 226-227
objetivos, 198, 223, 231
participantes
abandono, 213, 226
asignación, 199-201
control mediante enmascaramiento, 207, 207*t*
incumplimiento del tratamiento, 213-214
reclutamiento, 213
y retención, 222
selección, 199
prevención del cáncer de mama con tamoxifeno, 230-231, 230*f*-231*f*
recopilación de datos, 205-208
resultado, 206
tratamiento, 206
registro, 234
resultados, 219*f*
eficacia, 222
generalización, 223-224, 224*f*
investigación de la eficacia comparativa, 226
médico responsable del paciente individual, 224-226, 225*f*
métodos de expresión, 222-223
número de pacientes que necesitarían ser tratados, 223
riesgo relativo, 222-223

utilidad del tratamiento, 222
tamaño de la muestra, 216-222, 217*c*, 217*f*-219*f*, 219*c*
tratamiento del accidente cerebrovascular, 343*f*
validez
externa, 290
interna, 224, 224*f*, 290
vitamina C, 207*t*
de fase I, 226
de fase II, 226
de fase III, 226
de fase IV, 227
Entrevistas sustitutas, 293
Epidemia, 23-26, 24*f*
curva, 31
Epidemiología, 397
aspectos éticos, 396-397
datos descriptivos, 8, 8*f*
de campo, 14-15
definición, 2
en los tribunales, 390-391
enfoque, 7-9, 9*f*
función, 377
genética, 328*t*-331*t*
modelo tradicional de factores de riesgo, 378, 378*f*
molecular, 386-387
objetivos, 2-5, 397
práctica clínica y, 7, 7*f*-8*f*
prevención y, 377-379, 378*f*-379*f*, *Véase también* Prevención
y salud pública, 377-394
Error
de tipo I, α, 218, 219*t*
de tipo II, β, 218, 219*t*

Esclerosis
lateral amiotrófica, 170-171
múltiple, 315-316
Escorbuto, 198
Especificidad, 95, 96*t*
de la asociación, 279
evaluación, 372, 373*f*
neta, 101*t*-103*t*, 103-104
cálculo, 101
con dos pruebas simultáneas, 103-104, 105*t*, 106*f*, 107*t*
pruebas secuenciales, 101, 102*f*
valor predictivo positivo y, 110-111, 110*f*, 111*t*
Esperanza de vida, 5, 5*f*
Espina bífida, niveles de α-fetoproteína en el líquido amniótico como prueba de cribado, 109, 109*f*
Esquizofrenia, 311, 314*t*
Estratificación, 296-297
Estudio(s)
con asignación aleatoria, 202*f*, 204, 204*f*
evaluaciones mediante cribado, 364-367, 365*f*-366*f*
con diseño
de casos y cohortes, 189, 190*f*
de casos y controles
anidados, 187-189, 188*f*-189*f*, 194
características, 195*t*
casos
definición, 157-174
incidentes, 159-160
pacientes hospitalizados, 161
prevalentes, 160
cociente de posibilidades, 243*t*, 245-248, 246*t*, 247*f*
comparaciones equiparadas, 251-253, 252*f*-253*f*, 291*t*

con datos emparejados, 251-253, 252*f*-253*f*, 291*t*
consumo de café y cáncer pancreático, 163, 163*t*-164*t*, 164*f*-165*f*
controles
  correspondientes
    de mejor amigo, 162
    de vecindario, 161
  definición, 157-174
  fuentes, 161-165
  múltiples, 169-171
  personas
    hospitalizadas como, 162-163, 162*f*, 169
    no hospitalizadas como, 161-162
  problemas, 163-165
  selección, 160-165
definición, 174*t*, 270
diseño, 157-159, 157*f*, 158*t*, 171*f*, 241*f*, 365*f*
  anidado, 194
  de estudios de cohortes, comparación con, 158-159, 193-194, 194*f*, 196*f*
  transversal de casos, 172-174, 173*f*
ejemplo, 158*t*
emparejamiento, 167-169
  definición, 167
  grupal, 167-168
  individual, 168-169
evaluaciones
  de los servicios asistenciales, 350
  mediante cribado, 363-364
indicaciones, 171*f*
recuerdo
  limitaciones, 165-166
  problemas, 165-167

sesgo asociado, 166-167
sesgos
de información, 165-167, 166*t*-167*t*
de selección, 159-167
de cohortes, 178-192, 179*f*
cálculos del riesgo relativo, 243-245, 243*t*
cáncer de mama y deficiencia de progesterona, 182-183, 183*f*
cociente de posibilidades, 243*t*, 246-248, 247*f*
Collaborative Perinatal Study, 184
concurrentes, 180
definición, 174*t*, 270-271
del tabaquismo y la coronariopatía, 178, 179*t*
diseño, 178-179, 179*f*, 179*t*, 186*f*, 241*f*
ejemplos, 181-183, 182*t*
ensayos clínicos con asignación aleatoria, comparación con, 186
estudio(s)
con diseño de casos y controles
anidados, 187-189, 188*f*-189*f*
comparación con, 158-159, 187-190, 187*f*, 193-194, 194*f*, 196*f*
de Framingham, 181-182, 244
histórico, 181
periodo de seguimiento, 181
poblaciones de estudio, 179-180, 179*f*-180*f*
problemas asociados, 181, 181*f*
prospectivo, 174*t*, 181, 181*f*, 195*t*
retrospectivo, 158-159, 174*t*, 181, 181*f*, 195*t*
salud y enfermedad infantiles, 183-185, 184*f*
sesgo, 185-186
de información, 186
de selección, 185
sin asignación aleatoria, 347*f*-348*f*, 365*f*
sin emparejamiento, 250-251, 250*f*, 251*t*

tipos, 180-181, 180*f*
de observación, 149-177
definición, 178
diseño de estudios
de casos y controles, *Véase* Estudio(s), con diseño, de casos y controles
de cohortes, *Véase* Estudio(s), con diseño, de cohortes
longitudinal, 180
transversal, 154-157, 155*f*, 174*t*, 208-211, 210*f*
no planificados, 208-211, 210*f*
planificados, 208, 209*f*
de asociación, 318-320
del genoma completo (GWAS), 328*t*-331*t*
de casos clínicos, 199
del proceso, 334-335
ecológicos
de los servicios asistenciales, 340-341, 342*c*
en migrantes, 314-316, 314*t*, 316*t*
epidemiológicos, diseños, 309
prospectivo no concurrente, 181
relativos a la actividad laboral, 400, 400*f*
retrospectivo con diseño de cohortes, 158-159, 174*t*, 181, 181*f*, 195*t*
simultáneos sin asignación aleatoria, 345-46
sin asignación aleatoria, 347*f*
descripción, 204*f*, 344-350
diseño
antes-después, 344-345
de cohortes, 347*f*-348*f*
evaluaciones mediante cribado, 363-364
simultáneos
con y sin programa, 345-346
no usuarios, 347, 347*t*

poblaciones
elegibles, 347-348
no elegibles, 347-348
usuarios, 347
transversal de casos, 172, 173*f*
descripción, 150-154, 152*f*
Etiología, de la enfermedad, *Véase también* Causas
descripción, 2
en los grupos de población, 270-271, 270*f*-271*f*
enfoques para el estudio, 269-271
estudios epidemiológicos, 336, 337*f*
Evaluación
del riesgo, 384-387
definición, 384
pasos implicados, 384-385
problemas metodológicos, 385
dosis-respuesta, 384-385
Evidencia anecdótica, 392-393
Evolución natural, *Véase también* Pronóstico
aspectos fundamentales, 356, 357*f*
de la enfermedad, 43*f*, 123-146, 355-357, 355*f*
del cáncer cervical uterino, 357*f*
descripción, 2, 123
representación esquemática, 123, 124*f*
Exceso de riesgo, 241*t*-242*t*
Experimentos
naturales, 270
no planificados, 270
Exposición(es)
a un vehículo común, 26
enfermedad y, asociación entre, 178, 179*f*, 241-242, 270*f*, 299
evaluación, 385-387

fuentes de datos, 385*c*
macroambientales, 379
microambientales, 379, 379*f*
única, epidemia con vehículo común, 31

F

Factor(es)
ambientales, factor genético y, 307-328, 320*f*, 321*t*, 322*f*
de riesgo, 2
genéticos
enfermedades complejas, 308-313
estudios
de adopción, 312-313, 314*t*
de asociación, 318-320
en gemelos, 309-312, 310*f*-312*f*, 310*t*-311*t*
en migrantes, 314-316, 314*t*, 316*t*
familiares
análisis de relación genética, 317-320, 317*f*-319*f*
estudios de adopción, 312-313, 314*t*
factores ambientales y, 307-328, 320*f*, 321*t*, 322*f*
genética tradicional, 307-308
medicina de precisión, 322-323
perspectivas, 323-324, 324*f*
V Leiden, mutación, 320-321
Falacia ecológica, 150-151
Familiares en primer grado, riesgo de enfermedad, 317
Faringitis por estreptococos β-hemolíticos del grupo A, 37
Fase preclínica, 6, 355-356, 356*f*, 359
Feminine Forever, 382
Fenilcetonuria, 308
Fenotipo, 328*t*-331*t*
α-Fetoproteína, 109, 109*f*

Fiabilidad, pruebas de cribado, 94-122
validez y, relación entre, 117, 117*f*
variación
interobservador, 113-117, 114*f*
intraindividual, 112, 112*f*
intraobservador, 113, 113*f*
Fibrosis quística, 319*f*
Fiebre
puerperal, 9-12, 11*f*, 12*t*
reumática, 349, 349*f*-350*f*
Fluoración, 9, 10*f*
Framingham, estudio, 181-182, 244
Frost, Wade Hampton, 333
Frye, prueba, 390-391

G

Galton, Francis, 197
Gastroenteritis aguda, 26, 27*f*
Gemelos
dicigóticos, 310, 312*t*
idénticos, 328*t*-331*t*
Gen, 328*t*-331*t*
candidato, 328*t*-331*t*
Generalización, 223-224, 224*f*, 290, 338
Genética tradicional, 307-308
Genoma, 328*t*-331*t*
Genómica, 328*t*-331*t*
Genotipo, 328*t*-331*t*
Gonorrea, 32, 33*f*
Gripe, leucemia linfocítica aguda y, 151-152, 151*t*
Grupo(s)
de pacientes remitidos, 227

equiparados, 167-168
étnico, 402-403
expuesto y grupo no expuesto, diseño de estudios de cohortes, 193, 194*f*, 196*f*
Guillain-Barré, síndrome, 171

H

H5N1, virus, 24-25, 25*f*
Health Insurance Plan, 364-365, 365*f*, 367*f*-368*f*
Health Insurance Portability and Accountability Act de 1996, 401
Healthy People 2020, indicadores de salud, 342*c*
Heart and Estrogen/Progestin Replacement Study (HERS), 383
Helicobacter pylori, 280-281, 280*f*, 281*c*, 395
Henderson, D. A, 13-14, 13*f*
Hepatitis C, prevalencia, 153, 154*f*
Herencia mendeliana, 328*t*-331*t*
HERS, *Véase* Heart and Estrogen/Progestin Replacement Study (HERS)
Heterocigoto, definición, 328*t*-331*t*
Hill, Austin Bradford, 279
Hipertensión, 380, 380*f*-381*f*
Hipócrates, 399
Histerectomía(s), 57-58, 58*f*
laparoscópicas, 345*t*
Historias clínicas de los pacientes, 58*c*
objetivos, 399
Homocigoto, definición, 328*t*-331*t*
Humo del tabaco en el ambiente, 16
Hypertension Detection and Follow-up Program, 227-229, 228*f*, 228*t*

I

Identificación de riesgos, 384
Incertidumbre, 391-393, 391*f*, 392*c*

significativa, 406
Incidencia
cálculo, objetivos, 49-50
cáncer
de mama, 50*f*, 60, 60*f*
tiroideo, 50*f*, 75-76, 77*f*-78*f*
datos de mortalidad utilizados como índice, 73
ejemplos, 52*f*
mortalidad, comparación con, 74, 75*f*
prevalencia y, 51, 51*t*, 53*f*, 59-62, 59*f*-60*f*, 59*t*
problemas, 55-58
tendencias cronológicas, 313-317, 315*f*
Inclusión, ventajas, estudio con diseño de casos y controles, 190
Incumplimiento del tratamiento, 213-214
Indicadores de salud, 341, 342*c*
Infarto de miocardio
contaminación del aire y, 172, 173*f*
coronariopatía y, 245*f*
Infección(es), *Véase también* las infecciones específicas
aspectos de confidencialidad, 397
estreptocócica, 16-17
materna, durante el embarazo, 167
periodos de incubación, 29, 31*f*
transmitidas por los alimentos, 241, 241*t*
Inferencias causales, 289-306
de confusión, 271-272, 294-299, 295*f*-296*f*, 295*t*-296*t*, 296*c*
edad de inicio del consumo de alcohol y consumo de alcohol a lo largo de la vida, 281-282, 282*f*
ejemplos, 280-282, 281*c*, 282*f*
interacción, 299-304, 299*t*-302*t*
modificaciones de las directrices, 282-286, 283*t*, 284*c*
sesgo, *Véase* Sesgo

úlceras pépticas y cáncer gástrico, relación con la infección por Helicobacter pylori, 280-281, 280*f*, 281*c*
Informes de casos clínicos, 149-150
Inmunidad, 26-27
de grupo, 27-28, 28*f*
Inmunoanálisis enzimático, 372
Interacción, 299-304, 299*f*, 299*t*-302*t*
genética, 328*t*-331*t*
Interdependencia, 395-396, 396*f*
Interpretación de los hallazgos, 405-407, 405*f*-406*f*, 406*t*-407*t*
Investigación(es)
de los resultados, 337-339, 344
en gemelos, 309-312, 310*f*-312*f*, 310*t*-312*t*, 390
Investigadores, obligaciones de estudiar a los participantes, 397-399, 398*f*

J

Jenner, Edward, 12-14, 12*f*-13*f*
Junta de revisión institucional, 401
Juramento Médico, hipocrático, 399

K

Kaplan-Meier, método, 134-136, 134*f*-136*f*, 134*t*, 282*f*, 343*f*
Kappa, estadístico, 114-117
Kass, Edward, 17

L

Leucemia
campos electromagnéticos y, 386
en los niños, 137, 137*f*-138*f*
linfocítica aguda, exposición prenatal a la gripe y, 151-152, 151*t*
Ligado
al cromosoma X, definición, 328*t*-331*t*

al sexo, definición, 328*t*-331*t*
Lind, James, 198
Línea de células germinales, 328*t*-331*t*
Locus cromosómico, 328*t*-331*t*
Logaritmo del tiempo, 31, 31*f*
Lyme, enfermedad, 35, 35*f*

M

Malaria, 43
Mamografía, 365, 367-370, 368*f*, 373
Mapa genético, 328*t*-331*t*
Marcadores
biológicos, 387, 387*f*
genéticos, 317, 328*t*-331*t*
Mediana de supervivencia, 141
Medicare, 339, 344-345, 346*f*
Medicina de precisión, 322-323
Mellizos, 328*t*-331*t*
Mendel, Gregor Johann, 328*t*-331*t*
Meningitis aséptica, 33-34, 34*f*
Metaanálisis, 387-388, 388*f*
Micrometástasis, 139
Migración en fases, 138-139, 139*f*-140*f*
Modelo
multiplicador, de la interacción, 300-301
tradicional de factores de riesgo, epidemiología, 378, 378*f*
Modo de transmisión de las enfermedades, 20-21, 21*c*, 21*f*
directo, 20, 21*c*
indirecto, 20, 21*c*
Morbilidad, parámetros, 45-62
fuentes, 55, 55*c*
limitaciones, 55

prevalencia, *Véase* Prevalencia
tasa
de ataque, 50-51
de incidencia, *Véase* Tasa(s), de incidencia
Mortalidad, 69
cálculo, 69
cáncer de mama, 66, 75*f*-76*f*
de pulmón, 66
tiroideo, 75-76, 77*f*
uterino, 66
comparaciones de poblaciones, 80-86
ajuste por edad
directo, 82-84, 82*t*
indirecto, 84-85, 85*t*
efecto cohorte, 85-86, 85*t*-86*t*
interpretación de los cambios observados, 86
definición, 125
en los niños, 68*f*
envejecimiento como factor predictivo, 80
evitable, 341
expresión del pronóstico, 125
incidencia, comparación con, 74, 75*f*
interpretación de los cambios observados, 86, 87*c*, 87*t*
no ajustada, 80-82
parámetros, 65-78
años de vida potencial perdidos, 71-73, 72*f*, 73*t*
cáncer, 65, 66*f*-67*f*
mortalidad proporcional, 69-71, 70*f*-71*f*, 70*t*-71*t*
ajuste directo respecto a la edad, 82-84
por cáncer, 65, 66*f*-67*f*
de mama, 366, 366*f*
proporcional, 69-71, 70*f*-71*f*, 70*t*

razones del estudio, 65, 73-76
tasa de mortalidad, comparación con, 69*c*
total, 80-82
Moubarac, Jean-Claude, 403
Muerte(s)
cambios en la definición, 77-78
causa(s)
principales, 67*f*, 265*f*
subyacente, 76
codificación, 76-77
Muestreo sesgado por la cantidad, 359-360, 359*f*-360*f*
Multiple Risk Factor Intervention Trial, 229-230, 229*f*-230*f*
Mutación, 328*t*-331*t*

N

Nefropatía crónica, prevalencia, 156, 157*f*
Neuroblastoma, 371-372, 371*f*, 372*t*
Niños
causas principales de muerte, 68*f*
diseño de estudios de cohortes, investigación de la salud y la enfermedad infantiles, 183-185, 184*f*
leucemia, 137, 137*f*-138*f*
No maleficencia, 400
Norovirus, 26
Notificación pasiva, 42
Numeradores
de la mortalidad, 69
problemas con, 55-56
Número(s)
aleatorios, 202, 202*c*, 202*t*
de pacientes que necesitarían ser tratados, 223
necesario para dañar (NND), 223

O

Obesidad, 54-55

Observaciones, 9-16

Ochsner, Alton, 270

Oración, 197

Organización Mundial de la Salud

iniciativas para la erradicación de la viruela por, 13-14

P

p53, mutaciones, 321

cervical, 81*f*, 165-166, 357*f*, 370-371

chequeos, recomendaciones de la American Cancer Society, 374*c*

colon, 50*f*, 361, 361*f*

epidermoide, tabaquismo y, 321, 322*f*

estómago, 280-281, 313, 314*t*, 315*f*

hígado, 304*t*

oral, 302*t*-303*t*

pacientes con, 233-234, 233*f*-234*f*

cribado, 367-370, 368*f*

deficiencia de progesterona y, 182-183, 183*f*

prevención, usando el tamoxifeno, 230-231, 230*f*-231*f*

tasas

de mortalidad, 16, 16*f*, 66, 75*f*-76*f*

de supervivencia, 367*f*

pancreático, consumo de café y, 163, 163*t*-164*t*, 164*f*-165*f*, 271, 272*f*, 295*f*

papilar, 75-76, 78*f*

piel, tasas de incidencia, 50*f*

próstata, 50*f*

pulmón, 50*f*

exposición a radiación y, 303*t*

incidencia, 49-50

tabaquismo y, 16, 158-159, 159*t*, 162-163, 265*t*, 277, 279*f*, 297*f*, 297*t*, 302*t*-

303*t*
tasas de mortalidad, 16, 16*f*, 66, 68, 278*f*, 297-298, 298*f*
tasas de mortalidad, 65, 66*f*-67*f*
tiroides, 50*f*, 75-76, 77*f*-78*f*
tuberculosis y, 160, 160*t*
uterino, tasas
de incidencia, 50*f*
de mortalidad, 66
Países en desarrollo
aparición de enfermedades, 3-5
vigilancia, 42-43
Pandemia, 23-26
Papanicolaou, prueba, 357-358, 370
Parámetros de resultado, 206
Paré, Ambroise, 198
Pares
concordantes, 251
discordantes, 251
Participaciones, 213
Participantes
con enfermedad (casos) y sin enfermedad (controles)
estudio
con diseño de casos y controles, 196*f*
de cohortes, 194*f*
ensayos clínicos con asignación aleatoria
abandono, 213
asignación, 201-205
incumplimiento del tratamiento, 213-214
reclutamiento, 213
y retención, 222
selección, 199
obligaciones de los investigadores del estudio, 397-399, 398*f*

Patrón de referencia, 95-96
Pearl, Raymond, 160
Periodo
  de espera
    evaluación de cribado, 356, 356*f*
    y supervivencia a los 5 años, 127, 127*f*-128*f*, 361-363, 361*f*-362*f*
  de incubación
    definición, 28
    enfermedad y, 28-32, 30*f*
Personas-años, 125-126, 126*f*
Personas-tiempo, 45, 46*f*-48*f*
Peste negra, 29
Piel
  agentes infecciosos que entran a través de la, 21
  cáncer, tasas de incidencia, 50*f*
Placebo, 207
Población
  comparaciones de la mortalidad, 80-86
    ajuste por edad
      directo, 82-84, 82*t*
      indirecto, 84-85, 85*t*
    efecto cohorte, 85-86, 85*t*-86*t*
    interpretación de los cambios observados, 86
  de estudio, 223-224
  definida, 223
  enfoques etiológicos, 270-271, 270*f*-271*f*
  total, 223
    cálculo del riesgo atribuible, 261-262, 266*t*
    descripción, 223
  variación biológica, 94-95, 95*f*
Polimorfismo(s), 328*t*-331*t*
  de nucleótido único (SNP), 328*t*-331*t*

Poliomielitis, 28
Poliovirus, inactivado, vacuna, 28
Política sanitaria pública
  efectos de la incertidumbre, 393, 393*f*
  epidemiología, 3, 377-394
  riesgo y, 393-394
Porcentaje de concordancia, 113-114, 113*t*, 114*f*
Portador(es), 23, 328*t*-331*t*
  cribado, 328*t*-331*t*
  sano de los patógenos asociados con la fiebre tifoidea, 23
Potencia estadística, 219, 219*t*
Práctica clínica, 7, 7*f*-8*f*
Prehipertensión, definición, 380
Prevalencia, 51-55
  de la nefropatía crónica, 54*f*
  de la obesidad, 54-55
  de periodo, 51
  del asma, 53-54
  ecuación para el cálculo, 51
  ejemplos, 52*f*, 53-55, 54*f*
  incidencia y, 51, 51*t*, 53*f*, 59-62, 59*f*-60*f*, 59*t*
  problemas con, 55-58
  proporción, 56*f*
  puntual, 51, 59*t*
Prevención
  enfoque
    basado en la población, 6-7, 380-382, 382*f*
    de alto riesgo, 6-7, 380-382
  epidemiología y, 377-379, 378*f*-379*f*
  primaria, 6, 6*t*, 353, 355-356, 377
  secundaria, 6, 6*t*, 353, 355-356, 378
  terciaria, 6, 6*t*, 355-356

tipos, 6, 6*t*
tratamiento y, 18-19, 18*f*
Principio de utilidad, 397-399
Privacidad, 399-401, 400*f*
Privacy Protection Study Commission, 401
Probabilidad, 246
Pronóstico
descripción, 2
generalización de los datos de supervivencia, 142-143, 143*f*
métodos de expresión, 143*c*
descripción, 124-125
método de Kaplan-Meier, 134-136, 134*f*-136*f*, 134*t*, 282*f*
mortalidad, 125
personas-años, 125-126, 125*f*-126*f*
supervivencia
a los 5 años, 126-128, 126*f*-128*f*
observada, 128-134
relativa, 141-142, 141*t*, 142*f*
tiempo de supervivencia mediana, 141
Proporción(es)
de la incidencia total, 263
objetivo, 46
relativa, *Véase* Cociente de posibilidades
tasas, comparación con, 50-51
Prueba(s)
con variables continuas, 98-101, 98*f*
de cribado
análisis de la relación coste-beneficio, 373-374
aplicaciones, 108
cáncer cervical uterino, 370-371
dicotómicas, 95-97, 97*t*, 101*f*
eficacia, 354*c*

en dos fases
descripción, 101-102
ejemplo, 102*f*
especificidad
descripción, 95, 96*t*
evaluación, 372, 373*f*
evaluación
aspectos metodológicos, 358-363
cribado
del cáncer cervical uterino, 370-371
del neuroblastoma, 371-372, 371*f*, 372*t*
mediante mamografía, 367-370, 368*f*
diseños de los estudios, 363-372
estudios
con asignación aleatoria, 364-367, 365*f*-366*f*
con diseño de casos y controles, 363-364
sin asignación aleatoria, 363-364
Health Insurance Plan Study, 364-365, 365*f*, 367*f*-368*f*
muestreo sesgado por la cantidad, 359-360, 360*f*
resultados que demuestran el beneficio nulo del cribado, 373
sesgo
de clasificación errónea, 363
de periodo de espera, 360-361, 360*f*, 362*f*
de remisión, 358-359, 359*f*
de sobrediagnóstico, 363
supervivencia a los 5 años, 361-363
fiabilidad, 94-122
validez y, relación entre, 117, 117*f*
variación
interobservador, 113-117, 114*f*
intraindividual, 112, 112*f*
intraobservador, 113, 113*f*

mamografía, 365, 367-370, 368*f*, 373
múltiples, 101-105, 101*t*-103*t*, 102*f*-104*f*, 105*t*, 106*f*-107*f*, 107*t*
neuroblastoma, 371-372, 371*f*, 372*t*
prueba de Papanicolaou, 357-358, 370
resultados de estudios que no muestran beneficios, 373
secuencial
descripción, 101-102
ejemplo, 102*f*
pruebas simultáneas y, comparaciones entre, 104-105
sensibilidad
descripción, 95, 96*t*
evaluación, 372, 373*f*
simultáneo
descripción, 102
especificidad neta, 101*t*-103*t*, 103-104, 106*f*
pruebas secuenciales y, comparaciones entre, 106*f*
sensibilidad neta, 101*t*-103*t*, 103-104, 104*f*, 105*t*
validez, 94-122, 96*t*-97*t*, 117*f*
valor predictivo, 105-111, 107*t*
negativo, 107
positivo, 106-108, 108*t*
prevalencia de la enfermedad y, 108-110, 108*f*-109*f*, 108*t*
variables continuas, 98-101, 98*f*
diagnósticas, validez y fiabilidad, 94-122
estadística unilateral, 220
Punto crítico, 356, 357*f*

R

Radiación
exposición pediátrica, 170, 183-184
tumores cerebrales y, 170
Rasgo, 328*t*-331*t*

hereditario, 328*t*-331*t*
poligénico, 328*t*-331*t*
Raza, 402-403
Recuerdo
diferencial, 167
limitaciones, 165-166
problemas, 165-167
sesgo asociado, 166-167, 293
Relación(es)
causales
evidencia, 275-276
necesarias pero no suficientes, 274-275, 275*f*
necesarias y suficientes, 274, 274*f*
ni suficientes ni necesarias, 275, 275*f*-276*f*
suficientes pero no necesarias, 275, 275*f*
coste-beneficio, 373
dosis-respuesta, 277, 345
genética, 328*t*-331*t*
análisis, estudios familiares, 317-320, 317*f*-319*f*
Replicación de los hallazgos, 277
Resultado(s)
estudios, 335
falsamente
negativos, 96, 99
positivos, 96-97, 99
interpretación, 405-407, 405*f*-406*f*, 406*t*-407*t*
parámetros, 335-336, 336*c*, 354, 354*c*
verdaderamente
negativos, 96
positivos, 96
Revisión
bibliográfica, 390

sistemática, 388
Revistas de medicina con revisión editorial por especialistas, 405, 406*f*
Riesgo
absoluto, 240
antecedentes, 259-260
aspectos normativos, 393-394
atribuible
a la población, 261-262, 266*t*
cálculo, 259
definición, 259
descripción, 300
grupos expuestos, 259-261, 260*f*, 266*t*
representación esquemática, 261*f*
riesgo relativo, comparación con, 264-266
de fondo, 259-260
estimación, 240-258
función del epidemiólogo en la comunicación, 406-407
involuntario, 406*t*
potencial para la prevención, 259-268
relativo, 222-223, 242-245, 301
definición, 242
ecuación, 242
estimación mediante el cociente de posibilidades, 249-250, 249*f*-250*f*
estudios con diseño de cohortes, 243-245, 243*t*
importancia, 259
interpretación, 242-243, 243*t*
riesgo atribuible, comparación con, 264-266
voluntario, 407*t*

S

Salmonella typhimurium, 31, 31*f*
Salud comunitaria, 3-5

Seguimiento, 134-135
activo, 42
datos de mortalidad utilizados, 43
definición, 41-43
estudios con diseño de cohortes, 181
factores de riesgo ambientales evaluados, 43
países en desarrollo, 42-43
pasivo, 42-43
tras la comercialización, 227
Semmelweis, Ignáz, 9-12, 10*f*
Sensibilidad, 26-27, 95, 96*t*
evaluación, 372, 373*f*
factores determinantes, 378
neta, 101*t*-103*t*, 103-104, 105*t*
cálculo, 101
con dos pruebas simultáneas, 103-104, 103*t*, 104*f*, 105*t*
pruebas secuenciales, 101, 102*f*
Series de casos clínicos, 149-150, 199
Servicios asistenciales, evaluaciones, 333-352
datos
agrupados, 336-341, 337*f*
asistencia prenatal, 339
desventajas, 338
resultados de la investigación, 337-339
sesgos, 339-340
individuales, 341-350
diseño(s)
antes-después, 344-345
con asignación aleatoria, 342-344, 343*f*
de combinación, 348-350, 348*f*
programa-ausencia de programa, 345-346, 345*t*, 346*f*
efectividad, 335

eficacia, 335
eficiencia, 335
estudios
con diseño de casos y controles, 350
de evolución, 335
de proceso, 334-335
epidemiológicos, 336
mortalidad evitable, 341
parámetros de evolución, 335-336, 336*c*
tempranos, 334*f*
Sesgo
de clasificación errónea, 292, 363
de deseo, 293
de exclusión, 291
de información, 292-294
descripción, 292
estudios con diseño
de casos y controles, 165-167, 166*t*-167*t*
de cohortes, 186
tipos, 293*c*
de notificación, 293
de publicación, 234, 390
de remisión, 358-359, 359*f*
de seguimiento, 293
de selección, 289-292
de los participantes en un estudio, comparación con, 290
definición, 289
ejemplo, 214, 290
estudios con diseño
de casos y controles, 159-160
de cohortes, 185
muestreo sesgado por la cantidad, 359-360, 359*f*-360*f*

sesgo
de exclusión, 291
de periodo de espera, 360-361, 360*f*, 362*f*
de remisión, 358-359, 359*f*
de sobrediagnóstico, 363, 364*f*
del periodo de espera, 360-361, 360*f*, 362*f*
definición, 289
estudios de diseño
de casos y controles
información, 165-167, 166*t*-167*t*
selección, 159-160
de cohortes, 185-186
evaluaciones de los servicios asistenciales, uso de datos de grupos, 339-340
potencial, 404
recuerdo, 166-167, 293
sobrediagnóstico, 363, 364*f*
vigilancia, 293
voluntario, 366
SIDA, *Véase* Síndrome, de inmunodeficiencia adquirida (SIDA)
Síndrome
de inmunodeficiencia adquirida (SIDA), 33-34, 77-78, 81*f*
eosinofilia-mialgia, 278, 278*f*
Sinergia, 302
Sistema
de clasificación, 57*f*
de pago prospectivo, 344-345
Snow, John, 14-16, 14*f*, 397
Staphylococcus aureus, bacteriemia, 55, 56*f*, 56*t*
Supervivencia
a los 5 años, 126-128, 126*f*-128*f*, 361-363
observada, 128-134, 141*t*
relativa, 141-142, 141*t*, 142*f*

Surgeon General de EE. UU., directrices para las inferencias causales, 390-391

T

Tabaco, *Véase* Consumo de cigarrillos, *Véase también* Tabaquismo

Tabaquismo, 16

cáncer

de pulmón y, 16, 158-159, 159*t*, 162-163, 265*t*, 277, 279*f*, 297*f*, 297*t*, 302*t*-303*t*

esofágico y, 298*f*, 299

carcinoma epidermoide y, 321, 322*f*

coronariopatía y, 158, 158*t*, 243, 243*t*, 262, 263*t*, 264, 265*t*

diseño de un estudio de cohortes, 178, 179*t*

materno, 272*f*-273*f*

Tabla de mortalidad

cálculo, 132-134, 132*t*-133*t*

de las tasas de supervivencia, 128*t*-131*t*, 131*c*-132*c*, 132*f*

ejemplo de uso, 137-138, 137*f*-138*f*

fundamento, 128-131

suposiciones, 136-138

Tabulación cruzada, 36-38, 38*t*

Tamaño de la muestra, 216-222, 217*c*, 217*f*-219*f*, 219*c*, 219*t*, 221*t*

Tamoxifeno, prevención del cáncer de mama, 230-231, 230*f*

Tasa(s)

de ataque, 32, 37*c*, 38*t*, 50-51

específica de los alimentos, 32, 37*c*

secundaria, 32

de concordancia, en los gemelos, 310, 310*f*-312*f*

de discordancia, en los gemelos, 310

de incidencia, 45-49

cáncer

de mama, 75*f*-76*f*

tiroideo, 75-76, 77*f*-78*f*

uterino, 50*f*
casos recién detectados de la enfermedad utilizados para el cálculo, 45, 49*f*
denominador, 45
ecuaciones para el cálculo, 45
ejemplos, 45, 45*f*
tasa de mortalidad y, 73
de mortalidad, 68-69
ajuste por edad, 82-84, 339, 339*f*-340*f*
anual, 68
cáncer de pulmón, 68
con especificidad respecto
a la enfermedad, 68
a la raza, 80*t*-81*t*
denominador, 125
diabetes mellitus, 80*f*
específica, 68
por edad, 68, 81*t*, 83
mortalidad, comparación con, 69, 69*c*, 125
por causas específicas, 68
por grupo de edad, 68, 83, 277*f*
de respuesta, 220
de supervivencia
a los 5 años, 126-128, 126*f*-128*f*, 361-363
cálculo, probabilidades de supervivencia, 128*t*-131*t*, 131*c*-132*c*, 132*f*
específica, 68
objetivo, 45
proporciones, comparación con, 50-51
Tendencias cronológicas, incidencia de las enfermedades, 313-317, 315*f*
Tenofovir/emtricitabina, 18
Teoría miasmática, 14
Terrorismo biológico, viruela como, 13-14

THS, *Véase* Tratamiento, hormonal sustitutivo (THS)
Tos ferina, 32-33, 33*f*-34*f*
Tracto digestivo, 21
Transición epidemiológica, 89*f*
Transmisión
  directa, 20, 21*c*
  enfermedad, 20-40
    aspectos generales, 20, 41
    directa, 20, 21*c*
    gotículas, 21*f*
    indirecta, 20, 21*c*
    superficies corporales, 21, 22*f*
  indirecta, 20, 21*c*
  por gotículas, 21*f*
Trastornos autosómicos
  dominantes, 317-318, 318*f*
  recesivos, 317-318, 319*f*
Tratamiento
  asignado y recibido, 206
  hormonal sustitutivo (THS), 382, 383*f*
    Women's Health Initiative, 383, 383*f*
  prevención y, 18-19, 18*f*
  resultados de los ensayos clínicos con asignación aleatoria, 222
Tríada epidemiológica, 20, 21*f*
Tribunales, 390-391
Trombosis venosa, 320-321
Truvada, 18
Tuberculosis, 50*f*
  cáncer y, 160, 160*t*
Tumores cerebrales, 169, 169*f*-170*f*

U

Úlceras
duodenales, 281*c*
pépticas, 280-281, 281*c*
Umbral, 97, 99
U.S. Preventive Services Task Force, 283, 283*t*, 285*f*, 286*t*, 370
Utilidad, principio, 397-399

V

Vacuna de la poliomielitis oral, 28
Vacunación, 13
Validez
externa, ensayo clínico con asignación aleatoria, 223-224, 224*f*, 290
interna, ensayo clínico con asignación aleatoria, 224, 224*f*, 290
pruebas de cribado, 94-122, 96*t*-97*t*, 117*f*, 290
Valor predictivo, 105-111, 107*t*
negativo, 107
positivo, 106-108, 108*t*
descripción, 106-108
especificidad y, 110-111, 110*f*, 111*t*
prevalencia de la enfermedad y, 108-110, 108*t*
prevalencia de la enfermedad y, 108-110, 108*f*-109*f*, 108*t*
Variación
interobservador, 113-117, 114*f*
intraindividual, 112, 112*f*
intraobservador, 113, 113*f*
Variolización, 12
Verosimilitud biológica, 277
Vibrio cholerae, 15-16
Vigilancia
activa, 42
pasiva, 42
VIH, *Véase* Virus, de la inmunodeficiencia humana (VIH)

Viruela, 12-14
bovina, 13-14
Virus
de la inmunodeficiencia humana (VIH)
aspectos de confidencialidad, 397
datos de mortalidad, 73-74
en los países industrializados, 3-5
Truvada, 18
del Nilo occidental, 35-36, 36*f*-37*f*
del papiloma humano (VPH), 395
VPH, *Véase* Virus, del papiloma humano (VPH)

W

Western blot, prueba, 372
Will Rogers, fenómeno, 141
Wilson, Robert, 382
Women's Health Initiative, 383, 383*f*

Y

Yerushalmy, Jacob, 272-274

Z

Zika, virus
áreas con riesgo, 149, 151*f*
durante los embarazos, 149, 150*f*

Instrucciones para el acceso en línea

Gracias por su compra. Este libro electrónico de Elsevier incluye el acceso a contenido online complementario. Por favor, haga clic aquí (o vaya a http://ebooks.elsevier.com) para solicitar un código de activación y siga las instrucciones de registro para acceder al contenido en línea.

Índice

Instrucciones para el acceso en línea 989
Portada 2
Índice de capítulos 3
Página de créditos 12
Dedicatoria 14
Prefacio 15
Agradecimientos 19
Sección 1: Abordaje Epidemiológico de la Enfermedad y la Intervención 22
Introducción 23
Capítulo 1: Introducción 25
¿Qué es la epidemiología? 27
Objetivos de la epidemiología 28
Epidemiología y prevención 34
Epidemiología y práctica clínica 37
Planteamiento epidemiológico 40
De observaciones a acciones preventivas 44
Cuando la frecuencia de una enfermedad disminuye, ¿de quién es el mérito? 59
Integración de la prevención y el tratamiento 62
Conclusión 65
Capítulo 2: Dinámica de la transmisión de las enfermedades 68
Modos de transmisión 72
Enfermedad clínica y subclínica 75
Estado de portador 79
Endémico, epidémico y pandémico 80
Brotes de enfermedad 84
Inmunidad y susceptibilidad 87
Inmunidad grupal 88
Periodo de incubación 91
Tasa de ataque 96
Exploración de la ocurrencia de la enfermedad 98
Investigación del brote 106
Conclusión 110
Capítulo 3: La ocurrencia de la enfermedad: I. Vigilancia de la enfermedad

y medidas de la morbilidad
Vigilancia 119
Etapas de la enfermedad en un individuo y en una población 123
Medidas de la morbilidad 127
Conclusión 160
Capítulo 4: La ocurrencia de la enfermedad: II. Mortalidad y otras medidas del impacto de la enfermedad 167
Medidas de la mortalidad 170
Comparación de la mortalidad en diferentes poblaciones 202
Otras medidas del impacto de la enfermedad 212
Conclusión 216
Capítulo 5: Evaluación de la validez y fiabilidad de las pruebas diagnósticas y de cribado 227
Variación biológica de las poblaciones humanas 230
Validez de las pruebas de cribado 233
Uso de pruebas múltiples 243
Valor predictivo de una prueba 255
Fiabilidad (repetibilidad) de las pruebas 264
Relación entre validez y fiabilidad 274
Conclusión 276
Apéndices del capítulo 5 278
Capítulo 6: La historia natural de la enfermedad: formas de expresar el pronóstico 285
Tasa de letalidad 291
Personas-años 292
Supervivencia a cinco años 296
Supervivencia observada 299
El método Kaplan-Meier 310
Suposiciones empleadas al usar tablas de vida y el método de Kaplan-Meier 315
Efectos aparentes sobre el pronóstico de las mejoras diagnósticas 321
Mediana de supervivencia 325
Supervivencia relativa 326
Generalización de los datos de supervivencia 328
Conclusión 330
Sección 2: Empleo de la Epidemiología Para Identificar las Causas de las Enfermedades 336
Introducción 337
Capítulo 7: Estudios observacionales 341

Capítulo 7: Estudios observacionales 341
Casos clínicos y series de casos 344
Estudios ecológicos 347
Estudios transversales 353
Estudios de casos y controles 359
Conclusión 391
Capítulo 8: Estudios de cohortes 399
Diseño de un estudio de cohortes 402
Selección de las poblaciones de estudio 404
Tipos de estudios de cohortes 406
Ejemplos de estudios de cohortes 409
Estudios de cohortes para investigar la salud y las enfermedades infantiles 414
Posibles sesgos en los estudios de cohortes 418
¿Cuándo se requiere un estudio de cohortes? 420
Estudios de casos y controles basados en una cohorte definida 423
Conclusión 429
Capítulo 9: Comparación de los estudios de casos y controles y de cohortes 435
Capítulo 10: Valoración de las medidas preventivas y terapéuticas: ensayos aleatorizados 444
Selección de sujetos 450
Asignación de sujetos a grupos de tratamiento sin aleatorización 451
Asignación de sujetos utilizando aleatorización 456
Recogida de datos sobre los sujetos 464
Diseño cruzado 469
Diseño factorial 474
Falta de cumplimiento 478
Conclusión 481
Capítulo 11: Ensayos aleatorizados: algunos aspectos adicionales 485
Tamaño de la muestra 488
Reclutamiento y retención de participantes del estudio 500
Formas de expresar los resultados de los ensayos aleatorizados 502
Interpretación de los resultados de los ensayos aleatorizados 504
Las cuatro fases para probar nuevos fármacos en Estados Unidos 509
Cinco ensayos aleatorizados importantes en Estados Unidos 511
Ensayos aleatorizados para evaluar intervenciones ampliamente aceptadas 520
Registro de ensayos clínicos 526
Consideraciones éticas 528

Conclusión 530
Epílogo 531
Capítulo 12: Estimación del riesgo: ¿existe una asociación? 540
Riesgo absoluto 544
¿Cómo se puede determinar si una enfermedad concreta se asocia a una exposición específica? 545
Riesgo relativo 548
Cociente de posibilidades (posibilidades relativas) 554
Conclusión 572
Apéndice del capítulo 12 578
Capítulo 13: Conceptos adicionales sobre el riesgo: estimación de las posibilidades de prevención 583
Riesgo atribuible 585
Comparación de riesgo relativo y riesgo atribuible 596
Conclusión 599
Preguntas de revisión del capítulo 13 601
Apéndice del capítulo 13: fórmula de Levin para calcular el riesgo atribuible para la población total 603
Capítulo 14: De la asociación a la causalidad: derivación de inferencias a partir de estudios epidemiológicos 605
Estrategias para el estudio de la etiología de las enfermedades 608
Tipos de asociaciones 612
Tipos de relaciones causales 620
Evidencia de una relación causal 625
Directrices para juzgar si una asociación observada es causal 627
Derivación de inferencias causales: dos ejemplos 634
Modificaciones de las directrices para las inferencias de causalidad 641
Conclusión 648
Preguntas de repaso del capítulo 14 651
Capítulo 15: Más sobre inferencias causales: sesgos, confusión e interacción 654
Sesgo 657
Confusión 668
Interacción 678
Conclusión 690
Capítulo 16: Identificación de los factores genéticos y ambientales en la causalidad de la enfermedad 697
Genética tradicional 700
Enfermedades complejas 702
Tendencias temporales en la incidencia de la enfermedad 713

Tendencias temporales en la incidencia de la enfermedad 713
Análisis de vinculación en estudios familiares 720
Interacción entre los factores genéticos y los factores de riesgo ambientales 725
Medicina de precisión 730
Perspectivas de futuro 732
Conclusión 735
Glosario de los términos genéticos del capítulo 16 742
Sección 3: Aplicación de la Epidemiología a la Evaluación y al Desarrollo de Políticas Sanitarias 747
Introducción 748
Capítulo 17: Uso de la epidemiología para evaluar los servicios de salud 750
Estudios de proceso y de resultado 755
Eficacia, efectividad y eficiencia 757
Medidas de resultado 758
Comparación entre los estudios epidemiológicos sobre la etiología de la enfermedad y la investigación epidemiológica que evalúa la efectividad de los servicios de salud 760
Evaluación mediante el uso de datos de grupo 761
Evaluación utilizando datos individuales 773
Conclusión 790
Capítulo 18: Enfoque epidemiológico en la evaluación de los programas de cribado 797
Historia natural de la enfermedad 803
Patrón de progresión de la enfermedad 810
Problemas metodológicos 813
Diseños de estudio para la evaluación del cribado: estudios no aleatorizados y aleatorizados 825
Problemas en la evaluación de la sensibilidad y la especificidad de las pruebas de cribado 840
Interpretación de los resultados de estudios que no muestran beneficios del cribado 843
Análisis de rentabilidad del cribado 844
Conclusión 846
Capítulo 19: Epidemiología y políticas de salud pública 854
Epidemiología y prevención 858
Estrategias poblacionales frente a estrategias de alto riesgo para la prevención 863
Epidemiología y medicina clínica: tratamiento hormonal sustitutivo en 867

Evaluación del riesgo 871
Metaanálisis 877
Sesgo de publicación 881
Epidemiología en la vía judicial 883
Fuentes y efecto de la incertidumbre 885
Aspectos políticos referentes al riesgo: ¿cuáles deberían ser los objetivos? 890
Conclusión 891
Capítulo 20: Aspectos éticos y profesionales en epidemiología 896
Problemas éticos de la epidemiología 902
Obligaciones de los investigadores con los sujetos de estudio 903
Protección de la privacidad y la confidencialidad 906
Acceso a los datos 912
Raza y etnicidad en los estudios epidemiológicos 913
Conflicto de intereses 917
Interpretación de los resultados 919
Conclusión 924
Respuestas a las preguntas de repaso 927
Índice alfabético 946