ActivEpi Español

Libro de texto de acompañamiento para uso suplementario con el CD-ROM de ActivEpi Español

David G. Kleinbaum, Dinorah L. Calles y Kevin M. Sullivan

Junio del 2013

David G. Kleinbaum
Department of Epidemiology
Emory University
Atlanta, GA 30322
USA
dkleinb@sph.emory.edu

Dinorah L. Calles
Epidemic Intelligence Service Officer,
Centers for Disease Control and Prevention
Consultant, Emory University
Atlanta, GA 30322
USA
dcalles@alum.emory.edu

Kevin M. Sullivan
Department of Epidemiology
Emory University
Atlanta, GA 30322
USA
cdckms@sph.emory.edu





Tabla de Contenidos

Lección 1 Introducción 1

I. OBJETIVOS Y MÉTODOS EN INVESTIGACIÓN EPIDEMIOLÓGICA

Lección 2 Investigación Epidemiológica: una revisión 7
Lección 3 Diseños de Estudios Epidemiológicos 30
Lección 4 Medidas de Frecuencia de Enfermedad 68
Lección 5 Medidas de Efecto 112
Lección 6 Medidas de Impacto Potencial 150

II. VALIDEZ DE LA INVESTIGACIÓN EPIDEMIOLÓGICA

Lección 7 Validez de la Investigación Epidemiológica 195
Lección 8 Sesgo de Selección 221
Lección 9 Sesgo de Información 264
Lección 10 Confusión Epidemiológica 328
Lección 11 Confusión Involucrando Varios Factores de Riesgo 362

III. ANÁLISIS EPIDEMIOLÓGICO

Lección 12 Inferencia Estadística Sobre Medidas de Efecto 385
Lección 13 Control de Factores Externos 456
Lección 14 Análisis Estratificado 497
Lección 15 Equiparamiento 565

ÍNDICE 614

Prefacio

Este libro de texto de epidemiología es un complemento al CD-ROM de *ActivEpi Español para Windows*. El CD-ROM de *ActivEpi Español para Windows* ofrece una presentación multimedia de conceptos epidemiológicos que comúnmente se abordan en un curso de introducción a la epidemiología. El CD-ROM utiliza una serie de efectos multimedia para motivar, explicar, visualizar y aplicar conceptos introductorios epidemiológicos, integrando vídeo, animación, narración, texto y sesiones de preguntas y respuestas interactivas. Dadas las diferencias en aprendizaje de los lectores, el CD-ROM y el libro de texto de acompañamiento ofrecen a los usuarios distintas opciones, no obstante entrelazadas, para facilitar su aprendizaje de epidemiología. El CD-ROM ofrece una forma dinámica de presentar conceptos epidemiológicos a través del uso de la animación. El libro de texto de acompañamiento, por otro lado, puede ser utilizado como una referencia impresa de los materiales de texto que contiene el CD-ROM, como un recurso para los ejercicios de práctica, como referencia general, o incluso como un libro de texto independiente y auto-suficiente.

En general, todo el material en el CD-ROM está incluido en el libro de texto de acompañamiento. Esta primera versión del CD-ROM y el libro de texto de acompañamiento *ActivEpi Español* se elaboró en varias etapas de producción y a través de distintos niveles de revisión. Por lo tanto, algunos guiones de la narración en el CD-ROM se alteraron en el libro de texto de acompañamiento durante ese proceso. Esta diferencia se produce principalmente cuando la narración del CD-ROM se enfoca en contenido animado en la pantalla. Solicitamos al lector que detectara alguna discrepancia, se sirva escribir a los autores (apoyoactivepi@gmail.com) para disipar dudas y efectuar correcciones necesarias en las subsecuentes ediciones.

LECCIÓN 1

ActivEpi Español

1-1 Primeros pasos: Libro de estudio

INTRODUCCIÓN

La epidemiología es el estudio de la salud y enfermedades en poblaciones humanas o de otra naturaleza (por ejemplo, veterinarias). En este libro de texto, analizamos problemas de salud y enfermedades del mundo real y mostramos cómo los conceptos y métodos epidemiológicos nos permiten estudiar, comprender, y resolver dichos problemas. Y, más importante aún, es el hecho de que aplicamos cada concepto o método nuevo a medida que lo desarrollamos para ayudar al lector a adquirir una mayor comprensión del tema.

Acerca de los autores del libro de texto *ActivEpi Español*:
David G. Kleinbaum, Dinorah L. Calles y Kevin M. Sullivan

David G. Kleinbaum es Profesor de Epidemiología en la Escuela de Salud Pública Rollins de la Universidad Emory, en Atlanta, Georgia, y ha sido reconocido internacionalmente como experto en la enseñanza de conceptos y métodos bioestadísticos y epidemiológicos a todo nivel. Es autor de varios libros de texto muy reconocidos, incluyendo *Applied Regression Analysis and Other Multivariable Methods*; *Epidemiologic Research: Principles and Quantitative Methods*; *Logistic Regression-A Self-Learning Text*; y *Survival Analysis-A Self-Learning Text*.



El Dr. Kleinbaum tiene más de 40 años de experiencia en el dictado de más de 230 cursos cortos sobre métodos estadísticos y epidemiológicos para una gran variedad de audiencias internacionales. Además, es autor de varias publicaciones tanto en el campo metodológico como en el de salud pública aplicada. También es un experimentado y solicitado asesor y, en la actualidad, se desempeña como asesor *ad hoc* de todo el personal de investigación de los Centros para el Control y la Prevención de Enfermedades (CDC, por sus siglas en inglés) Ha recibido numerosos premios en docencia, y recibió el primer Premio Pfizer de la Asociación de Escuelas de Salud Pública en 2005 por su carrera distinguida en docencia.

En cuanto a lo personal, el Dr. Kleinbaum es un consumado flautista de jazz y toca regularmente en Atlanta con su conjunto de jazz, *The Moonlighters Jazz Band*.

La Dra. Dinorah L. Calles es una oficial del Servicio de Inteligencia Epidémica (*EIS Class of 2013*) de los CDC, Atlanta, Georgia. La Dra. Calles obtuvo su PhD en Epidemiología en la Universidad Emory, y ha coordinado todas las etapas de producción del CD-ROM y el libro de texto de acompañamiento de *ActivEpi Español*. Su voz introduce el CD-ROM de *ActivEpi Español.*



Ha dictado varios cursos y talleres de epidemiología como una docente asociada e instructora titular para distintas audiencias de profesionales de salud. Recibió el premio *J. William Fulbright* para desempeñar su trabajo de campo de doctorado en Rio de Janeiro, Brasil. Recientemente, se destacó como asistente de coordinación de los *Módulos de Princípios de Epidemiologia para o Controle de Enfermidades* para plataformas de educación a distancia junto a la OPS Brasil.

Posee maestría en Salud Pública (University of North Texas Health Science Center, Fort Worth, Texas) y formación en música (Dartmouth College, Hanover, New Hampshire). Es originaria de San Salvador, El Salvador.

El Dr. Kevin M. Sullivan es Profesor Adjunto de Epidemiología en la Escuela de Salud Pública Rollins de la Universidad Emory, y ha trabajado en el campo de la epidemiología y salud pública por más de 30 años, presentado más de 80 trabajos en publicaciones y publicado capítulos en varios libros. Es uno de los desarrolladores de los programas informáticos OpenEpi (www.OpenEpi.com) y Epi Info (www.cdc.gov/EpiInfo) y autor del libro de texto *ActivEpi* en su versión original en inglés.

Agradecimientos

La versión original en inglés de *ActivEpi* brinda agradecimientos a varios contribuyentes al desarrollo de esa versión. En su adaptación al español, los autores reservan sus agradecimientos a aquellos que han contribuído al desarrollo de *ActivEpi Español*.

Los autores agradecen especialmente a la Organización Panamericana de la Salud (OPS), Oficinia Regional para las Américas de la Organización Mundial de la Salud (OMS), por la prestación de apoyo financiero para completar las etapas finales del desarrollo del CD-ROM de auto-aprendizaje y libro de texto de acompañamiento de *ActivEpi Español*, además de la anticipada distribución de *ActivEpi Español* en países integrantes de las Américas.

Los autores extienden su gratitud al Dr. Oscar Mujica, Asesor Regional de la OPS en Epidemiología Social, por su entusiasmo y reconocimiento del enorme potencial de este producto como herramienta de instrucción en la capacitación de metodología epidemiológica en América Latina.

Agradecemos a la Dra. Karen Andes del Departamento *Hubert* de Salud Global de la Escuela de Salud Pública de la Universidad Emory, investigadora principal del proyecto cuya iniciativa dio inicio al ahora producto final de *ActivEpi Español.* El proyecto de la Dra. Andes fue obtenido de *Higher Education for Development*, con presupuesto de *USAID Paraguay*, y se intitula "*Emory/Paraguay Health Information Partnership for Progress."*

Reconocemos la contribución del Dr. Kevin M. Sullivan del Departamento de Epidemiología de la Escuela de Salud Pública de la Universidad Emory, quien es el autor del libro de texto *ActivEpi* en su versión original en inglés y el cual es la fuente principal de la versión en español del texto de *ActivEpi Español*. Por esta aportación, el Dr. Sullivan es nombrado como el tercer autor del libro de texto de acompañamiento de *ActivEpi Español*.

Los autores además agradecen a *Trusted Translations, Inc.,*quien produjo el borrador traducido al español de la versión original del libro de texto de *ActivEpi* al español.

Primeros pasos

Este curso utiliza recursos multimedia para:

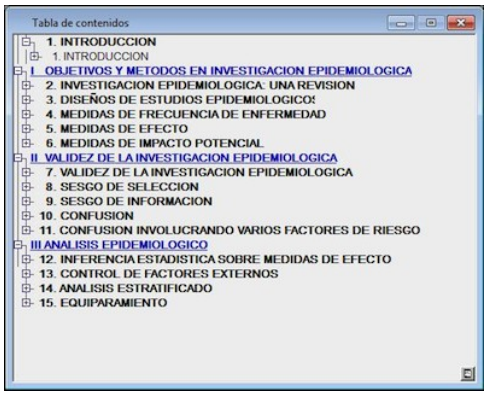
- *mostrar ejemplos reales;*
- *permitirle aplicar métodos a medida los aprende;*
- *permitirle controlar su aprendizaje en privado; y*
- *suministrar una gran variedad de ejercicios, auto-evaluaciones y ejemplos para que pueda llevar su aprendizaje más allá del libro de texto y el CD-ROM de ActivEpi Español.*

La epidemiología es el estudio de la salud y enfermedades en poblaciones humanas. Por ejemplo, un ensayo clínico aleatorizado realizado por epidemiólogos en la Facultad de Salud Pública de Harvard demostró que la ingesta de aspirina reduce de 20% a 30% el riesgo de sufrir ataques cardíacos. Según se demostró en estudios de salud pública de la década de los cincuenta, fumar cigarrillos provoca cáncer pulmonar. Los epidemiólogos ambientales han estado evaluando la evidencia para determinar si los niños que viven cerca de las líneas de alta tensión corren un alto riesgo de padecer leucemia. Los investigadores del cáncer se preguntan por qué es menos probable que las mujeres mayores se sometan a un examen de cáncer de mama, en comparación con las más jóvenes. Todos estos casos son ejemplos de investigaciones epidemiológicas, puesto que todos intentan describir la relación que existe entre un resultado de salud y una o más explicaciones o causas de dicho resultado. Todos estos ejemplos comparten varios desafíos: deben elegir un diseño de estudio apropiado, deben ser cuidadosos para evitar sesgos y deben emplear métodos estadísticos adecuados para analizar los datos. La epidemiología enfrenta cada uno de estos tres desafíos.

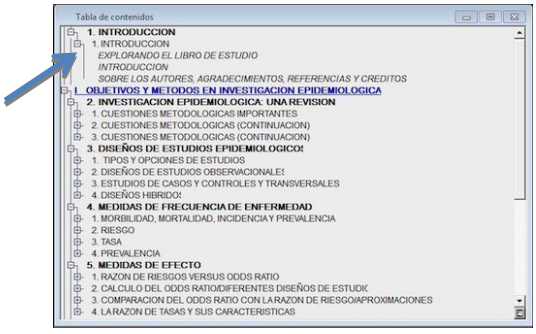
1-2 Funciones globales

Funciones globales del curso en CD-ROM

Tabla de contenidos

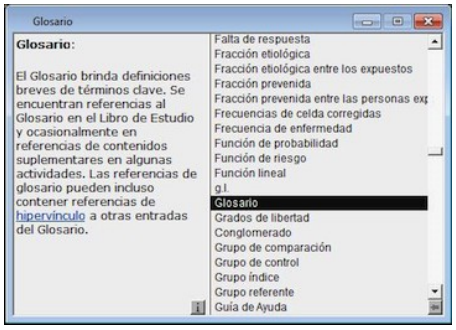


La tabla de contenidos cumple 3 funciones importantes: primero, ofrece una descripción general del curso mediante una lista de las lecciones en el orden en que aparecen. Segundo, proporciona una forma rápida de acceder a cualquier lección del curso. Sólo haga clic en el título de cualquier lección para dirigirse a la primera página de dicha lección. Finalmente, la tabla de contenidos también le muestra el punto del curso en el que se encuentra.



La tabla de contenidos se abre para mostrar cada una de las páginas de la lección actual y abre la página actual para mostrar cada una de las actividades de esa página. Las actividades marcadas como vistas en la página también se identifican con marcas de verificación en la Tabla de contenidos. Haga clic en cualquier renglón de la tabla para dirigirse a la página correspondiente o incluso a una actividad específica. Para cerrar la tabla de contenidos, haga clic en el botón .

Glosario

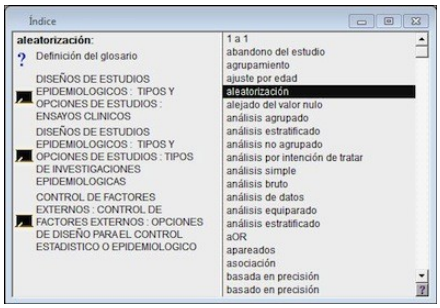


A lo largo del libro de estudio del CD-ROM encontrará palabras resaltadas con color y subrayadas. Estas palabras son términos del glosario. Haga clic en cualquier término del glosario para abrirlo en la definición correspondiente.

La ventana del glosario contiene las definiciones de los términos que se presentan en el curso. Haga clic en el término deseado en el panel de la derecha para acceder a su definición. La definición se muestra en el panel adyacente de la izquierda. Puede desplazarse por la lista en orden alfabético para encontrar un término específico, o bien presione cualquier tecla para buscar un término que empiece con esa letra.

Si una definición fuera demasiado larga para la ventana, desplácese por la ventana o ajuste su tamaño. La mayoría de las definiciones hacen referencia a otras definiciones. Haga clic en cualquier término de color para ver su definición. Para cerrar el glosario, haga clic en el botón .

Índice



La ventana del Índice muestra una lista de términos en orden alfabético por la que puede desplazarse. Haga clic en el término que desea ubicar. En el panel adyacente, se mostrarán vínculos a las referencias en el curso. El ícono junto a la referencia indica el tipo de actividad, igual que en el caso de la tabla de contenidos. La referencia incluye el nombre de la lección, el nombre de la página y el nombre de la actividad. Haga clic en el ícono para dirigirse a la página correspondiente del libro de estudio del CD-ROM.

Muchos de los términos indexados se definen en el glosario. Cuando exista una definición del glosario, aparece tanto en las referencias como al final de la barra de desplazamiento de las referencias. Para encontrar un término específico, puede desplazarse por toda la lista en orden alfabético. Para cerrar el índice, haga clic en el botón que se encuentra en la esquina superior derecha de la ventana.

1-3 Aprendizaje eficaz

Aprendizaje eficaz en este curso

Para trabajar de un modo eficaz en este curso, debe dominar las partes claves de su aprendizaje. En especial, debería:

Tomar apuntes: la experiencia física de tomar notas tiene un verdadero valor mnemotécnico. Se recomienda tomar apuntes en papel del modo tradicional. Recuerde que puede pausar una explicación o un video en cualquier momento haciendo clic en el botón stop/play (detener/reproducir) o presionando la barra espaciadora. Si no entiende algo o desea verlo de nuevo, deslice la barra de progreso hasta ese punto.

Controlar las explicaciones: todos nos distraemos en algún momento. Incluso, por más que prestemos mucha atención, hay partes del material que pueden parecer incomprensibles la primera vez que se presentan (la verdad es que nadie entiende estos conceptos al verlos por primera vez; una desventaja de una clase convencional en un aula es que, por lo general, uno no puede hacer que el disertante haga una pausa o rebobine). Aquí usted tiene

pleno control de la explicación del material nuevo. Puede detenerla en cualquier momento para pensar un minuto y asimilar una idea nueva, tomar apuntes, remitirse a la sección correspondiente del texto o dialogar con otro alumno. Puede revisar cualquier parte las veces que lo desee y trabajar a voluntad con cualquier herramienta.

Hacer los ejercicios: como a usted nadie lo está observando, es fácil saltear el material de repaso. No lo haga. En estos ejercicios se presenta material importante del curso. Si pasa por alto estos ejercicios, se perderá conceptos primordiales.

Trabajar en orden: éste es un curso multimedia, con hipertexto y muchas opciones. Pero los métodos epidemiológicos, al igual que la estadística, requieren cierto orden. Puede navegar por el curso con total libertad, pero se dará cuenta de que el material es mucho más comprensible cuando primero se aprenden las ideas elementales y luego se profundizan. Y sepa que el viejo truco de tratar de hacer primero la tarea y luego volver atrás para intentar encontrar un texto similar y copiarlo en la respuesta simplemente no funciona con los materiales multimedia. El contenido a menudo se encuentra dentro de una actividad o ejercicio, por lo que perderá demasiado tiempo buscándolo en lugar de ahorrarse trabajo.

LECCIÓN 2

INVESTIGACIÓN EPIDEMIOLÓGICA: DESCRIPCIÓN GENERAL

2-1 Cuestiones metodológicas importantes

En sus inicios, el campo de la epidemiología se ocupaba de proporcionar una base metodológica para el estudio y el control de epidemias en la población. Ahora, sin embargo, la epidemiología tiene un alcance mucho más amplio, incluyendo el estudio de enfermedades tanto agudas como crónicas, la calidad de la atención médica y los problemas de salud mental. Al ampliarse el foco de la investigación epidemiológica, también se ha expandido la metodología. Esta lección proporciona una descripción general en la que se incluyen ejemplos de investigaciones epidemiológicas y se presentan varias preguntas metodológicas importantes que suelen considerarse en investigaciones de esta naturaleza.

Estudio de usuarios de las playas de Sydney

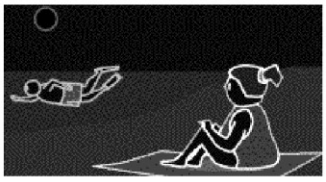
*La **epidemiología** se ocupa principalmente de identificar los factores o variables importantes que influyen en un resultado de salud significativo. En el estudio de usuarios de las playas de Sydney, la pregunta clave era "¿Hay alguna relación entre bañarse en las playas de Sydney y un mayor riesgo de enfermedades infecciosas agudas?".*

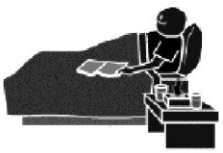
En Sydney, Australia, a lo largo de la década de los ochenta, en los medios de comunicación locales hubo quejas de personas que manifestaron que las playas públicas más concurridas alrededor de la ciudad eran cada vez más inseguras. Básicamente, la preocupación se centraba en la sospecha de que las playas estaban cada vez más contaminadas debido a métodos incorrectos de eliminación de desechos.

En 1989, el Departamento de Salud de New South Wales decidió realizar un estudio para investigar en qué medida el hecho de bañarse y la posible contaminación en 12 playas concurridas de Sydney incidían en la salud pública, en especial, durante los meses de verano, época en la que las playas recibían más personas. La pregunta principal de interés de la investigación era: *¿corren un mayor riesgo de desarrollar una enfermedad infecciosa aguda las personas que se bañan en las playas de Sydney?*

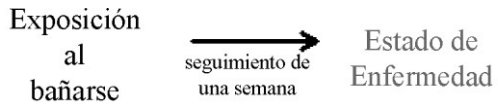
La pregunta de interés:

¿ Corren un mayor riesgo de desarrollar una enfermedad infecciosa aguda las personas que se bañan en las playas de Sydney?





El estudio se llevó a cabo seleccionando individuos en las playas durante los meses de verano de 1989-90. Los individuos que reunieron los requisitos para participar de esta entrevista inicial recibieron una llamada telefónica de seguimiento una semana después para determinar su exposición al bañarse el día de la entrevista en la playa, así como la condición de enfermedad subsiguiente durante la semana posterior a la entrevista.



También se realizaron mediciones de la calidad del agua de las playas cada día que se tomaron muestras de individuos a fin de establecer una correlación entre la exposición al bañarse y los niveles de contaminación en las playas.

Los análisis de los datos del estudio llevaron a la conclusión general de que bañarse en aguas contaminadas implicaba un riesgo 33% mayor de contraer una enfermedad infecciosa que bañarse en aguas no contaminadas, un porcentaje significativo en términos estadísticos. Tanto los funcionarios del departamento de salud como el público en general se basaron en estos resultados para confirmar que bañarse en las playas de Sydney representaba un problema importante para la salud. Por consiguiente, los departamentos de salud del estado y la localidad, junto con otros organismos ambientales del área de Sydney, lanzaron un programa para reducir las fuentes de contaminación del agua de la playa, el cual mejoró la calidad del agua en las playas durante la década de los noventa.

Resumen

- El estudio de usuarios de las playas de Sydney es un ejemplo de cómo se pueden aplicar los principios y métodos epidemiológicos para investigar un problema localizado de salud pública.
- La pregunta clave del estudio de usuarios de las playas de Sydney era:
  - ¿Representa el hecho de bañarse en las playas de Sydney, Australia, (en 1989-90) un mayor riesgo de contraer enfermedades infecciosas agudas?
  - La conclusión fue: "sí, un riesgo 33% mayor".

Cuestiones metodológicas importantes

Proporcionamos un panorama general de la investigación epidemiológica subrayando varias preguntas generales que surgen en el transcurso de la mayoría de este tipo de investigaciones.

Existen muchos problemas preocupantes a la hora de planificar un estudio de investigación epidemiológica (vea el recuadro a continuación). En esta actividad comenzaremos a describir una lista de cuestiones metodológicas generales que deben abordarse. Ilustraremos cada una de las cuestiones utilizando el mencionado estudio de usuarios de las playas de Sydney de 1989.

Cuestiones que deben considerarse al planificar un estudio de investigación epidemiológica	
Pregunta	Definir una pregunta de interés y una serie de variables claves
Variables	Qué medir y cómo; exposición **E**, enfermedad **D** y variables de control **C**
Diseño	¿Qué diseño de estudio y marco de muestreo?
Frecuencia	Medidas de frecuencia de enfermedad
Efecto	Medidas de efecto

Sesgo	Fallas en el diseño de estudio, la recolección de datos o el análisis
Análisis	Realización de análisis apropiados

La primera acción a seguir es definir claramente la **pregunta** de interés del estudio, lo cual incluye especificar las variables claves que deben medirse. Por lo general, preguntamos: *¿cuál es la relación entre uno o más determinantes hipotéticos y una enfermedad o un resultado de salud de interés?*



Los **determinantes** suelen denominarse **variables de exposición** y se expresan con la letra **E**. La enfermedad o el resultado de salud se indican con la letra **D**. Por lo general, además de la exposición y la enfermedad, deben tenerse en cuenta otras variables que pueden predecir el resultado de salud. Estas variables suelen denominarse **variables de control** y se expresan con la letra **C**.

A continuación, debemos determinar cómo se medirán estas **variables**. Este paso requiere la identificación de los instrumentos de recolección de datos y los cuestionarios que deben obtenerse o desarrollarse.

Se obtienen datos de:
- encuestas
- entrevistas
- muestras
- laboratorio

Luego, se debe seleccionar un **diseño de estudio** apropiado e idear un plan de muestreo para inscribir individuos en el estudio. La elección del diseño de estudio y el plan de muestreo depende de la viabilidad y el costo, así como de diversas características de la población que se desea estudiar y el propósito del estudio.

Términos para aprender:
- ensayo clínico
- transversal
- caso y control
- cohorte

Después, medidas de frecuencia de enfermedad y efecto se deben elegir en función del diseño de estudio. Una medida de **frecuencia de enfermedad** aporta información cuantitativa sobre la periodicidad con que se produce un resultado de salud en subgrupos de interés. Una **medida de efecto** permite la comparación entre los subgrupos.

Términos para aprender:

tasa	**razón de riesgo**
proporción	**odds ratio**
riesgo	**razón de prevalencia**
odds	**razón de tasas**
prevalencia	
incidencia	

Además, debemos considerar los **sesgos** potenciales de un estudio. ¿Existe alguna falla en el diseño de estudio, los métodos de recolección de datos o los métodos de análisis de datos que puedan llevar a conclusiones falsas sobre la relación entre la exposición y la enfermedad?

Términos para aprender:
sesgo de selección
sesgo de información
sesgo de confusión

Por último, debemos realizar un **análisis de datos** apropiado, lo cual incluye la estratificación y la modelación matemática según corresponda. El análisis de datos epidemiológicos a menudo incluye la consideración de otros factores de riesgo que ya se conocían para el resultado de salud. Si se omite este último paso, es probable que los resultados se distorsionen y se obtengan conclusiones incorrectas.

Términos para aprender
Regresión Logística
Factores de Riesgo
Confusión
Modificación de Efecto

Resumen: preguntas metodológicas importantes

- ¿Cuál es la pregunta de estudio?
- ¿Cómo deberían medirse las variables del estudio?
- ¿Cómo debería diseñarse el estudio?
- ¿Qué medidas de frecuencia de enfermedad deberían utilizarse?
- ¿Cuáles son los tipos de sesgos probables?
- ¿Cómo analizamos los datos del estudio?

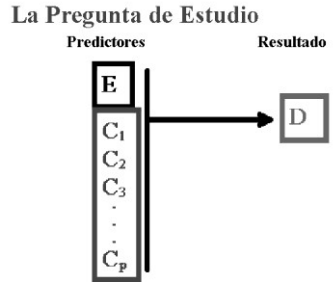
La pregunta de estudio

La epidemiología se ocupa principalmente de identificar los factores o variables trascendentales que influyen en un resultado de salud significativo. Por lo tanto, un primer paso importante en un estudio de investigación epidemiológica es plantear detenidamente la pregunta clave del estudio.

La pregunta de estudio debe ser clara y formularse lo antes posible, en especial, para indicar las variables que se observarán o medirán. Una pregunta típica de una investigación epidemiológica describe la relación entre una variable de resultado de salud, **D**, y una variable de exposición, **E**, teniendo en cuenta los efectos de otras variables que, según ya se sabe, pueden predecir el resultado (**C**, variables de control).

D = resultado de salud
E = variable de exposición
C = variable de control

Una situación simple, nuestro enfoque principal a lo largo del curso, se produce cuando hay una sola **D** y una sola **E**, con varias variables de control. Luego, la pregunta típica de investigación puede expresarse según se muestra a continuación, donde la flecha indica que las variables **E** y los controles **C** de la izquierda son las variables que se evaluarán como elementos predictores del resultado **D** de la derecha.



En el estudio de usuarios de las playas de Sydney, la variable de resultado de salud **D** que les interesaba a los investigadores era si una persona que se bañaba en una playa de Sydney desarrollaba o no una enfermedad infecciosa aguda, como, por ejemplo, tos, resfrío, gripe, infección de oído o infección ocular, en el plazo de una semana después de bañarse en la playa.

Los individuos del estudio podían clasificarse de la siguiente manera:

D=0 para los que no se enfermaron, o
D=l para los que se enfermaron.

Una opción lógica para la variable de exposición es: *condición del bañista*, la cual se clasifica en:

E=0 para los que no se bañaron, y
E=1 para los que se bañaron durante el período del estudio

(Adviértase que, si bien pueden utilizarse otros sistemas de codificación aparte de 0/1 tales como 1/2, S/N, o +/-, nosotros utilizaremos 0/1).

Las variables de control podrían incluir el nivel de contaminación en la playa, la edad del individuo y el tiempo que éste se bañó.

En términos generales, un estudio no es muy útil a menos que se pueda formular algún tipo de pregunta o hipótesis que justifique el tiempo y los gastos que conlleva el estudio.

Así, la pregunta de investigación de este estudio pretende describir la relación entre bañarse y el desarrollo de una enfermedad infecciosa, a la vez que tiene en cuenta los efectos de diversas variables de control relevantes tales como el nivel de contaminación, la edad del individuo y el tiempo que éste se bañó.

Como hay muchas variables involucradas, es muy probable que se requiera una serie de análisis complejos para estudiar todas las relaciones posibles entre dichas variables.

Resumen: la pregunta de estudio

- Un primer paso importante en un estudio de investigación epidemiológica es plantear detenidamente la pregunta clave del estudio.
- La pregunta general: ¿en qué medida existe una relación entre una o más variables de exposición **E** y un resultado de salud **D**, si tenemos en cuenta (es decir, controlamos) la posible influencia de otras covariables **C** importantes?
- Es muy probable que se requiera una serie de análisis complejos para estudiar todas las relaciones posibles entre las variables involucradas.



Auto-evaluación (P2.1)

En el estudio de usuarios de las playas de Sydney, la exposición se definió en forma alternativa distinguiendo las personas que se bañaron en aguas contaminadas de las que se bañaron en aguas no contaminadas y de las que

directamente no se bañaron. Sobre la base de esta situación potencial, introduzca la información faltante en la siguiente afirmación:

1. La variable de exposición tiene **???** categorías, una de las cuales es **???**

Opciones:
2 3 4 5 no se bañó agua contaminada se bañó agua no contaminada

2. Si se consideran juntas tanto la actividad de bañarse como la contaminación, ¿cuál de las siguientes opciones es adecuada para definir la variable de exposición en el estudio de usuarios de las playas de Sydney? **???**

Opciones:
a) E=0 si no se bañó, E=1 si se bañó en aguas contaminadas.
b) E=0 si no se bañó, E=1 si se bañó en aguas no contaminadas.
c) E=0 si no se bañó, E=1 si se bañó en aguas contaminadas, E=2 si se bañó en aguas no contaminadas.
d) E=0 si no se bañó, E=1 si se bañó.

En el estudio de usuarios de las playas de Sydney, el resultado de salud era si se desarrollaba o no una enfermedad infecciosa aguda al cabo de 1 semana de haberse bañado en la playa. Por otro lado, además de la edad, existía la variable de control de si un individuo del estudio se había bañado o no otros días aparte del día de la entrevista. Introduzca la información faltante:

3. El resultado de salud tiene **???** categorías.
4. Existen al menos **???** variables de control.
5. ¿Cuál de las siguientes opciones no es una variable de control? **???**
    a) Edad
    b) Condición de bañista en otros días
    c) Condición de bañista el día de la entrevista

Opciones:
2 3 4 5 a b c

2-2 Cuestiones metodológicas (continuación)

Medición de variables

Otra pregunta importante es: ¿cómo medimos las variables del estudio? En esta sección se presentan varios temas relacionados con la medición.

Después de establecer la pregunta de estudio, los investigadores deben determinar la forma de medir las variables identificadas para el estudio, así como cualquier otra información necesaria. Por ejemplo, ¿cómo se medirá la variable de exposición? Si un individuo se metió en el agua, pero jamás sumergió la cabeza, ¿se considera que se bañó? ¿Cuánto tiempo debe permanecer una persona en el agua para considerarse que se bañó? ¿Es factible observar la condición de bañista de cada individuo el día de la entrevista inicial? De lo contrario, ¿cómo se debe determinar la condición de bañista?

Después de considerar estas preguntas, el equipo del estudio definió la actividad de bañarse como *cualquier inmersión de la cara y la cabeza en el agua*. Se decidió que la única forma factible de obtener información sobre la condición de bañista de los individuos era considerar las declaraciones de cada individuo sobre su propia condición.

Medición de las variables de exposición

Definición de bañarse:

Cualquier inmersión de la cara y la cabeza en el agua

Medición del estado de bañarse:

Declaración de cada individuo

¿Cómo se debe medir el resultado de salud? ¿Debe la existencia de una enfermedad determinarse según la información proporcionada por los individuos (lo cual puede ser un método impreciso), o mediante la confirmación de un médico (el cual puede no estar disponible)? El equipo del estudio decidió utilizar información autodeclarada sobre síntomas mediante entrevistas telefónicas individuales al cabo de 7 a 10 días de la entrevista inicial.

Otra cuestión relacionada con la medición era cómo determinar la calidad del agua de la playa. ¿Es necesario tomar muestras de agua? ¿A qué hora del día se deben tomar dichas muestras? ¿Cómo se vinculará esta información con los individuos del estudio? El equipo del estudio decidió que los inspectores del departamento de salud tomaran muestras matutinas y vespertinas en un punto medio de cada uno de los tres sectores de las playas.

Los individuos del estudio se entrevistarían del modo más inmediato posible durante el período en que se tomaran las muestras de agua. Se estableció un protocolo estándar para la cantidad de agua que debía contener cada muestra, así como para el modo en que debía evaluarse la calidad del agua de las muestras.

Un último aspecto de medición se relacionaba con el tipo de información que debía recolectarse de las personas entrevistadas en las playas para su posible inclusión en el estudio. El equipo del estudio decidió recolectar datos demográficos básicos tales como edad, sexo y código postal, para preguntarle a cada encuestado si se había bañado o no durante los últimos 5 días y si presentaba alguna afección que le impidiera bañarse en la playa el día de la entrevista.

Variables de Entrevista

edad
sexo
código postal

estado de bañarse
estado de salud

Se excluyeron del estudio los individuos que informaron haberse bañado en el mar durante los últimos 5

días o que manifestaron padecer una enfermedad que les impidió bañarse en el mar. Se incluyeron en el estudio los individuos de, al menos, 15 años de edad que accedieron a una entrevista inicial en la playa y una entrevista telefónica de seguimiento.

Todas las cuestiones de medición anteriores deben resolverse antes de la recolección de datos para garantizar la obtención de información estandarizada y brindar un estudio eficiente en términos de costo y tiempo.

Preguntas de estudio (P2.2)

1. ¿Qué otras variables consideraría como variables de control en el Estudio de usuarios de las playas?
2. ¿Cómo decidimos las variables que se deben medir como variables de control?
3. ¿Por qué debería considerarse la edad?
4. ¿Cómo trataría a los individuos que fueron a la playa más de un día?

Resumen: medición de las variables

Preguntas generales de medición:

- ¿Cómo se operacionaliza la forma en que se realiza una medición?
- ¿Deben utilizarse los informes autodeclarados sobre la exposición y/o el resultado de salud?
- ¿Cuándo deben realizarse las mediciones?
- ¿Cuántas mediciones deben realizarse de cada variable y cómo deben combinarse las diversas mediciones?
- ¿Cómo se vinculan las mediciones ambientales con individuos en nuestro estudio?

El diseño de estudio, incluyendo el plan de muestreo

*Otra pregunta importante es: ¿qué **diseño de estudio** debe emplearse y cómo debemos seleccionar a los individuos del estudio? En esta sección se presentan varios temas relacionados con el diseño de estudio.*

En epidemiología se utiliza una gran variedad de diseños de estudio. En el estudio de usuarios de las playas de Sydney se utilizó un diseño de **cohorte**. Una característica clave de este tipo de diseños es que a los individuos que no presentan el resultado de salud se les hace un seguimiento durante un cierto tiempo para determinar si lo desarrollan. La selección de los individuos requirió el relevamiento de 12 playas concurridas de Sydney a lo largo de 41 días de muestreo. Se realizó en la playa una entrevista inicial con los individuos del estudio para confirmar su participación y obtener información demográfica.

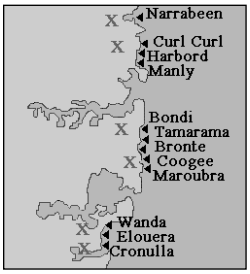
Quedaron excluidas del estudio las personas que no pudieron bañarse en el mar ese día a causa de una enfermedad y las que se habían bañado en el mar durante los últimos 5 días. Se llegó a la conclusión de que no era factible determinar la condición de bañista de cada individuo el día de la entrevista inicial. Por consiguiente, se realizó una entrevista telefónica de seguimiento entre los 7 y 10 días posteriores, a fin de obtener informes personales sobre la exposición por bañarse y la condición de enfermedad de cada individuo.

Preguntas de estudio (P2.3)

1. ¿Qué críticas merece la elección de usar informes autodeclarados sobre exposición y enfermedades?
2. ¿Qué críticas merece la decisión de determinar la condición de bañista mediante una entrevista telefónica realizada entre los 7 y 10 días posteriores a la entrevista en la playa?

Se empleó un diseño de encuestas por muestreo complejo para obtener los casi 3000 participantes del estudio. Se eligió un día cualquiera para relevar seis playas seleccionadas, las cuales incluían 2 de cada zona

(norte, este y sur). Cada playa se dividió en tres sectores, según la posición de las banderas colocadas por los guardavidas para delimitar las áreas aptas para bañarse. Los individuos fueron reclutados por entrevistadores capacitados, quienes comenzaron por el centro de cada sector y se desplazaron en sentido horario hasta alcanzar una cuota determinada para dicho sector. Los individuos potenciales debían guardar una distancia de, al menos, 3 metros entre sí.



Preguntas de estudio (P2.4)

1. ¿Por qué piensa que se especificó que los individuos potenciales en un sector determinado de la playa debían guardar, al menos, 3 metros de distancia entre sí?
2. ¿Por qué el estudio de usuarios de las playas de Sydney es un estudio de cohorte?
3. Una cohorte fija es un grupo de personas identificadas al inicio de un estudio a las que se les hace un seguimiento a lo largo del tiempo para determinar si desarrollaron el resultado. ¿En el estudio de usuarios de las playas de Sydney se utilizó una cohorte fija? Desarrolle su respuesta.
4. Un estudio de casos y controles comienza con individuos tanto enfermos como sanos y realiza un análisis retrospectivo para determinar los antecedentes de exposición de ambos grupos. ¿Por qué el estudio de usuarios de las playas de Sydney *no* es un estudio de casos y controles?
5. En un estudio transversal se observan tanto la condición de exposición como la de enfermedad, al mismo tiempo que se seleccionan individuos para el estudio. ¿Por qué el estudio de usuarios de las playas de Sydney no es un estudio transversal?

Resumen: diseño de estudio

- Dos preguntas generales de diseño de estudio:
  - ¿cuál de los posibles **diseños de estudio epidemiológico** debería utilizarse (por ejemplo, de cohorte, de casos y controles, transversal)?
  - ¿Cuál es el **plan de muestreo** para la selección de individuos?

Medidas de frecuencia de enfermedad y efecto

*Otra pregunta importante es: ¿qué **medida de frecuencia de enfermedad** y qué **medida de efecto** deben utilizarse? En esta sección se incluye una introducción breve de estos términos.*

Después de determinar el diseño de estudio, se pueden especificar las medidas de frecuencia de enfermedad y efecto adecuadas. Una medida de frecuencia de enfermedad aporta información cuantitativa sobre la periodicidad con que se ha producido el resultado de salud en un subgrupo de interés.

Por ejemplo, en el estudio de usuarios de las playas de Sydney, si queremos medir la frecuencia con la que los bañistas desarrollaron la enfermedad de interés, podemos determinar la cantidad de individuos que se enfermaron y se bañaron, y dividirla por la cantidad total de individuos que se bañaron. El denominador representa la cantidad total de individuos del estudio entre los bañistas que tuvieron la posibilidad de enfermarse. El numerador expresa la cantidad de individuos del estudio entre los bañistas que efectivamente se enfermaron.

Del mismo modo, si quisiéramos medir la frecuencia de la enfermedad entre las personas que *no* se bañaron, podríamos dividir la cantidad de individuos que se enfermaron y no se bañaron por la cantidad total de individuos que no se bañaron.

Medidas de frecuencia de enfermedad
Estudio de usuarios de playas de Sydney

Bañistas: $\frac{\#\ \text{enfermos bañistas}}{\#\ \text{total de bañistas}}$

No bañistas: $\frac{\#\ \text{enfermos no bañistas}}{\#\ \text{total de no bañistas}}$

La información necesaria para realizar los cálculos anteriores puede describirse en una tabla de doble entrada tal como se muestra a continuación. La tabla de doble entrada proporciona un resumen simple de la información requerida. En este caso, se muestra la cantidad de individuos que se enfermaron entre "bañistas" y "no bañistas". Podemos calcular que la proporción de enfermos entre los bañistas es de 0,277, o bien del 27,7%. También podemos calcular que la proporción de enfermos entre los no bañistas es de 0,165, o bien del 16,5%.

		Se bañó		
		Sí	No	Total
Se enfermó	Sí	532	151	683
	No	1392	764	2156
	Total	1924	915	2839

proporción enfermos (bañistas): $\frac{532}{1924} = .277$ or 27.7%

proporción enfermos (no bañistas): $\frac{151}{915} = .165$ or 16.5%

Cada proporción es una medida de frecuencia de enfermedad y se denomina **riesgo**. **R(E)** denota el riesgo que corren los individuos expuestos de desarrollar el resultado de salud. **R(no E)** [o **R($\bar{E}$)**] indica el riesgo entre los ***no*** expuestos. En este curso se describen otras medidas de frecuencia de enfermedad, aparte del riesgo. La elección de la medida (por ejemplo, riesgo, *odds*, prevalencia o tasa) depende principalmente del tipo de diseño de estudio que se utilice y el objetivo del estudio de investigación.

Si queremos comparar dos riesgos u otras medidas de frecuencia de enfermedad, podemos dividir los dos valores entre sí; por ejemplo, podemos dividir el riesgo de los bañistas por el riesgo de los no bañistas. En nuestro estudio, el coeficiente de estos riesgos es 1,68, lo cual significa que el riesgo de los bañistas de desarrollar la enfermedad es 1,68 veces mayor que el de los no bañistas.

Riesgo

proporción enfermos (bañistas): 27.7%

proporción enfermos (no bañistas): 16.5%

$$\frac{R(E)}{R(\text{no } E)} = \frac{27.7\%}{16.5\%} = 1.68$$

$$\text{Riesgo}_{(\text{bañistas})} = 1.68 \times \text{Riesgo}_{(\text{no bañistas})}$$

Dicha medida se denomina **medida de efecto**. En este ejemplo, el efecto de interés se refiere al efecto que produce la condición de bañista de una persona en su estatus de salud. Si dividimos un riesgo por el otro, la medida de efecto o la asociación se denomina **razón de riesgos**. En este curso se describen otras medidas de efecto (por ejemplo, razón de riesgos, *odds ratio*, razón de prevalencias, razón de tasas, diferencia de riesgos y diferencia de tasas). Al igual que en el caso de las medidas de frecuencia de enfermedad, la elección de la medida de efecto depende del tipo de diseño de estudio y del objetivo del estudio de investigación.

Resumen: medidas de la frecuencia de la enfermedad y el efecto

- Una **medida de frecuencia de enfermedad** cuantifica la periodicidad con que se ha producido el resultado de salud en un subgrupo de interés.
- Una **medida de efecto** cuantifica una comparación entre medidas de frecuencia de enfermedad para dos o más subgrupos.
- La elección de la medida de frecuencia de enfermedad y la medida de efecto depende del tipo de diseño de estudio utilizado y del objetivo del estudio de investigación.

Sesgo

Otra pregunta importante es: ¿cuáles son los sesgos potenciales del estudio? En esta sección se incluye una introducción breve del concepto de sesgo.

La siguiente cuestión metodológica comprende los sesgos potenciales de un estudio. Un sesgo es una falla en el diseño de estudio, los métodos de recolección de datos o los métodos de análisis de datos que puede derivar en conclusiones falsas sobre la relación entre la exposición y la enfermedad. Un sesgo puede producirse como consecuencia de: la **selección** de los individuos del estudio; la recopilación de información incorrecta sobre los individuos del estudio; o la omisión de ajustes por variables que no sea la variable de exposición, lo cual en la epidemiología suele denominarse **confusión**.

Sesgo

Una falla en

1. el diseño del estudio,
2. los métodos de recolección de datos
3. los métodos de análisis de datos

que puede derivar en conclusiones falsas

Fuentes de sesgo

1. Selección
2. Información
3. Confusión

En el estudio de usuarios de las playas de Sydney se consideraron las 3 fuentes de sesgos. Por ejemplo, para evitar el **sesgo de selección**, se excluyeron del análisis a los individuos que ya estaban enfermos el día de la entrevista. Esto garantizó que la muestra representara sólo a aquéllos que eran lo suficientemente saludables como para bañarse el día de la entrevista. En ocasiones, el sesgo de selección es inevitable. Por ejemplo, si un individuo no completaba la entrevista de seguimiento, se le excluía del estudio. Este **sesgo de falta de respuesta** puede hacer que la muestra sea menos representativa.

También existía la posibilidad de que se produjera un **sesgo de información**, puesto que tanto la condición de bañista como la de enfermedad se basaban en informes autodeclarados de los individuos del estudio. La condición de bañista se determinó mediante informes autodeclarados de los individuos, al menos, siete días después de que se bañaron. Asimismo, el registro del resultado de salud no comprendía ninguna confirmación clínica de los síntomas informados.

Con el propósito de evaluar el **sesgo de confusión** en la fase análitica del estudio de usuarios de las playas

de Sydney, se tuvo en cuenta la medición de todas las variables relevantes, además de la condición de bañista y la exposición a niveles de contaminación del agua de mar. Estas variables incluían la edad; el sexo; la duración de la exposición al agua de mar, en el caso de los que se bañaron en la playa; y si el individuo se bañó o no en la playa otros días después de la entrevista inicial en la playa. La razón principal por la que se midieron estas variables fue garantizar que ninguno de los resultados de salud observados potencialmente asociados con el hecho de bañarse en el mar pudiera explicarse por medio de estas otras variables.

Resumen

- Un sesgo puede resultar en conclusiones falsas sobre la relación entre la exposición y la enfermedad debido a una falla en: (1) el diseño de estudio; (2) los métodos de recolección de datos; o (3) la metodología analítica de datos que puede derivar en conclusiones falsas sobre.
- Hay tres fuentes generales de sesgos que se producen en:
  - la selección de los individuos del estudio;
  - la recopilación de información incorrecta sobre los individuos del estudio; y
  - la omisión de ajustes por variables que no sean la variable de exposición (confusión).

2-3 Cuestiones metodológicas (continuación)

Análisis de datos

Otra pregunta importante es: ¿cómo realizamos el análisis de datos? En esta sección se incluye una introducción breve de algunas ideas básicas sobre el análisis de datos.

La última cuestión metodológica comprende el análisis de datos. Debemos realizar un análisis adecuado una vez que se hayan recabado y procesado los datos del estudio. Puesto que los datos suelen provenir de una muestra de individuos, por lo general, el análisis de datos requiere el uso de procedimientos estadísticos para dar cuenta de la inherente variabilidad de la información. En epidemiología, el análisis de datos suele comenzar por el cálculo y la evaluación estadística de medidas simples de frecuencia de enfermedad y efecto. El análisis a menudo avanza de tal manera que requiere técnicas más sofisticadas tales como la estratificación y la modelación matemática. Estos últimos métodos con frecuencia se utilizan para controlar uno o más factores de confusión potenciales.

Estadísticas

Frecuencia:	Efecto:
riesgo	razón de riesgo
proporción	odds ratio
tasa	razón de prevalencia

Estratificación

Modelación matemática

Tomemos el análisis de datos en el estudio de usuarios de las playas de Sydney. Antes comparamos los bañistas con los no bañistas. Ahora, podríamos abordar la pregunta más específica de si las personas que se bañaron en aguas contaminadas corrían un riesgo mayor de desarrollar la enfermedad que los que se bañaron en aguas no contaminadas. Esto puede hacerse separando a los bañistas en dos grupos. Los no bañistas representan un grupo de referencia en función del cual se pueden comparar los dos grupos de bañistas.

Si nos basamos en la tabla de doble entrada, podemos calcular el riesgo de enfermedad de cada uno de los tres grupos obteniendo la proporción de los que se enfermaron con respecto al total de cada grupo. Los tres valores estimados para el riesgo son 0,357, 0,269 y 0,165; lo cual se traduce en 35,7%, 26,9% y 16,5%, respectivamente.

Estudio de usuarios de playas de Sydney

		Se bañó			
		Sí-C	Sí-NC	No	Total
Se enfermó	Sí	55	477	151	683
	No	99	1293	764	2156
	Total	154	1770	915	2839
riesgo de enfermedad:		35.7%	26.9%	16.5%	

La razón de riesgos que compara el grupo de los que se bañaron en aguas contaminadas (Sí-C) con el grupo de los que se bañaron en aguas no contaminadas (Sí-NC) es igual a 1,33; lo cual indica que el primer grupo corría un riesgo 33% mayor que el segundo grupo.

razón de riesgo: (C vs NC) $\frac{35.7\%}{26.9\%} = 1.33$

Asimismo, las razones de riesgos obtenidas al dividir los riesgos de cada grupo por el riesgo de los no bañistas son 2,16; 1,63 y 1. Esto sugiere lo que denominamos un efecto dosis-respuesta; es decir, cuanto mayor la exposición, mayor el riesgo.

razón de riesgo: $\frac{35.7\%}{16.5\%}$ $\frac{26.9\%}{16.5\%}$ $\frac{16.5\%}{16.5\%}$

2.16 **1.63** **1.00** (referencia)

Efecto dosis-respuesta

El análisis anterior se denomina análisis "crudo", o análisis bruto, porque no tiene en cuenta los efectos de otros factores conocidos que también podrían incidir en el resultado de salud que se está estudiando. Una lista de este tipo de variables podría incluir la edad del individuo, el tiempo durante el cual se bañó en el mar y si también se bañó en el mar durante otros días. Las conclusiones obtenidas a partir de un análisis crudo podrían sufrir modificaciones drásticas después de hacer los ajustes necesarios para contemplar estas variables potenciales de confusión.

Cuando se considera el control de muchas variables, surgen varias preguntas:

- ¿Cuáles de las variables que se están considerando deben controlarse?
- ¿Qué se gana o se pierde cuando las variables que se controlan son demasiadas o muy pocas?
- ¿Qué debemos hacer si tenemos tantas variables para controlar en un análisis estratificado o modelo matemático que nos quedamos sin observaciones suficientes en nuestra muestra?
- ¿Qué es lo que realmente implica realizar un análisis estratificado o una modelación matemática para controlar diversas variables?
- ¿Qué puntos de comparación tienen los diferentes métodos de control tales como la estratificación y la modelación matemática?

Estas preguntas se abordarán en actividades posteriores.

Preguntas de estudio (P2.5)

1. ¿Cómo se interpreta una razón de riesgos de 1,33 (en términos generales, antes de evaluar significancia estadística, por supuesto)?
2. ¿La razón de riesgos 1,33 indica que bañarse en aguas contaminadas representa un riesgo para la salud?
3. Puesto que las 154 personas que se bañaron en aguas contaminadas constituyen una cantidad relativamente pequeña, ¿qué pregunta estadística se debe responder sobre la importancia de la razón de riesgos obtenida cuando ésta es de 1,33?

Resumen: análisis de datos

- Por lo general, el análisis de datos requiere el uso de procedimientos estadísticos para tomar en cuenta la inherente variabilidad de la información.
- En epidemiología, el análisis de datos suele comenzar por la evaluación y comparación de medidas simples de la frecuencia de enfermedad y efecto.
- El análisis a menudo avanza de tal manera que requiere técnicas más sofisticadas tales como la estratificación y la modelación matemática.

Consumo de alcohol y cáncer de mama en el estudio de salud de enfermeras

La Facultad de Salud Pública de Harvard estudió a una cohorte de alrededor de 100.000 enfermeras de diferentes partes de los Estados Unidos durante la década de los ochenta y principios de la década de los noventa. Los investigadores que participaron en el estudio de salud de enfermeras estaban interesados en evaluar la posible relación entre la dieta y el cáncer. Una pregunta en particular era en qué medida el consumo de alcohol se relacionaba con el desarrollo de cáncer de mama.

A las enfermeras identificadas como "sin enfermedades" al momento de inscribirse en el estudio se les preguntó por la cantidad de alcohol que bebían en forma habitual. También se determinaron otros factores relevantes, incluso la edad y antecedentes de tabaquismo. El seguimiento de enfermeras duró cuatro años, momento en el que se determinó quiénes habían desarrollado cáncer de mama y quiénes no. En la publicación *New England Journal of Medicine* de 1987, se incluyó un informe de estos resultados.

Recuerde que la primera cuestión metodológica consiste en definir la **pregunta de estudio**. ¿Cuál de las siguientes preguntas de estudio aborda mejor la pregunta de interés en este estudio?

A. ¿Existe alguna relación entre el consumo de alcohol y el desarrollo del cáncer de mama?
B. ¿Están relacionados el consumo de alcohol, la edad y el tabaquismo con el desarrollo del cáncer de mama?
C. ¿Están relacionados la edad y el tabaquismo con el desarrollo del cáncer de mama, después de controlar el consumo de alcohol?
D. ¿Está relacionado el consumo de alcohol con el desarrollo del cáncer de mama, después de tomar en cuenta otras variables relacionadas con el desarrollo del cáncer de mama?

La mejor respuesta es la D: "¿está relacionado el consumo de alcohol con el desarrollo del cáncer de mama, después de tomar en cuenta otras variables relacionadas con el desarrollo del cáncer de mama?". No obstante, también es correcta la respuesta A: "¿existe alguna relación entre el consumo de alcohol y el desarrollo del cáncer de mama?".

Al formular la pregunta de estudio de interés, debemos identificar las variables principales que se medirán.

Preguntas de estudio (P2.6)

Determine si cada uno de los siguientes factores es:
una variable de resultado de salud (D);
una variable de exposición (E); o
una variable de control (C).

1. Antecedentes de tabaquismo
2. El hecho de que un individuo desarrolle o no cáncer de mama durante el seguimiento
3. Alguna medida del consumo de alcohol
4. Edad

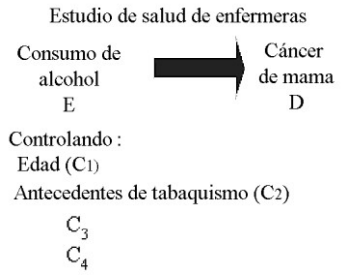
Después de especificar las variables apropiadas para el estudio, debemos determinar cómo medirlas. La variable de resultado de salud en este ejemplo, **D**, es simplemente *sí* o *no*, según el hecho de que a una persona se le haya diagnosticado cáncer de mama. Los investigadores de Harvard entrevistaron a las enfermeras del estudio acerca de su consumo habitual de alcohol, **E**, y plantearon una medición cuantitativa de la cantidad de alcohol, expresada en gramos por día, que se consumía en una semana promedio cercana al momento de inscripción en el estudio. Una pregunta importante que se consideró fue cómo tratar esta variable para propósitos del análisis de los datos del estudio. Un enfoque consistía en categorizar la medición del consumo de alcohol en "excesivo" y "moderado". Otro enfoque proponía categorizar el consumo de alcohol en 4 grupos: abstemio; menos de 5 gramos por día; entre 5 y 15 gramos por día; y 15 gramos o más por día.

La edad, denotada como $\mathbf{C_1}$, es una variable inherentemente cuantitativa, si bien muchos de los análisis la trataron como una variable categórica para formar tres grupos etarios tal como se muestra a continuación:

34 a 44 años
45 a 54 años
55 a 59 años

La variable antecedentes de tabaquismo, $\mathbf{C_2}$, se categorizó de varias maneras; una fue *nunca* fumó, en contraposición a fumó *alguna vez*.

Así, la pregunta de investigación en el estudio de salud de enfermeras puede describirse como un intento para determinar si existe alguna relación entre el consumo de alcohol, **E**, y el cáncer de mama, **D**, controlando los efectos de la edad, $\mathbf{C_1}$, los antecedentes de tabaquismo, $\mathbf{C_2}$, y otras variables posibles (C_3, C_4, etc.).



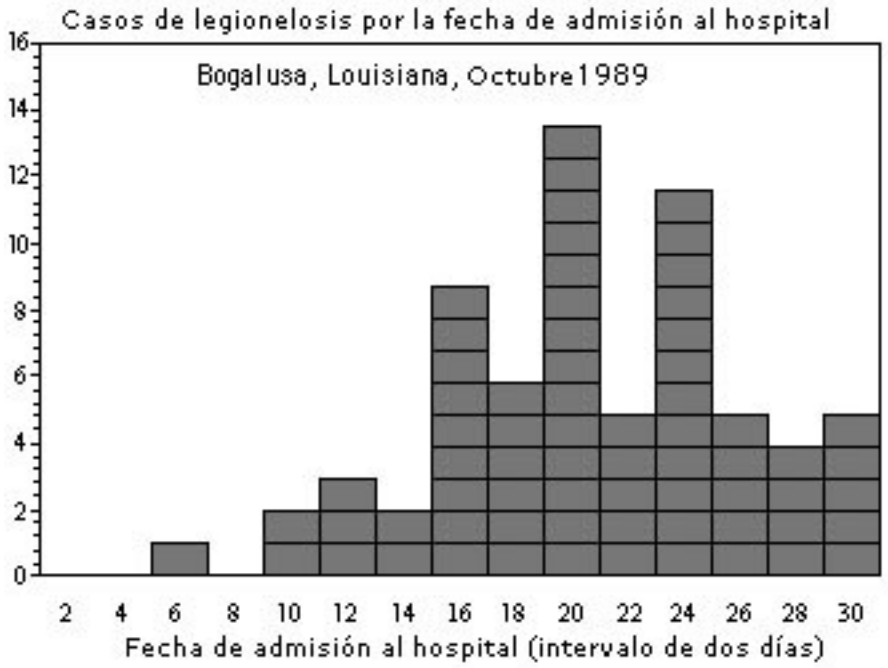
Si bien aquí no se describe un análisis detallado, los datos proporcionaron evidencia de una relación entre el consumo de alcohol y el desarrollo de cáncer de mama. En comparación con las personas abstemias, las mujeres con mayor consumo de alcohol corrían un riesgo 80% superior de desarrollar cáncer de mama. El riesgo de las bebedoras moderadas era un 50% mayor y el de las bebedoras ocasionales, un 20% mayor.

	Comparación con las personas abstemias
Bebedoras excesivas	80% riesgo superior
Bebedoras moderadas	50% riesgo superior
Bebedoras ocasionales	20% riesgo superior

Nota: El estudio de salud de enfermeras muestra un ejemplo en el que la variable de exposición (el consumo de alcohol) puede dividirse en varias categorías en lugar de ser sólo binaria. Además, las variables de control "edad" y "antecedentes de tabaquismo" pueden ser una combinación de diferentes tipos de variables. En el estudio de salud de enfermeras, la edad se divide en tres categorías y los antecedentes de tabaquismo se tratan como una variable binaria.

El brote de Bogalusa

El 31 de octubre de 1989, dos médicos en Bogalusa, Louisiana, informaron al Departamento de Salud del Estado de Louisiana que desde mediados hasta fines de ese mes, en el plazo de tres semanas, se habían detectado más de 50 casos graves de neumonía y que seis personas habían muerto. La información que los médicos habían obtenido de varios pacientes sugería que la enfermedad podría haber sido la enfermedad del legionario, o legionelosis.



En 1989, Bogalusa era una ciudad de unos 16.000 habitantes. El empleador más grande era una papelera ubicada en el centro de la ciudad junto a la calle principal. La papelera incluía cinco torres de enfriamiento prominentes. La fábrica también tenía tres máquinas de papel que emanaban grandes volúmenes de aerosol a lo largo de la calle principal de la ciudad. Muchas personas sospechaban que las torres de enfriamiento y/o la papelera eran la causa del brote, puesto que eran fuentes importantes de aerosoles atmosféricos donde podía haberse alojado la bacteria que causa la enfermedad del legionario.

Recuerde que la primera cuestión metodológica consiste en definir la **pregunta de estudio** de interés. ¿Cuál de las siguientes preguntas de estudio aborda mejor la pregunta de interés de este estudio?

A. ¿Fue la papelera el origen del brote de la legionelosis en Bogalusa?
B. ¿Cuál fue el origen del brote de legionelosis en Bogalusa?
C. ¿Por qué fue la papelera la causa del brote de legionelosis en Bogalusa?
D. ¿Hubo un brote de legionelosis en Bogalusa?

La pregunta de estudio más apropiada es la B: "¿cuál fue el origen del brote de legionelosis en Bogalusa?". Si bien se sospechaba que el origen era la papelera, el estudio no se limitó sólo a esa variable. De lo contrario, podría haber sido incapaz de recabar información sobre el verdadero origen del brote.

Al formular la pregunta de estudio, identificamos las variables principales que se considerarán en el estudio.

Preguntas de estudio (P2.7)

Determine si cada una de las siguientes variables es la variable de resultado de salud, **D**, una variable de exposición, **E**, o una variable de control, **C**:

1. exposición a las torres de enfriamiento de la papelera
2. exposición a las emanaciones de las máquinas de papel
3. edad del individuo
4. visita al almacén A
5. visita al almacén B
6. diagnóstico de legionelosis
7. visita a la farmacia A
8. visita a la farmacia B
9. consumo de alimentos en el restaurante A

La variable de resultado de salud, **D**, indica si a un individuo del estudio se le diagnosticó o no legionelosis durante el período de tres semanas desde mediados hasta fines de octubre. La variable de exposición es, desde el punto de vista conceptual, cualquier variable que indique el origen principal del brote. Puesto que, en esencia, esta variable se desconoce al inicio del estudio, existe una amplia colección de variables de exposición que deben identificarse como parte del diseño de estudio e investigarse como candidatas para convertirse en el origen principal del brote. Estas variables de exposición de interés reciben aquí las denotaciones $\mathbf{E_1}$ a $\mathbf{E_7}$. Una posible variable de control de interés fue la edad, la cual recibe aquí la indicación $\mathbf{C_1}$.

Por consiguiente, la pregunta de investigación general de interés en el brote de Bogalusa puede describirse como un intento de evaluar la relación entre una o más variables de exposición y el hecho de que un individuo del estudio desarrolle o no la enfermedad del legionario, controlando la edad.

Se realizó un **estudio de casos y controles** en el que se estableció una comparación entre 28 **casos** con diagnóstico confirmado de enfermedad del legionario y 56 no casos o **controles**. De esta investigación surgió la hipótesis de que el origen del brote era una máquina atomizadora para verduras de un almacén. Se retiró la máquina del almacén y se le envió al Centro de Control de Enfermedades, donde el personal del laboratorio fue capaz de aislar los organismos *legionella* de los aerosoles que producía la máquina. Antes de esto, se desconocía que dicho origen pudiese ser un vehículo para la transmisión de la bacteria *legionella*.

Nota: El estudio de Bogalusa muestra un ejemplo en el que existen diversas variables de exposición con posibilidades de ser el origen principal del resultado de salud bajo estudio. Se espera que los investigadores sean capaces de identificar al menos una variable de exposición implicada en la aparición del brote. Incluso es posible que se identifique más de una variable de exposición como origen potencial.

El estudio de casos y controles de este y muchos brotes a menudo puede considerarse una generación de hipótesis. Con frecuencia se requieren otros estudios, que a menudo comprenden métodos de laboratorio, así como diagnósticos clínicos y técnicas de muestreo ambiental, para confirmar una presunta exposición como el origen principal del brote. Los Centros para el Control y la Prevención de Enfermedades cuentan con una gran variedad de científicos que satisfacen los diferentes requerimientos de experiencia, conocimiento y trabajo en equipo que se presentan, como en el caso del estudio de Bogalusa.

El estudio de Rotterdam

El estudio de Rotterdam ha investigado los determinantes de diversas enfermedades discapacitantes crónicas, incluso el mal de Alzheimer, desde la década de los noventa.

A principios de los años noventa, el Departamento de Epidemiología de la Universidad de Rotterdam, en

los Países Bajos, comenzó a desarrollar el estudio de Rotterdam. Se seleccionó una cohorte de unas 8000 personas de edad avanzada. Su seguimiento se ha prolongado hasta el día de hoy. El objetivo del estudio consiste en investigar los determinantes de diversas enfermedades discapacitantes crónicas tales como el mal de Alzheimer y las enfermedades cardiovasculares. Una de las preguntas de estudio de interés era si el tabaquismo incrementa el riesgo de desarrollar el mal de Alzheimer.

En el estudio se incluyeron individuos que no presentaban signos de demencia en el examen inicial. Mediante este examen, se excluía a toda persona con diagnóstico de Alzheimer o cualquier otra forma de demencia a causa de factores orgánicos o psicológicos. Al cabo de unos dos años, se les pidió a los participantes que rindieran un breve examen cognitivo. Si obtenían una calificación positiva, se sometían a una revisión del neurólogo. Los investigadores luego podían determinar si un participante había desarrollado o no el mal de Alzheimer, la variable de resultado de salud **D** de interés, desde el inicio del seguimiento.

Los antecedentes de tabaquismo eran la principal variable de exposición, **E**. Se consideraron tres categorías de tabaquismo: fumadores al momento de la entrevista, ex fumadores, y personas que nunca habían fumado (no fumadores). Las variables de control que se consideraron en este estudio incluían: edad, sexo, educación y consumo de alcohol.

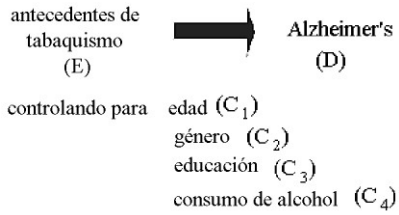
Estudio de Rotterdam

Individuos del estudio:

- no presentan signos de demencia en el examen inicial
- examen cognitivo - 2 años más tarde
- revisión del neurólogo (si examen +)
- resultado de salud: Alzheimer's (D)
- variable de exposición: antecedentes de tabaquismo (E)

3 categorías
fumadores al momento, ex fumadores, y no fumadores

La pregunta de estudio de interés se define de la siguiente manera: *¿existe alguna relación entre los antecedentes de tabaquismo y el mal de Alzheimer, controlando los efectos de la edad, sexo, educación y consumo de alcohol?*



Recuerde que una de las cuestiones metodológicas más importantes consiste en determinar el diseño de estudio.

¿Cómo definiría el diseño de este estudio?

1. Diseño de cohorte
2. Diseño de casos y controles
3. Diseño transversal
4. Ensayo clínico

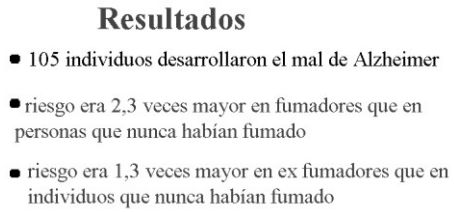
Este diseño es de cohorte porque a los participantes que no presentan el resultado de salud de interés (en este caso, el mal de Alzheimer) se les hace un seguimiento a lo largo del tiempo para ver si desarrollan el resultado con posterioridad.

¿En cuál de los siguientes factores influye el diseño de estudio?

A. La evaluación del sesgo de confusión.
B. La elección de medidas de frecuencia de enfermedad y efecto.
C. Una decisión con respecto al uso de un análisis estratificado.
D. El diseño de estudio empleado no influye en modo alguno en el análisis.

La respuesta es la B. Determinamos las medidas apropiadas para la frecuencia de enfermedad y efecto en función de las características del diseño de estudio. Las opciones A y C son incorrectas, porque se suele considerarlas independientemente del diseño de estudio empleado.

Los investigadores descubrieron que 105 individuos desarrollaron el mal de Alzheimer. Después de tener en cuenta las variables de control, se determinó que el riesgo de desarrollar el mal de Alzheimer era 2,3 veces mayor en fumadores que en personas que nunca habían fumado. En el caso de los individuos que habían fumado en el pasado, pero que dejaron de hacerlo antes del inicio del estudio, el riesgo de desarrollar el mal de Alzheimer era 1,3 veces mayor que el de los individuos que nunca habían fumado.



Preguntas de estudio (P2.8)

Si tenemos en cuenta los resultados anteriores:

1. ¿cuál es el *incremento porcentual* en el riesgo de los fumadores en comparación con el riesgo de los que nunca fumaron?
2. ¿Cuál es el *incremento porcentual* en el riesgo de los ex fumadores en comparación con el riesgo de los que nunca fumaron?

Puesto que estos resultados eran importantes desde el punto de vista estadístico y se habían controlado los elementos predictores de Alzheimer previamente establecidos, el estudio apoyó la hipótesis de que los antecedentes de tabaquismo constituían un factor de riesgo significativo en el desarrollo del mal de Alzheimer.

Nomenclatura

C	Variable de control o covariable
D	Enfermedad o variable de resultado
E	Variable de exposición
R(E)	Riesgo de los expuestos de desarrollar el resultado de salud
R(no E) o **R(Ē)**	Riesgo de los ***no*** expuestos de desarrollar el resultado de salud
RR	Razón de riesgos

Referencias

Para el estudio de usuarios de las playas de Sydney:
Corbett SJ, Rubin GL, Curry GK, Kleinbaum DG. *The health effects of swimming at Sydney Beaches*. The Sydney Beach. Users Study Advisory Group. Am J Public Health. 1993;83(12): 1701-6.

Para el estudio de salud de enfermeras:

Willett WC, Stampfer MJ, Colditz GA, Rosner BA, Hennekens CH, Speizer FE. *Moderate alcohol consumption and the risk of breast cancer*. N Engl J Med. 1987;316(19):1174-80.

Para el brote de Bogalusa:

Mahoney FJ, Hoge CW, Farley TA, Barbaree JM, Breiman RF, Benson RF, McFarland LM. *Communitywide outbreak of Legionnaires' disease associated with a grocery store mist machine*. J Infect Dis. 1992;165(4):736-9.

Para el estudio de Rotterdam:

Hofman A, Grobbee DE, de Jong PT, van den Ouweland FA. *Determinants of disease and disability in the elderly: the Rotterdam Elderly Study*. Eur J Epidemiol. 1991;7(4); 403-22.

Ott A, Slooter AJ, Hofman A, van Harksamp F, Witteman JC, Van Broeckhoven C, van Duijin CM, Breteler MM. *Smoking and risk of dementia and Alzheimer's disease in a population-based cohort study: the Rotterdam Study*. Lancet. 1998;351(9119):1840-3.

Ejercicios

EPI-1. ¿Qué es la epidemiología? ¿Cuál es el origen de la palabra "epidemiología" (y por qué no tiene nada que ver con el estudio de la piel)?

EPI-2. Causalidad. Para cada uno de los fragmentos que se incluyen a continuación, indique cuál(es) de los criterios de causalidad (propuestos por A. B. Hill, cerca de 1964) se está(n) analizando (puede elegir más de uno). Adviértase que estos criterios se presentan en la lección 3, página 38:

A. Grado de asociación
B. Consistencia
C. Temporalidad
D. Dosis-respuesta, o gradiente biológico
E. Plausibilidad biológica
F. Especificidad
G. Coherencia
H. Experimento
I. Analogía

1. [Extraído de un estudio que analiza si los hispanos* tienen mayores probabilidades de experimentar una discapacidad que los blancos] "Según la encuesta de salud realizada entre 1978 y 1980, los participantes mexicano-estadounidenses tenían más probabilidades de registrar limitaciones en su actividad que los blancos no hispanos. Los datos de la Encuesta nacional sobre gastos médicos de 1987 sugirieron lo contrario, ya que los hispanos presentaron menos limitaciones funcionales que los blancos no hispanos. Haan y Weldon brindaron datos que sugerían que la discapacidad en hispanos puede ser más evidente en personas con, al menos, dos de las siguientes enfermedades crónicas: diabetes, accidentes cerebrovasculares e hipertensión".

2. [Extraído del estudio anterior] "De las personas que viven en comunidades, la probabilidad de que los hispanos necesitaran asistencia con actividades instrumentales de la vida diaria (IADL, por sus siglas en inglés) era de 2 a 5 veces mayor que en el caso de los blancos no hispanos. No obstante, había una mayor proporción de blancos no hispanos discapacitados en hogares de ancianos y las estimaciones que incluían

* La nomenclatura por clasificación étnica utilizada en los Estados Unidos denomina *hispanos* a personas de descendencia Latina, independientemente de su raza o color. Sin embargo, la clasificación de otros grupos es basada en raza (e.g., personas de descendencia caucasiana son denominados *blancos*, etc.)

residentes de hogares de ancianos sugerían un excedente hispano más modesto que, por lo general, era de menos del doble".

3. [Extraído de un estudio sobre la exposición paterna a rayos x previa a la concepción y su impacto sobre los nacimientos] "La variable de exposición se generó a partir de un punto del cuestionario para padres en el que se preguntaba si el individuo se había sometido a ciertos estudios médicos con rayos x en algún momento durante los 12 meses anteriores a la concepción".

4. "El incremento pronunciado en el riesgo de desarrollar preeclampsia de las mujeres con diabetes tipo I concuerda con el de los informes anteriores y puede deberse a cambios microvasculares que afectan la perfusión placentaria. Nuestros resultados de que la diabetes tipo I guarda una relación significativa, aunque menos fuerte, con la hipertensión gestacional puede reflejar una vía metabólica común en la patogénesis de la preeclampsia y la hipertensión gestacional".

5. [Extraído de un estudio de factores predictores de enfermedad de la vesícula biliar en hombres] "Los mayores niveles de IMC (índice de masa corporal) demostraron una asociación progresiva con un incremento en el riesgo de desarrollar la enfermedad; los hombres con un IMC > = 24,0 unidades presentaron un riesgo del 46% mayor que aquéllos con un IMC < 20,0".

6. "Estudios recientes en los Estados Unidos y China han descrito una asociación entre el consumo de cigarrillos y los cánceres de cavidad nasal y senos paranasales en humanos. A la fecha, la limitada evidencia de dos estudios realizados en Japón sugiere que la exposición al humo de tabaco ambiental también es un factor de riesgo para el desarrollo del cáncer de senos nasales. El estudio que se menciona aquí se diseñó para probar la hipótesis de que la exposición al humo de tabaco ambiental en el hogar incrementa el riesgo de desarrollar cáncer de cavidad nasal y senos paranasales en perros domésticos. El riesgo de cáncer nasal también se examinó según el tipo histológico. Los perros con sarcomas presentaron un mayor riesgo ajustado que los perros con carcinomas para el tercil más alto del índice de exposición".

7. "Diversos estudios han descubierto una relación entre un mayor consumo de café y un menor riesgo de desarrollar cáncer de colon, si bien este resultado no es universal. La composición del café es bastante compleja y sus diversos componentes tienen propiedades potencialmente genotóxicas, mutagénicas y antimutagénicas. Además, el café modula diversos procesos fisiológicos tales como la motilidad del intestino grueso, lo cual podría alterar la exposición del colon a posibles agentes cancerígenos de la materia fecal".

8. [Extraído de un estudio sobre el café y el cáncer colorrectal] "Otra explicación posible de los resultados es que los individuos que tienen un alto riesgo de desarrollar cáncer colorrectal o que presentan síntomas no diagnosticados de cáncer de intestino grueso evitan el consumo de café. Rosenberg et al. obtuvieron resultados similares según se analizara el consumo de café del último año o de los 3 últimos años".

9. "Los estudios epidemiológicos observacionales sobre la ingesta de calcio y las fracturas son dispares. Se han realizado, al menos, 14 estudios de fractura de cadera e ingesta de calcio y sólo tres de ellos encontraron un efecto claramente protector. Por otro lado, dos pequeños ensayos aleatorizados han encontrado un porcentaje reducido de fracturas vertebrales en radiografías de individuos que recibieron suplementos de calcio, y otro estudio pequeño encontró una reducción insignificante en el riesgo de sufrir fracturas vertebrales y no vertebrales sintomáticas. Según un importante ensayo francés, la combinación de suplementos de calcio y vitamina D redujo a la mitad el porcentaje de fracturas de cadera en mujeres que vivían en hogares de ancianos".

EPI-3. Asociación causal entre la exposición y la enfermedad. ¿En qué circunstancias la asociación entre la exposición y la enfermedad podría ser causal sin ser biológicamente plausible?

EPI-4. Sitio web del CDC. El Centro para el Control y la Prevención de Enfermedades (CDC, por sus siglas en inglés) tiene un sitio web denominado **EXCITE**, que en inglés significa **Excelencia en la integración curricular a través de la enseñanza de epidemiología**. La dirección del sitio web es:

http://www.cdc.gov/excite/

Visite este sitio web y explore las diversas funciones y metas del sitio que se describen en la primera página. Luego, haga clic en el vínculo **Disease Detectives at Work** (detectives médicos en acción, en el menú de la izquierda) y lea los dos primeros artículos: *Public Health on Front Burner After Sept 11* y *USA's 'Disease Detectives' Track Epidemics Worldwide*. A continuación, haga clic en el vínculo **Classroom Exercises** (ejercicios para el salón de clases, en el menú de la izquierda) y haga el ejercicio sobre legionelosis en Bogalusa, Louisiana. La dirección específica de este ejercicio en el sitio web es:

http://www.cdc.gov/excite/legionnaires.htm

Respuestas a las preguntas de estudio y auto-evaluaciones

P2.1

1. 3, no se bañó
2. C
3. 2
4. 2
5. C

P2.2

1. Condición de salud general, condición de tabaquismo y dieta, incluso lo que un individuo podría haber comido en la playa.
2. Se seleccionan las variables que ya son determinantes conocidos del resultado de salud. Esto se analizará más adelante en la sección sobre **sesgo de confusión**.
3. Los individuos más jóvenes podrían tener menos probabilidades de enfermarse que los de más edad.
4. En el estudio en cuestión, los investigadores optaron por excluir del análisis a los sujetos que habían vuelto a la playa después del día de la entrevista inicial.

P2.3

1. Los datos de los informes personales podrían ser imprecisos y, por consiguiente, producir resultados falsos en el estudio.
2. Al igual que en la pregunta anterior, la información sobre la exposición se obtuvo mucho después del momento de la exposición, lo cual podría dar lugar a imprecisiones y resultados falsos en el estudio.

P2.4

1. Para minimizar la inclusión de familias o grupos sociales en el estudio.
2. Porque a los individuos que no presentaban el resultado de salud; es decir, los individuos saludables que se seleccionaron en la playa, se les hizo un seguimiento a lo largo del tiempo para determinar si habían desarrollado el resultado.
3. No, en el estudio de usuarios de las playas de Sydney no se utilizó una cohorte fija. Se fueron agregando individuos al estudio en forma progresiva durante el verano de 1989-90 para formar la cohorte.
4. Porque el estudio comenzó con individuos expuestos y no expuestos, en lugar de enfermos y no enfermos, y observó lo que pasó después, no antes, para determinar la condición de enfermedad.

5. Las condiciones de exposición y enfermedad se observaron en diferentes momentos para distintos individuos. Además, se seleccionó cada individuo una semana antes de que se determinaran sus condiciones de exposición y enfermedad.

P2.5

1. Se calcula que el riesgo de desarrollar la enfermedad es 1,33 veces mayor en personas que se bañaron en aguas contaminadas con respecto a las personas que se bañaron en aguas no contaminadas.
2. No necesariamente. La importancia de cualquier razón de riesgos obtenida depende del criterio clínico de los investigadores y del tamaño de otras razones de riesgo similares que se hayan obtenido en estudios anteriores.
3. ¿La razón de riesgos de 1,33 es relevantemente diferente de una razón de riesgos de 1? Dicho de otra manera, ¿cabe la posibilidad de que la razón de riesgos de 1,33 haya sido fortuita?

P2.6

1. C
2. D
3. E
4. C

P2.7

1. E
2. E
3. C
4. E
5. E
6. D
7. E
8. E
9. E

P2.8

1. El riesgo 2,3 veces mayor significa que los fumadores corren un riesgo 130% mayor que los no fumadores. El riesgo 1,3 veces mayor significa que los ex fumadores corren un riesgo 30% mayor que los no fumadores.

LECCIÓN 3

DISEÑOS DE ESTUDIO EPIDEMIOLÓGICO

*Una etapa clave de la investigación epidemiológica es el **diseño del estudio**. Éste se define como el proceso mediante el cual se planifica una investigación empírica para evaluar una **hipótesis conceptual** sobre la relación entre una o más **exposiciones** y un **resultado de salud**. El objetivo del diseño del estudio es transformar la hipótesis conceptual en una **hipótesis operacional** que pueda comprobarse empíricamente. Puesto que todos los diseños de estudio pueden tener fallas, es importante comprender las virtudes y limitaciones específicas de cada diseño. La mayoría de los problemas o errores serios de esta etapa no pueden corregirse en las etapas posteriores del estudio.*

3-1 Tipos de estudio/opciones

Tipos de investigaciones epidemiológicas

*La investigación epidemiológica puede dividirse en dos categorías generales, según se utilice **aleatorización** o no: los estudios **experimentales** utilizan aleatorización; los estudios **observacionales** no incluyen aleatorización.*

Existen dos tipos generales de estudios epidemiológicos, los **experimentales** y los **observacionales**. Un estudio experimental utiliza la **aleatorización** para asignar individuos a diferentes categorías de exposición. Un estudio observacional no utiliza aleatorización (para obtener más información sobre la aleatorización, remítase al final de esta actividad). En los estudios experimentales, el investigador, a través de la aleatorización, determina la condición de exposición de cada individuo, hace un seguimiento de éstos y documenta el resultado de salud subsiguiente. En un estudio observacional, los propios individuos o, probablemente, su genética determinan su exposición; por ejemplo, si fuman o no fuman. El investigador queda relegado al simple rol de observar la condición de exposición y el resultado de salud subsiguiente.

Los estudios experimentales en epidemiología suelen adoptar la forma de **ensayos clínicos** y **ensayos de intervención comunitaria**. El objetivo de la mayoría de los *ensayos clínicos* es evaluar el efecto posible (es decir, la eficacia) de un tratamiento terapéutico o preventivo, como un nuevo medicamento, fisioterapia o régimen dietario para tratar o prevenir una enfermedad. El objetivo de la mayoría de los *ensayos de intervención comunitaria* es evaluar la eficacia de un programa de prevención. Por ejemplo, uno podría estudiar la eficacia de la fluoración, la educación sexual o el intercambio de agujas.

La mayoría de los estudios epidemiológicos son observacionales. Existen dos tipos generales de estudios observacionales: los **descriptivos** y los **analíticos**. Los estudios descriptivos se realizan para describir el desarrollo natural de una enfermedad, determinar la asignación de recursos de atención médica y sugerir hipótesis sobre la causalidad de la enfermedad. Los estudios analíticos se llevan a cabo para evaluar diversas hipótesis sobre los determinantes de una enfermedad u otro problema de salud y comparten la meta ideal de determinar la causalidad (para obtener más información sobre la causalidad de una enfermedad, remítase al final de esta actividad).

Resumen

- Existen dos tipos generales de estudios epidemiológicos: los experimentales y los observacionales.
- Los estudios experimentales utilizan la aleatorización de exposiciones.
- Los estudios observacionales **no** utilizan la aleatorización de exposiciones.
- En los estudios experimentales, el investigador cumple un rol activo en la determinación de la condición de exposición de cada individuo.
- En los estudios observacionales, el individuo determina su propia condición de exposición.
- Los estudios experimentales suelen ser ensayos clínicos o ensayos de intervención comunitaria.
- Los estudios observacionales son descriptivos o analíticos.

Aleatorización

La aleatorización es un procedimiento de asignación que distribuye a los individuos en los diferentes grupos de exposición que se están comparando, de modo que cada individuo tenga las mismas probabilidades de estar en un grupo. u otro. La aleatorización tiende a hacer similares las características demográficas, conductuales, genéticas, etc. de los grupos comparativos, salvo por su condición de exposición. Por consiguiente, si el estudio descubre alguna diferencia en el resultado de salud entre los grupos comparativos, dicha diferencia sólo puede atribuirse a su diferencia en estatus de exposición.

Por ejemplo, si en forma aleatoria se asigna a los individuos un medicamento nuevo o uno convencional para el tratamiento de la hipertensión, se espera que otros factores, tales como la edad y el sexo, tengan aproximadamente la misma distribución en el caso de los individuos que reciben el medicamento nuevo y en el caso de los que toman el medicamento convencional. En realidad, con la aleatorización, no existe ninguna garantía de que la distribución de edades y otros factores sea equitativa en los dos grupos de tratamiento. El investigador siempre tiene la posibilidad de consultar los datos para ver qué sucedió en relación con una de estas características, siempre y cuando dicha característica se haya medido u observado en el estudio. Si se determina que la distribución de edades no ha sido equilibrada entre los dos grupos de tratamiento, el investigador puede tenerlo en cuenta en el análisis, por ejemplo, mediante la estratificación por edad.

La ventaja de la aleatorización es lo que ofrece en relación con las características que no se miden en el estudio. Por supuesto, las variables que no se miden no pueden considerarse en el análisis. La aleatorización hace posible, aunque no garantiza, que estas variables no medidas se distribuyan de manera uniforme entre los diferentes grupos de exposición. Por otro lado, en los estudios observacionales, el investigador puede dar cuenta únicamente de las variables que se miden, en cuyo caso es más probable que se obtengan conclusiones falsas debido a los efectos desconocidos de importantes variables no medidas.

Causalidad

En cualquier campo de investigación que involucre la realización de investigaciones científicas y el análisis de los datos obtenidos a partir de ellas para evaluar hipótesis etiológicas, la determinación de la **causalidad** es un tema complicado. En particular, la capacidad para hacer **inferencias causales** en las ciencias de salud suele depender de la síntesis de los resultados de varios estudios, tanto epidemiológicos como no epidemiológicos (por ejemplo, resultados clínicos o de laboratorio).

Mediante un esfuerzo patrocinado por el gobierno estadounidense para evaluar los efectos del tabaquismo en la salud, diversos científicos de la salud a finales de la década de los cincuenta y principios de los sesenta comenzaron a considerar la posibilidad de definir criterios objetivos para la evaluación de la causalidad. El esfuerzo se centraba especialmente en cómo abordar la causalidad en función de los resultados de los estudios que consideran exposiciones que no se pueden asignar en forma aleatoria; es decir, los estudios observacionales.

En 1964, el Departamento de Salud, Educación y Bienestar de los Estados Unidos publicó un informe en el que se analizaban los resultados de las investigaciones de los efectos del tabaquismo sobre la salud, con el objetivo de determinar si el tabaquismo podía identificarse como una "causa" del cáncer de pulmón y probablemente otras enfermedades. El tipo de síntesis que se realizó en este informe se describió, en la década de los noventa, como un metaanálisis, de modo que este informe fue básicamente uno de los primeros ejemplos de **metaanálisis** que se desarrollaron en las ciencias de la salud.

Muchas de las conclusiones del documento de 1964 sobre la causalidad del tabaquismo se basaron en una lista de criterios generales formalizados por Bradford Hill, los cuales se incorporaron luego en un famoso libro

de texto de Hill en 1971. Los criterios son los siguientes:

1. **Grado de asociación:** cuanto más fuerte sea la asociación observada, menores serán las probabilidades de que la asociación se deba a un sesgo; las asociaciones más débiles no ofrecen un gran respaldo para una interpretación causal.
2. **Efecto dosis-respuesta:** si la frecuencia de enfermedad aumenta con la dosis o el nivel de exposición, esto sirve de respaldo para una interpretación causal (no obstante, debe tenerse en cuenta que la ausencia de un efecto dosis-respuesta podría no descartar la causalidad mediante explicaciones alternativas tales como un efecto de umbral).
3. **Ausencia de ambigüedad temporal**: la causa hipotética debe ser anterior a la aparición de la enfermedad.
4. **Consistencia de los resultados**: si todos los estudios que tratan una relación determinada producen resultados similares, esto promueve la generación de una interpretación causal. Nota: las inconsistencias podrían deberse a diferencias en las características del diseño del estudio, por lo que es posible que cada estudio deba ponderarse de algún modo.
5. **Plausibilidad biológica de la hipótesis**: si el efecto hipotético tiene sentido en el contexto de los principios biológicos aceptados en ese momento, esto sirve de apoyo para una interpretación causal (no obstante, debe tenerse en cuenta que el estado actual de los conocimientos biológicos podría ser inadecuado para determinar la plausibilidad biológica).
6. **Coherencia de la evidencia**: si los resultados no plantean una contradicción seria con nuestra comprensión de la historia natural de la enfermedad u otros hechos aceptados sobre la aparición de la enfermedad, esto sirve de apoyo para una interpretación causal.
7. **Especificidad de la asociación:** si se descubre que el factor de estudio se asocia a una única enfermedad o que la enfermedad se asocia a un único factor, esto sirve de apoyo para una interpretación causal (sin embargo, este criterio no puede descartar una hipótesis causal, puesto que existen varios factores que tienen múltiples efectos y la mayoría de las enfermedades tienen varias causas). Algunos ejemplos incluyen cloruro de vinilo y angiosarcoma hepático; y consumo de DES en mujeres [durante el embarazo] y cáncer vaginal en sus hijas.
8. **Experimentación:** uso de evidencia experimental, lo que incluye ensayos clínicos en humanos, modelos animales y experimentos de laboratorio *in vitro*. Cuando dicha evidencia está disponible puede respaldar teorías causales, pero su ausencia no excluye la causalidad.
9. **Analogía:** cuando se han demostrado relaciones similares por medio de otras relaciones entre exposiciones y enfermedades. Por ejemplo, las hijas de mujeres a las que se les había dado DES durante el embarazo tenían mayores probabilidades de desarrollar cáncer vaginal. Por analogía, es posible que otros medicamentos para embarazadas produzcan cáncer en sus hijos.



Auto-evaluación (P3.1)

Complete los espacios vacíos con **Experimental** u **Observacional**

1. Una de las fortalezas del estudio **???** es la capacidad que tiene el investigador de controlar la asignación de individuos a grupos de tratamiento.

2. Una ventaja potencial de un estudio **???** es que suele desarrollarse en entornos más naturales, de modo que la población del estudio es más representativa de la población que se está analizando.

3. La principal limitación del estudio **???** es que le otorga al investigador un control muy reducido de la situación de estudio. Por lo tanto, los resultados suelen ser más susceptibles a influencias que distorsionan la realidad.

4. Uno de los puntos débiles del estudio **???** es que la distribución aleatoria en grupos de tratamiento podría

ser poco ética en el caso de que a un grupo arbitrario de individuos deba negársele un tratamiento que se considere beneficioso.

5. En un estudio de epidemiología de lesiones, se seleccionó a una comunidad de un estado para una campaña mediática en la que los individuos recibían descuentos en los cascos con la compra de cualquier bicicleta. A 50 kilómetros de ese lugar, se identificó una comunidad similar, la cual se utilizó como comunidad comparativa. Los epidemiólogos compararon la incidencia de las lesiones por uso de bicicletas mediante encuestas telefónicas y vigilancia de salas de emergencia. Éste es un ejemplo de estudio **???**

6. Los investigadores de un estudio entregaron un cuestionario a todos los alumnos nuevos de una universidad estatal importante. El cuestionario incluía preguntas sobre diversos hábitos tales como el uso del cinturón de seguridad, actividad física, tabaquismo y consumo de alcohol. Los investigadores planean distribuir cuestionarios de seguimiento en la graduación y luego cada cinco años para obtener información sobre episodios y afecciones relacionadas con la salud, incluso diabetes y enfermedades cardiacas. Éste es un ejemplo de estudio **???**

Direccionalidad

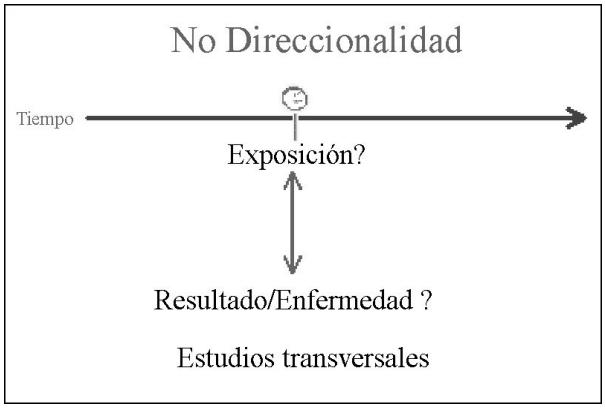
La direccionalidad de un estudio hace referencia al momento en que se observa la variable de exposición con respecto al momento en que se observa el resultado de salud. En un estudio con direccionalidad hacia adelante, el investigador empieza por determinar la condición de exposición de los individuos seleccionados a partir de una cierta población de interés; luego hace un seguimiento de estos individuos a lo largo del tiempo para determinar si desarrollan o no el resultado de salud. Los estudios de cohorte y los ensayos clínicos siempre tienen direccionalidad hacia adelante.



En un diseño con direccionalidad hacia atrás, el investigador selecciona a los individuos según presenten o no el resultado de salud de interés y luego obtiene información sobre sus exposiciones anteriores. Los estudios de casos y controles siempre tienen direccionalidad hacia atrás.



En un diseño no direccional, el investigador observa tanto el factor de estudio como el resultado de salud en forma simultánea, de modo que ninguna de las variables puede identificarse exclusivamente como anterior a la otra. Los estudios transversales son siempre no direccionales.



La direccionalidad de un estudio incide en la capacidad del investigador para distinguir lo precedente de lo subsiguiente. Esto es importante para evaluar la **causalidad**. Además, la direccionalidad afecta el modo en que pueden seleccionarse los individuos del estudio. Los diseños hacia atrás o los no direccionales son más susceptibles a **sesgos de selección** que los diseños hacia adelante. Los sesgos de selección se analizan en profundidad en una lección posterior de este programa.

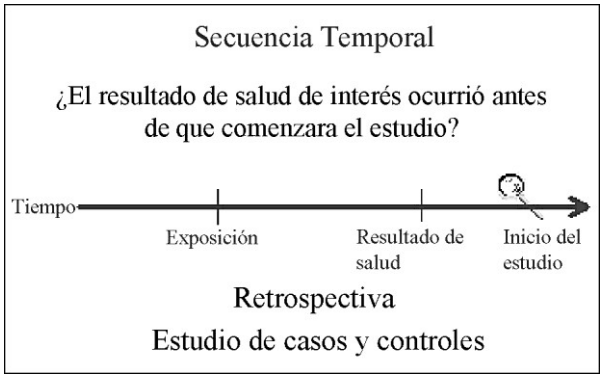
Resumen

- La direccionalidad responde a la pregunta: ¿cuándo se observó la variable de exposición con respecto al momento en que se observó el resultado de salud?
- Un estudio puede ser hacia adelante, hacia atrás o no direccional.
- La direccionalidad incide en la capacidad del investigador para distinguir lo precedente de lo subsiguiente.
- La direccionalidad también repercute en el hecho de que un estudio presente o no sesgos de selección.

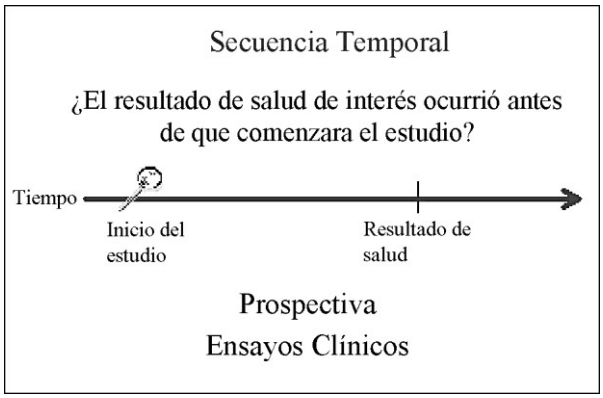
Secuencia temporal

La secuencia temporal responde a la pregunta de si el resultado de salud de interés ocurrió antes de que comenzara el estudio. Si el resultado de salud es anterior al inicio del estudio, la secuencia temporal es **retrospectiva**. Por ejemplo, supongamos que se inicia un estudio de casos y controles para investigar casos de una enfermedad que se produjo el año anterior. Esto sería un ejemplo de un estudio **retrospectivo** de casos y

controles.



Por el contrario, si el resultado de salud se presenta después del inicio del estudio, la secuencia temporal es **prospectiva**. Los ensayos clínicos son siempre *prospectivos*.



Los estudios de cohorte y los estudios de casos y controles pueden ser retrospectivos *o* prospectivos, puesto que el estudio puede comenzar antes o después de la aparición del resultado de salud. La secuencia temporal de un estudio puede tener implicaciones importantes para la calidad de los datos. Los datos retrospectivos suelen basarse en recuerdos personales o en registros hospitalarios o laborales y, por lo tanto, tienen más probabilidades de incluir errores de medición que los estudios prospectivos. Según se explica en una lección posterior de este programa, los errores de medición frecuentemente producen sesgos de información.

Algunos estudios pueden tener elementos tanto **prospectivos** como **retrospectivos**, lo cual en ocasiones se denomina secuencia temporal **mixta**.

Resumen

- La secuencia temporal responde a la pregunta: ¿el resultado de salud de interés ocurrió antes de que comenzara el estudio?
- Si el resultado de salud es anterior al inicio del estudio, la secuencia temporal es retrospectiva.
- Si el resultado de salud es posterior al inicio del estudio, la secuencia temporal es prospectiva.
- La secuencia temporal incide en los errores de medición y los sesgos de información.

Ensayos clínicos

El ensayo clínico es el diseño epidemiológico que más se parece a un experimento de laboratorio. El objetivo principal consiste en evaluar el posible efecto de una intervención terapéutica o preventiva.

El ensayo clínico es un *estudio experimental* diseñado para comparar los beneficios terapéuticos o médicos de dos o más tratamientos. El objetivo principal de un ensayo clínico es evaluar la eficacia de una intervención **preventiva** o **terapéutica**. El fin de un ensayo preventivo es prevenir una enfermedad, mientras que el de un ensayo terapéutico es curar o controlar una enfermedad. Algunos ejemplos de ensayos preventivos incluyen estudios de eficacia de vacunas, consumo de aspirina para prevenir una enfermedad cardiaca, cese del hábito de fumar, modificación en la dieta y actividad física. Las compañías farmacéuticas suelen realizar ensayos terapéuticos para probar nuevos medicamentos contra una enfermedad.

Las características claves de cualquier ensayo clínico son la **aleatorización**, el **enmascaramiento**, las **cuestiones éticas** y el uso de **análisis por intención de tratar**. La **aleatorización** se utiliza para distribuir a los individuos en grupos de tratamiento de modo tal que estos grupos puedan compararse con respecto a todos los factores, a excepción de la condición de exposición. El **enmascaramiento**, o **ciego**, significa que el paciente o el investigador no conocen el tratamiento asignado. En un estudio ciego, el paciente o el investigador desconocen el tratamiento asignado; en un estudio doble ciego, tanto el paciente como el investigador desconocen el tratamiento asignado. El enmascaramiento ayuda a eliminar el sesgo. El estudio debe ser **ético**, por lo que no se utilizan tratamientos perjudiciales para la salud. Se deben prever reglas de interrupción para dar por terminado el ensayo en forma prematura en el caso de que uno de los tratamientos sea claramente superior. Un **análisis por intención de tratar** exige que los investigadores "analicen lo aleatorizado"; es decir, el análisis debe compararse con los grupos de tratamiento originariamente aleatorizados, aun cuando los individuos cambien de tratamiento durante el período del estudio.

Resumen

- El objetivo principal de un ensayo clínico es evaluar la eficacia de una intervención preventiva o terapéutica.
- Las características claves de cualquier ensayo clínico son:
  - Aleatorización
  - Enmascaramiento
  - Cuestiones éticas
  - Análisis por intención de tratar

Eficacia, efectividad y eficiencia

La valoración de una modalidad de tratamiento o intervención médica nueva o existente comprende tres pasos (Detsky, 1995; Detsky y Naglie, 1990; Grimes y Schulz, 2002). Primero, se demuestra la *eficacia* (el logro del objetivo clínico establecido) en condiciones *óptimas* mediante un ensayo controlado aleatorizado y prospectivo. Luego, se evalúa la *efectividad* (la capacidad de producir más beneficios que daños) en condiciones *ordinarias* entre la población en general mediante un estudio de cohorte observacional prospectivo o un ensayo experimental de intervención comunitaria. Por último, se determina la *eficiencia* de la intervención médica (la mejoría en la condición de salud que se obtiene mediante la inversión de una cantidad determinada de recursos) a través de un análisis de costo-efectividad o un análisis de costo-utilidad. Por otra parte, una evaluación económica de esta naturaleza puede aportar información esencial sobre los recursos necesarios para desarrollar la intervención médica en una población específica (Kocher y Henley, 2003).

Ejemplo de ensayo clínico

En un ensayo clínico de 1993 en el que participaron 726 individuos, se desarrolló una comparación entre una terapia *convencional* con insulina y una terapia *intensiva* con insulina, la cual comprendía una mayor

frecuencia en las inyecciones de insulina y los análisis de glucosa en sangre para el tratamiento de la diabetes mellitus. El resultado que se estudiaba era una retinopatía con posterior ceguera, definido en términos de presencia o ausencia en cada paciente.

Los individuos se distribuyeron de manera aleatoria en grupos de tratamiento mediante el uso de un generador de números aleatorios por computadora. En este ensayo clínico no se podía usar el método de doble ciego, puesto que tanto el paciente como su médico conocerían el grupo de tratamiento del paciente. No obstante, los individuos que analizaban las imágenes de fondo de ojo para determinar la presencia o ausencia de retinopatía desconocían los grupos de tratamiento asignados. La aleatorización dio como resultado una distribución bastante equitativa con respecto a las características de referencia tales como la edad y el sexo, entre los grupos de terapia convencional e intensiva.

Se realizó un análisis por intención de tratar para comparar los grupos de tratamiento originariamente aleatorizados en relación con la aparición de la retinopatía. Se determinó que el 24% de los 378 individuos que recibieron la terapia convencional desarrollaron retinopatía, mientras que el 6,7% de los 348 individuos que recibieron la terapia intensiva desarrollaron retinopatía.

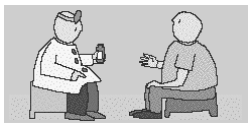
Ensayo Clínico		
Tratamiento:		
Terapia Convencional	Vs.	Terapia Intensiva
n=378		**n=348**
24%		**6.7%**

Según estos datos y otros análisis más complejos diseñados para controlar otros predictores importantes, la terapia intensiva demostró tener un riesgo de retinopatía mucho más bajo que la terapia convencional.

Resumen

- En un ensayo clínico de 1993 en el que participaron 726 individuos, se desarrolló una comparación entre una terapia convencional con insulina y una terapia intensiva con insulina.
- El resultado que se investigaba era una retinopatía con posterior ceguera, definido en términos de presencia o ausencia en cada paciente.
- El 24% de los individuos que recibieron la terapia convencional desarrollaron retinopatía, mientras que el 6,7% de los que recibieron la terapia intensiva desarrollaron retinopatía.

Auto-evaluación (P3.2)



Complete los espacios vacíos

1. En los ensayos **???** se incluyen individuos con una enfermedad determinada para evaluar su posible cura o control. Por ejemplo, podríamos determinar, si ello fuera posible, en qué medida un nuevo tipo de quimioterapia prolonga la vida de los niños que padecen leucemia linfática aguda.

2. Los ensayos **???** pueden incluir individuos o poblaciones enteras. Un ejemplo de esto es un estudio en el que se decidió que una comunidad (asignada en forma aleatoria) recibiría un agregado de fluoruro de sodio en el suministro de agua, mientras que la otra comunidad seguiría recibiendo agua sin dicho agregado. El estudio demostró que la comunidad que recibía fluoruro experimentó reducciones significativas en el desarrollo de caries.

Opciones

Preventivo **Terapéutico**

Para cada una de las siguientes características, seleccione la opción que aplique a los ***ensayos clínicos***:

1. Rol del investigador con respecto a la exposición: . . . **???**
   a. asignar b. observar

2. Distribución de individuos en grupos: . . . . . **???**
   a. selección personal b. aleatorización

3. Direccionalidad: . . . . . **???**
   a. hacia atrás b. hacia adelante c. no direccional

4. Secuencia temporal: . . . . . **???**
   a. prospectiva b. retrospectiva c. indistinta

5. Enmascaramiento: . . . . . **???**
   a. ciego b. doble ciego c. indistinto

6. Asunto: . . . . . **???**
   a. medicación b. vacuna c. indistinto

7. Análisis según: . . . . . **???**
   a. asignación original b. experiencia real

3-2 Diseños de estudios observacionales

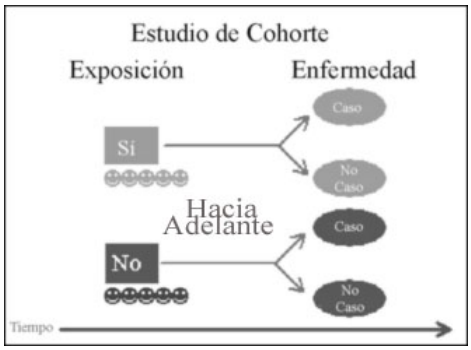
Existen tres categorías generales de diseños de estudios observacionales:

- Diseños básicos: de cohorte, de casos y controles, transversales.
- Diseños híbridos: de casos y controles anidados, de caso-cohorte.
- Diseños incompletos: ecológicos, proporcionales.

Estudios de cohorte

En 1948 se inició un estudio observacional de largo plazo en Framingham, Massachusetts. Se seleccionaron y examinaron 5.100 individuos que no padecían enfermedades cardiovasculares (CVD, por sus siglas en inglés) y se registró información sobre posibles factores de riesgo para el desarrollo de estas enfermedades. En los casos en que resultó posible, los individuos volvieron a examinarse cada 2 años durante los siguientes 50 años. Este estudio clásico se conoce con el nombre de *Estudio de Framingham* y ha dado origen a gran parte de nuestros conocimientos sobre los factores de riesgo de las enfermedades cardiovasculares. El estudio de Framingham es un ejemplo de estudio de cohorte **prospectivo**.

Un diseño de cohorte comienza con la selección de individuos que no presentan un resultado de salud de interés, a los cuales luego se les hace un futuro seguimiento para determinar la condición de salud. Un rasgo clave del estudio de cohorte es que los individuos se agrupan en función de sus características de exposición antes de observar el resultado de salud; es decir, la direccionalidad del estudio es siempre hacia adelante.



Un estudio de cohorte puede ser retrospectivo o prospectivo. El estudio de Framingham es un ejemplo de estudio prospectivo, puesto que comenzó antes de que se produjera el resultado de salud.

Resumen

- El estudio de Framingham es un ejemplo clásico de un estudio de cohorte.
- El diseño de cohorte es siempre un estudio de seguimiento con direccionalidad hacia adelante.
- Un estudio de cohorte puede ser prospectivo o retrospectivo.
- El estudio de Framingham es un estudio de cohorte prospectivo porque comenzó antes de que se produjera el resultado de salud.

Variable de exposición en estudios de cohorte

Si todas las variables de exposición de interés son relativamente comunes tales como las que se miden en el estudio original de Framingham, la cohorte suele determinarse realizando un muestreo de personas pertenecientes a una población numerosa y, después de excluir a las que ya presentan el resultado de salud, dividiendo la muestra restante en individuos expuestos y no expuestos.

Si la exposición es poco frecuente, como cuando se estudia una exposición ocupacional específica, los expuestos suelen muestrearse a partir de una población especial como, por ejemplo, el lugar de trabajo. Luego se determinan los no expuestos a partir de un grupo comparativo externo, el cual debe ser lo más parecido posible al grupo de individuos expuestos con respecto a los factores, a excepción de la exposición, que pueden relacionarse con la enfermedad.

Además, las personas empleadas, en promedio, suelen ser más saludables que las desempleadas. Por consiguiente, si los trabajadores expuestos se comparan con la población en general, es probable que se subestime el efecto de la exposición.

Ventajas del estudio de cohorte

La principal ventaja de un estudio de cohorte es su direccionalidad hacia adelante. El investigador puede estar seguro, en la medida de lo razonable, de que la causa hipotética es anterior a la aparición de la enfermedad. En un estudio de cohorte, la condición de enfermedad no puede influir en el método de selección de los individuos, de modo que un estudio de cohorte no tiene ciertos sesgos de selección que presentan limitaciones serias en otros tipos de estudios.

Un diseño de cohorte prospectivo es menos propenso a obtener información incorrecta sobre variables importantes que otros diseños de estudio observacionales. Los estudios de cohorte pueden utilizarse para estudiar varias enfermedades, puesto que el seguimiento permite identificar varios resultados de salud.

Los estudios de cohorte también sirven para examinar exposiciones *poco frecuentes*. Como el investigador selecciona a los individuos en función de su exposición, puede garantizar una cantidad suficiente de individuos expuestos. Un estudio de cohorte *retrospectivo* puede ser relativamente económico y rápido. Un ejemplo de esto son los estudios ocupacionales que se basan en registros laborales y actas de defunción o registros de indemnización laboral o de seguro.

Desventajas del estudio de cohorte

El estudio de cohorte prospectivo a menudo requiere una inversión considerable de tiempo y dinero. Un problema potencial de cualquier estudio de cohorte es la pérdida de individuos por migración, falta de participación, abandono del estudio y muerte. Tal desgaste de la cohorte durante el período de seguimiento puede producir resultados sesgados.

Un diseño de cohorte es ineficaz en términos estadísticos y prácticos para el estudio de las enfermedades *poco frecuentes* que tienen una latencia prolongada debido a la considerable extensión del período de seguimiento y la cantidad de individuos que se requiere para identificar una cantidad suficiente de casos. No obstante, un estudio de cohorte retrospectivo puede encontrar casos suficientes porque los episodios de interés del estudio ya se han producido.

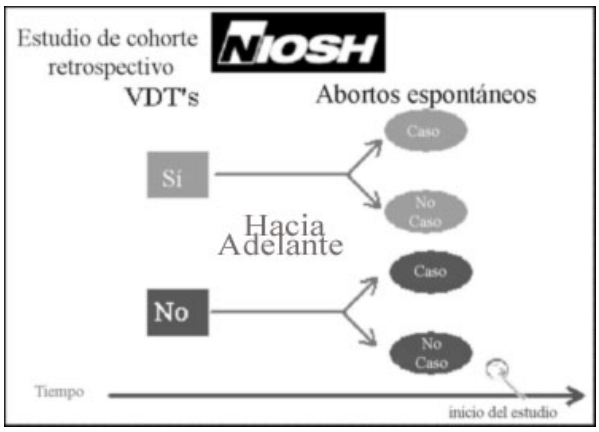
Otro problema de los estudios de cohorte es que los individuos expuestos pueden seguirse más de cerca que los no expuestos. Si esto sucede, es más probable que se diagnostique el resultado en los individuos expuestos. Esto podría dar la impresión de que existe una relación entre la exposición y la enfermedad cuando no la hay.

Resumen: ventajas (+) y desventajas (-) de los estudios de cohorte

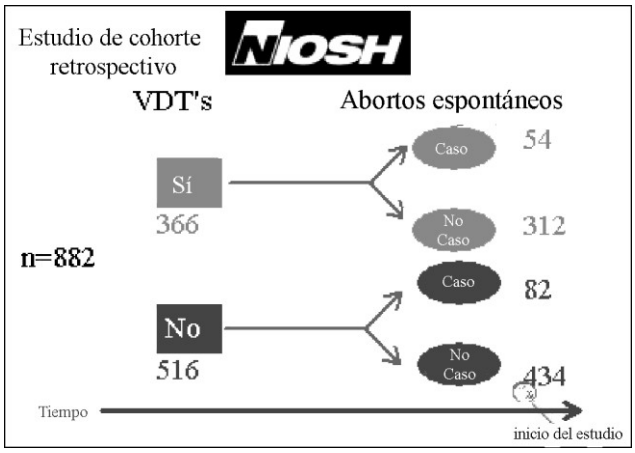
- (+) Estudio de cohorte prospectivo: menos propenso a sesgos que otros diseños de estudio observacionales.
- (+) Puede abordar varias enfermedades en el mismo estudio.
- (+) Estudio de cohorte retrospectivo: puede ser relativamente económico y rápido; es común en estudios ocupacionales.
- (-) La pérdida de individuos durante el seguimiento es una fuente potencial de sesgo.
- (-) Estudio de cohorte prospectivo: requiere una inversión considerable de tiempo y dinero; si la enfermedad es poco frecuente, la cantidad de casos podría ser insuficiente.
- (-) Si a los individuos expuestos se les sigue más de cerca que a los no expuestos, es más probable que el resultado se diagnostique en los expuestos.

Ejemplo de estudio de cohorte retrospectivo: terminales de video y abortos espontáneos

La relación entre los resultados adversos del embarazo y el uso de terminales de video (VDT, por sus siglas en inglés) se convirtió en un problema de salud pública en la década de 1980 cuando se informaron resultados adversos del embarazo en varios grupos de mujeres que usaban VDT. En la publicación *New England Journal of Medicine* de 1991 se incluyó un estudio más integral del efecto de los VDT. En este estudio realizado por el Instituto Nacional para la Seguridad y Salud Ocupacional (NIOSH, por sus siglas en inglés), se utilizó un diseño de cohorte retrospectivo para analizar la hipótesis de que la energía electromagnética producida por los VDT podía generar abortos espontáneos.



En el estudio del NIOSH se seleccionó una cohorte de telefonistas empleadas entre 1983 y 1986 a partir de registros de personal de dos compañías telefónicas en ocho estados del sudeste de los Estados Unidos. En esta cohorte había 882 mujeres cuyos embarazos cumplían con los criterios de inclusión en el estudio. De estas mujeres, los resultados de embarazo de las 366 telefonistas que proporcionaban información de guía telefónica y usaban VDT en el trabajo se compararon con los de las 516 telefonistas generales que no usaban VDT.



Preguntas de estudio (P3.3)

1. ¿Qué porcentajes de mujeres desarrollaron abortos espontáneos en el caso de las usuarias de VDT y en el de las que no los usaban?

Los resultados del estudio no demostraron riesgos significativos de sufrir abortos espontáneos entre las mujeres que usaron VDT durante el primer trimestre del embarazo. Tampoco se encontró una relación dosis-respuesta mediante el análisis de la cantidad de horas semanales de uso de VDT por parte de las mujeres. Además, cuando se consideraron otras características relevantes de los individuos del estudio, el riesgo por el uso de VDT resultó insignificante. Los investigadores llegaron así a la conclusión de que el uso de terminales de video y la exposición a los campos electromagnéticos que éstos producían no estaban relacionados con un mayor riesgo de sufrir un aborto espontáneo.

Resumen

- Un estudio de 1991 utilizó un diseño de cohorte retrospectivo para analizar la hipótesis de que la energía electromagnética producida por los terminales de video (VDT) podía generar abortos espontáneos.
- Los resultados de embarazo de las 366 telefonistas que proporcionaban información de guía telefónica y usaban VDT en el trabajo se compararon con los de las 516 telefonistas generales que no usaban VDT.
- Los resultados del estudio no demostraron riesgos significativos de sufrir abortos espontáneos entre las mujeres que usaron VDT.

Ejemplo de estudio de cohorte prospectivo: estudio de Rotterdam sobre el mal de Alzheimer

Se cree que la actividad inflamatoria en el cerebro contribuye al desarrollo del mal de Alzheimer. Esta hipótesis sugiere que el consumo prolongado de medicamentos antiinflamatorios no esteroideos, o AINE, puede reducir el riesgo de padecer esta enfermedad.

Esta hipótesis se investigó en el marco del estudio de Rotterdam, un estudio de cohorte con personas de edad avanzada que se inició en los Países Bajos en 1990. En ese momento, 7.000 participantes no padecían el mal de Alzheimer. Durante ocho años de seguimiento, 293 de estos individuos desarrollaron la enfermedad.

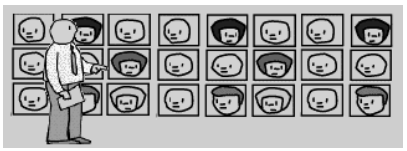
Preguntas de estudio (P3.4)

1. ¿Cuál es la direccionalidad de este estudio?
2. ¿La secuencia temporal es prospectiva o retrospectiva?

A fin de evitar el sesgo de información proveniente de la medición de los AINE, los investigadores utilizaron registros farmacéuticos computarizados en lugar de datos de entrevistas para determinar la cantidad total de meses durante los cuales los participantes habían consumido AINE después del inicio del estudio. Tras controlar los factores de edad, sexo y tabaquismo, los investigadores establecieron que el riesgo de Alzheimer de los participantes que habían consumido AINE por más de 24 meses era significativamente menor que el de los participantes que habían consumido AINE durante 24 meses o menos. Los investigadores llegaron a la conclusión de que el consumo prolongado de AINE reduce el riesgo de desarrollar el mal de Alzheimer.

Resumen

- El estudio de Rotterdam analizó la hipótesis de que el consumo prolongado de medicamentos antiinflamatorios no esteroideos (AINE) puede reducir el riesgo de desarrollar el mal de Alzheimer.
- En el estudio se empleó un diseño de cohorte prospectivo y se les hizo un seguimiento a 7.000 participantes sin Alzheimer durante 8 años a partir de 1990.
- El riesgo de Alzheimer de los individuos que consumieron AINE por más de 24 meses resultó significativamente menor que el de los participantes que consumieron AINE durante 24 meses o menos.



Auto-evaluación (P3.5)

Complete los espacios vacíos
Para cada una de las siguientes características, seleccione la opción que aplique a los ***estudios de cohorte***:

1. Rol del investigador con respecto a la exposición: . . . **???**
   a. asignar b. observar
2. Distribución de individuos en grupos: . . . . . **???**
   a. selección personal b. aleatorización
3. Direccionalidad: . . . . . **???**
   a. hacia atrás b. hacia adelante c. no direccional
4. Secuencia temporal: . . . . . **???**
   a. prospectiva b. retrospectiva c. indistinta
5. Análisis según: . . . . . **???**
   a. asignación original b. experiencia real

Para cada una de las siguientes características (puntos fuertes o débiles) de un estudio, seleccione el tipo de estudio de cohorte que tenga dicha característica:

6. Más económico: . . . . . **???**
7. Más rápido: . . . . . **???**
8. La información sobre la exposición es más precisa: . . . **???**
9. Adecuado para el estudio de exposiciones poco frecuentes: . . **???**
10. Adecuado para el estudio de enfermedades poco frecuentes: . . **???**
11. Presenta problemas por la pérdida de individuos durante el seguimiento: . **???**
12. Mejor para enfermedades con latencia prolongada: . . . **???**

Opciones
Ambos **Ninguno** **De cohorte prospectivo** **De cohorte retrospectivo**

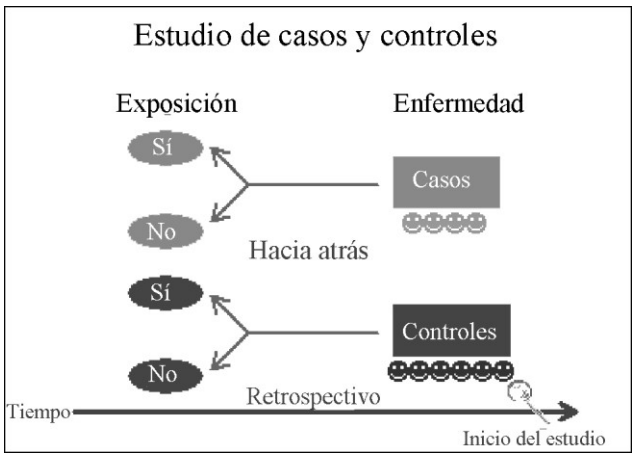
3-3 Estudios de casos y controles y estudios transversales

Estudios de casos y controles

El estudio de casos y controles es un diseño de estudio observacional básico que puede ser tanto retrospectivo como prospectivo. Su implementación suele ser bastante económica y rápida, pero es muy propenso a sesgos si se compara con un diseño de cohorte.

En los estudios de casos y controles, los individuos se seleccionan en función de su condición de enfermedad. El investigador primero selecciona *casos* de una enfermedad determinada y luego selecciona *controles* de entre personas que no padecen la enfermedad. Lo ideal es seleccionar los casos a partir de una población claramente definida, a menudo denominada **población fuente**, y escoger los controles de la misma

población que produjo los casos. Luego se determinan los antecedentes de exposición de los casos y controles. Por lo tanto, a diferencia de un estudio de cohorte, un estudio de casos y controles funciona hacia atrás, desde la condición de enfermedad hasta la condición de exposición previa. Si bien los estudios de casos y controles tienen siempre una direccionalidad hacia atrás, pueden ser prospectivos o retrospectivos en términos de secuencia temporal.



Además de ser más económicos y rápidos que los estudios de cohorte, los estudios de casos y controles presentan otras ventajas:

- son factibles con respecto a la obtención de casos suficientes para el estudio de enfermedades crónicas o poco frecuentes o enfermedades con períodos de latencia prolongados;
- suelen requerir una muestra más pequeña que otros diseños;
- permiten evaluar el efecto de una gran variedad de exposiciones diferentes.

No obstante, los estudios de casos y controles también presentan una serie de desventajas:

- a diferencia de los estudios de cohorte, no permiten evaluar varias enfermedades;
- no permiten hacer un cálculo directo del riesgo de la enfermedad porque funcionan hacia atrás, desde la condición de enfermedad hasta la condición de exposición;
- son más susceptibles al sesgo de selección que otros diseños, puesto que la exposición es anterior a la selección de los casos y controles;
- son más susceptibles al sesgo de información que los estudios de cohorte debido a su direccionalidad hacia atrás y porque pueden ser retrospectivos en términos de secuencia temporal;
- no son eficientes para el estudio de exposiciones poco frecuentes.

Resumen

- ❖ Comienzan con casos y no casos de una enfermedad u otro resultado de salud y se retrotraen en el tiempo para determinar los antecedentes de exposición.
- ❖ Uso difundido porque requieren una menor inversión de tiempo y dinero que los estudios de cohorte.
- ❖ Otras ventajas incluyen la capacidad de obtener una cantidad suficiente de casos de enfermedades poco frecuentes con períodos de latencia prolongados y permiten evaluar varias exposiciones.
- ❖ Entre las desventajas se incluyen la susceptibilidad a sesgos de selección e información, dificultades para calcular el riesgo, incapacidad de considerar más de una enfermedad e inviabilidad para el estudio de exposiciones poco frecuentes.

Casos incidentes y prevalentes en un estudio de casos y controles

Los casos pueden seleccionarse como incidentes o prevalentes. Los casos incidentes son los casos nuevos de una enfermedad que se desarrollan a lo largo del período cubierto por el estudio de casos y controles. Cuando se utilizan en estudios de casos y controles, los casos incidentes suelen obtenerse de una oficina de registro de enfermedades de base poblacional o institucional, como un registro del cáncer, o de una organización de mantenimiento de la salud que registre de manera constante las enfermedades nuevas en una población específica.

Los casos prevalentes son casos existentes de una enfermedad en un momento determinado. Cuando se utilizan en estudios de casos y controles, los casos prevalentes suelen obtenerse de registros hospitalarios o clínicos.

Una ventaja del uso de casos incidentes en estudios de casos y controles es que la relación entre la exposición y la enfermedad sólo puede vincularse al desarrollo, y no al pronóstico o la duración de la enfermedad.

Por el contrario, en los casos prevalentes, la exposición puede incidir en el pronóstico o la duración de la enfermedad. Por lo tanto, si se utilizaran casos prevalentes, el cálculo del efecto de la exposición en el desarrollo de la enfermedad podría estar sesgado debido a la omisión de los individuos que murieron antes de la selección de casos.

Selección de controles en un estudio de casos y controles

Cuando se realiza un estudio de casos y controles, se debe seleccionar un grupo comparativo o de control con sumo cuidado. El grupo de control ideal debería ser representativo de la población de la que se obtuvieron los casos, la cual suele denominarse **población fuente**. Este ideal a menudo es difícil de alcanzar cuando se seleccionan controles.

Dos de los tipos de controles más comunes son los **controles de base poblacional** y los **controles de base hospitalaria**. En los estudios de casos y controles de base poblacional, los controles se seleccionan de la comunidad. Los métodos empleados para seleccionar dichos controles incluyen el discado telefónico aleatorio, los amigos o vecinos y listados de departamentos de vehículos motorizados. Una ventaja de los estudios de casos y controles de base poblacional es que los casos y controles provienen de la misma población fuente, por lo que en cierto modo son similares. Una desventaja es la dificultad para obtener listas poblacionales y para identificar e inscribir a los individuos. El creciente uso de números telefónicos protegidos y contestadores automáticos aumenta la falta de respuesta por parte de los controles potenciales.

En un estudio de casos y controles de base hospitalaria, los controles se seleccionan a partir de pacientes de hospitales que padecen enfermedades *diferentes de la enfermedad de interés*. Los controles hospitalarios son más accesibles y suelen ser más cooperativos que los controles de base poblacional. Los estudios de base hospitalaria requieren una inversión de tiempo y dinero mucho menor que los estudios de base poblacional. No obstante, es probable que los controles de base hospitalaria no sean representativos de la población fuente que produjo los casos. Además, los controles de base hospitalaria padecen alguna enfermedad y la exposición de interés podría ser un determinante tanto de la enfermedad de control como de la enfermedad de interés. En este caso, es probable que no se encuentre una asociación real entre la exposición y la enfermedad de interés.

Resumen

- El grupo de control ideal debería ser representativo de la *población fuente* de la que se obtuvieron los casos.
- Dos de los tipos de controles más comunes son los *controles de base poblacional* y los *controles de base hospitalaria*.
- En los estudios de casos y controles de base poblacional, los casos y los controles provienen de la misma población fuente.
- Los controles hospitalarios suelen ser más accesibles, cooperativos y económicos.

- ❖ Por lo general, los controles hospitalarios no son representativos de la población fuente y pueden representar una enfermedad producida por la exposición.

Estudios de casos y controles: ejemplo del síndrome de Reye

Varios estudios en las décadas de los setenta y ochenta utilizaron un diseño de casos y controles para determinar si el consumo de aspirina se relacionaba con la aparición del síndrome de Reye en niños con enfermedades virales.

El síndrome de Reye es una enfermedad poco frecuente que afecta el cerebro y el hígado, y puede producir delirios, estados de coma y la muerte. Suele afectar a niños y, por lo general, se manifiesta después de una enfermedad viral. Para establecer si la aspirina era un determinante del síndrome de Reye, los investigadores de los años setenta y ochenta decidieron que el uso de un ensayo clínico sería poco ético.

¿Por qué podría ser poco ético un ensayo clínico sobre el consumo de aspirina y el síndrome de Reye?

A. Hay niños involucrados.
B. Efectos nocivos del consumo de aspirina.
C. Se puede usar el método de doble ciego.
D. Los ensayos clínicos siempre son poco éticos.

La respuesta es la B, por los efectos potencialmente nocivos del consumo de aspirina. También se decidió que un estudio de cohorte sería ineficiente:

¿Por qué podría ser ineficiente un estudio de cohorte de aspirina y el síndrome de Reye?

A. El resultado es poco frecuente (se requeriría una gran cantidad de individuos).
B. Requiere, al menos, 5 años de seguimiento.
C. La exposición es poco frecuente.
D. Los estudios de cohorte son siempre ineficientes.

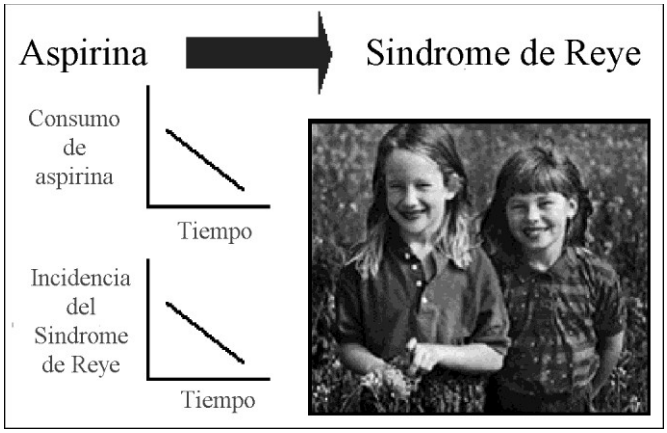
La respuesta es la C, porque el resultado es poco frecuente. Por consiguiente, se optó por un estudio de casos y controles puesto que, como dicho estudio podía completarse en un período más corto, permitía hallar una cantidad suficiente de casos y aun así requería menos individuos que un estudio de cohorte.

La investigación original del síndrome de Reye que identificó a la aspirina como un factor de riesgo fue un estudio de casos y controles realizado en Michigan de 1979 a 1980. Este estudio comprendía 25 casos y 46 controles. Los controles eran niños del mismo colegio: (1) ausentes del colegio; (2) que estaban en un grado similar; (3) que tenían diversos factores (p. ej., raza, año de nacimiento) en común; y (4) el mismo tipo de enfermedad antecedente. En 1982 se realizó un estudio de mayor envergadura para confirmar o refutar el resultado preliminar. Los investigadores utilizaron un sistema de vigilancia a nivel estatal para identificar todos los casos de síndrome de Reye en Ohio. En este estudio se utilizaron casos incidentes; es decir, casos de aparición reciente. Los controles de base poblacional se seleccionaron mediante la identificación y el posterior muestreo de individuos en la comunidad de todo el estado que habían padecido enfermedades virales similares a las informadas por los casos, pero que no habían desarrollado el síndrome de Reye. A los padres de los casos y los controles se les preguntó sobre los medicamentos que tomaron sus hijos durante la enfermedad.

Otro estudio publicado en 1987 seleccionó casos de niños con síndrome de Reye internados en cualquiera de los hospitales de atención terciaria que formaban parte de un grupo previamente seleccionado. Los controles de base hospitalaria se seleccionaron a partir de niños internados en estos mismos hospitales a causa de una enfermedad viral, pero que no habían desarrollado el síndrome de Reye. Se entrevistó a los padres para determinar el consumo previo de aspirina.

Como consecuencia de esta investigación de casos y controles sobre la relación entre el consumo de

aspirina y el síndrome de Reye, los profesionales de la salud recomendaron que ***no*** se utilizara aspirina para tratar los síntomas de enfermedades virales en niños. Más tarde, conforme se redujo el consumo infantil de aspirina, también se redujo la cantidad de casos de síndrome de Reye.



Resumen

- En las décadas de los setenta y ochenta se utilizaron estudios de casos y controles para determinar si el consumo de aspirina para el tratamiento de enfermedades virales era un determinante del síndrome de Reye.
- Estos estudios comenzaron con individuos que padecían esta enfermedad (el síndrome de Reye) e individuos similares que no la padecían.
- A los padres de los casos y los controles se les preguntó por el uso de medicamentos del hijo a lo largo de un período similar antes de que el niño experimentara los primeros síntomas del síndrome de Reye.
- Los profesionales de la salud recomendaron que no se utilizara aspirina para tratar síntomas de enfermedades virales en niños.
- Conforme se redujo el consumo infantil de aspirina, también se redujo la cantidad de casos del síndrome de Reye.

Estudios de casos y controles: ejemplo de la enfermedad de Creutzfeldt-Jakob (ECJ)

La enfermedad de Creutzfeldt-Jakob (ECJ) es una afección poco frecuente que se caracteriza por la demencia progresiva de evolución rápida. En la década de los noventa se descubrió en Europa una nueva variante de la ECJ en humanos después de una epidemia vacuna del mal de la vaca loca, la forma animal de la ECJ. Más tarde, la Unión Europea organizó un estudio para determinar si la inclusión de productos animales en la dieta era un factor de riesgo para el desarrollo de la ECJ.

Debido a que la ECJ es una enfermedad muy poco frecuente con un período de latencia prolongado, los investigadores optaron por emplear un diseño de estudio de casos y controles. Recabaron información sobre 405 casos de ECJ que se habían producido en la Unión Europea. Luego, recurrieron a los hospitales donde se había diagnosticado la ECJ a los pacientes para reclutar una cantidad equivalente de controles. A raíz del deterioro mental de los pacientes que padecían la enfermedad, la información sobre la alimentación de los casos debió recabarse entrevistando a uno de los parientes de cada caso.

¿De qué manera cree que se recabó información sobre la alimentación de los individuos de control? Si bien los participantes seleccionados a modo de control eran perfectamente capaces de brindar información sobre su propia alimentación, los investigadores optaron por entrevistar a un pariente de cada uno de ellos. De esta manera, la calidad de la información sobre la alimentación, tanto de los casos como de los controles, era similar y se reducía el sesgo de información.

Recuerde que una de las ventajas de un estudio de casos y controles es la posibilidad de evaluar el efecto de varias exposiciones diferentes. En este estudio, los investigadores examinaron por separado el riesgo de ECJ que representaba el consumo de embutidos, carne cruda, pescado crudo, productos derivados de la sangre animal, leche, queso y otros productos animales específicos. Puesto que ninguno de estos productos alimenticios demostró un incremento significativo en el riesgo de ECJ, los investigadores llegaron a la conclusión de que es poco probable que la ECJ se transmita de animales a humanos a través de productos animales.



Auto-evaluación (P3.6)

Para cada una de las siguientes características, seleccione la opción que aplique a los *estudios de casos y controles*:

1. Rol del investigador con respecto a la exposición: . . **???**
   a. asignar b. observar

2. Distribución de individuos en grupos: . . . . **???**
   a. selección personal b. aleatorización

3. Direccionalidad: . . . . . **???**
   a. hacia atrás b. hacia adelante c. no direccional

4. Secuencia temporal: . . . . . **???**
   a. prospectiva b. retrospectiva c. indistinta

5. Análisis según: . . . . . **???**
   a. asignación original b. experiencia real

La auto-evaluación continúa en la página siguiente

Para cada una de las siguientes características (puntos fuertes o débiles) de un estudio, seleccione el tipo de estudio que tenga dicha característica:

6. Más económico: . . . . . . **???**
7. Más rápido: . . . . **???**
8. La información sobre la exposición es más precisa: . . **???**
9. Adecuado para el estudio de exposiciones poco frecuentes: . **???**
10. Adecuado para el estudio de enfermedades poco frecuentes: . **???**
11. Permite estudiar varios resultados: . . . **???**
12. Requiere una muestra más pequeña: . . . . **???**
13. Permite calcular el riesgo: . . . . **???**

Opciones
De casos y controles **De cohorte prospectivo**

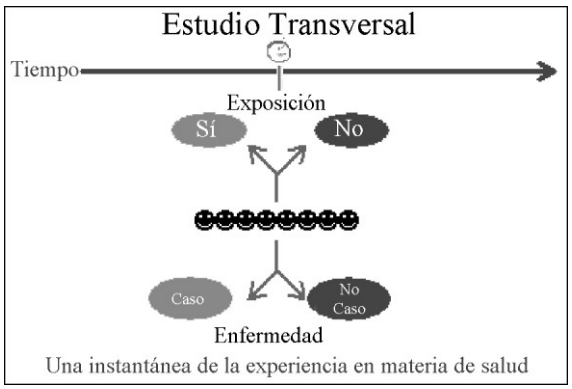
Determine si cada una de las siguientes afirmaciones es **verdadera** o **falsa**:

14. Lo ideal es seleccionar los controles de la misma población que produjo los casos. . **???**
15. Lo ideal es seleccionar los controles de pacientes internados en hospitales. . . **???**
16. Los controles de base poblacional incluyen únicamente vecinos y personas identificadas mediante el discado de números de teléfono elegidos al azar. . . . . . **???**

Estudios transversales

El estudio transversal es un diseño observacional básico en el que todas las variables se observan o miden en un único momento determinado. De todos los diseños de estudio observacionales, suele ser el más rápido y económico aunque también es muy propenso a sesgos si se compara con un diseño de cohorte.

En un estudio transversal, el muestreo de individuos se realiza en un momento establecido o durante un período breve. Todos los individuos participantes se examinan, observan e interrogan con respecto a su condición de enfermedad, sus exposiciones presentes o pasadas, al igual que otras variables relevantes. Un estudio transversal proporciona una instantánea de la experiencia en materia de salud que tiene una población en un momento específico y, por lo tanto, se utiliza para describir patrones de aparición de enfermedades. Una muestra transversal suele ser más representativa de la población general bajo estudio que en el caso de otros diseños de estudio. Los estudios transversales ofrecen una forma práctica y económica de observar las relaciones entre varias exposiciones y diversas enfermedades. Si la enfermedad de interés es relativamente común y tiene una duración prolongada, el estudio transversal permite obtener una cantidad suficiente de casos para generar hipótesis sobre las relaciones entre una exposición y la enfermedad. Para confirmar o refutar estas hipótesis, se utilizan otros tipos de estudios más costosos tales como los estudios de cohorte y los ensayos clínicos.



Una instantánea de la experiencia en materia de salud

Los estudios transversales tienen algunas desventajas. Por ejemplo, un estudio de esta naturaleza sólo puede identificar casos *existentes* o *prevalentes* en un momento determinado, y no casos *nuevos* o *incidentes* a lo largo de un período de seguimiento. Por consiguiente, un estudio transversal no permite determinar si la exposición es anterior a la enfermedad o si la enfermedad influyó en la exposición.

Puesto que sólo se admiten casos existentes, el estudio transversal incluye únicamente los casos que han sobrevivido lo suficiente como para disponer de ellos. Por ende, la exclusión de los que no sobrevivieron podría dar lugar a una conclusión errónea sobre la relación entre una exposición y la enfermedad (consulte la nota sobre este problema que se encuentra al final de esta actividad).

Las enfermedades de corta duración tales como un simple resfrío o gripe, en especial, las que se producen en una determinada estación, pueden tener una representación insuficiente en un estudio transversal que analiza la presencia de dicha afección en un momento específico.

Resumen: estudios transversales

- El muestreo de individuos se realiza en un momento determinado o durante un período breve: una instantánea.

Ventajas

- Práctico y económico.
- Permite considerar varias exposiciones y diversas enfermedades.
- Permite generar hipótesis.
- Suele ser representativo de la población en general.

Desventajas

- No permite determinar si la exposición es anterior a la enfermedad o si la enfermedad influyó en la exposición.
- Posible sesgo, puesto que los únicos individuos disponibles para el estudio son los sobrevivientes.
- La representación de enfermedades de corta duración podría ser insuficiente.

¿Qué sesgo conllevan los sobrevivientes en un estudio transversal?

Un estudio transversal puede estar sesgado porque sólo dispone de los individuos que han sobrevivido lo suficiente como para su inclusión en el estudio. Para ilustrar este punto, supongamos que todos los que padecen una enfermedad determinada y no realizan actividades físicas intensas en forma periódica mueren al poco tiempo. Supongamos también que los que padecen la enfermedad, pero realizan actividades físicas intensas en forma periódica, sobreviven por varios años.

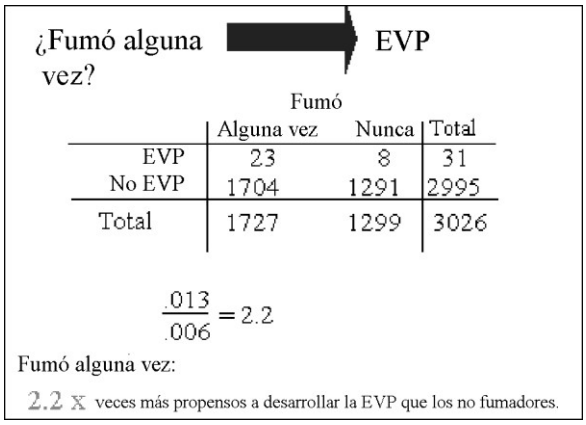
Ahora pensemos en un estudio transversal que tenga como objetivo determinar si existe alguna relación entre una rutina de ejercicios físicos intensos y la enfermedad. Como este tipo de estudio incluiría únicamente a los sobrevivientes, sería probable que encontráramos una baja proporción de casos entre las personas que no realizan ejercicios físicos intensos. Por el contrario, sería probable que encontráramos una proporción relativamente más alta de casos entre las personas que sí realizan ejercicios físicos intensos. Esto llevaría a la conclusión de que los ejercicios físicos intensos son perjudiciales para la salud, aun cuando, de hecho, la protegieran.

Ejemplo de estudio transversal: enfermedad vascular periférica, Escocia

Un estudio de 1991 examinó una muestra de 5.000 hombres escoceses para evaluar la presencia de la enfermedad vascular periférica (EVP). Durante el examen, también se determinaron otras características para cada individuo, incluso el hecho de que hubiesen fumado o no alguna vez.

El estudio era transversal porque todos los individuos se seleccionaron y observaron en un momento determinado. Si bien se realizaron exámenes físicos, la inversión de tiempo y dinero fue mucho menor que la que habría requerido el seguimiento de individuos sin la enfermedad a lo largo del tiempo para determinar su posterior condición de EVP. La muestra era representativa de la población masculina escocesa.

El estudio determinó que el 1,3% de los 1727 hombres que habían fumado alguna vez padecían la EVP, mientras que sólo el 0,6% de los 1299 hombres que nunca habían fumado la padecían. Si dividimos 0,013 por 0,006, vemos que los hombres que habían fumado alguna vez eran 2,2 veces más propensos a desarrollar la EVP que los que jamás lo habían hecho.



Estos resultados sugirieron que el tabaquismo puede contribuir al desarrollo de la EVP. Aun así, los resultados son sólo una instantánea de los individuos en un momento determinado de 1991. A los individuos sin EVP no se les ha hecho un seguimiento a lo largo del tiempo. Entonces, ¿cómo sabemos con esta instantánea si la EVP produce tabaquismo o si el tabaquismo produce la EVP? Esto ilustra uno de los problemas de los estudios transversales: son siempre **no-direccionales**. Además, las personas que murieron de EVP antes de la selección de individuos no se incluyen en el estudio. Por lo tanto, los resultados del estudio podrían estar sesgados porque sólo tienen en cuenta a los hombres que sobrevivieron a la EVP.

Resumen

- Un ejemplo de estudio transversal es un estudio de la enfermedad vascular periférica (EVP) realizado en Escocia en 1991.
- Los resultados demostraron que los hombres que habían fumado alguna vez eran 2,2 veces más propensos a desarrollar la EVP que los que jamás lo habían hecho.
- Este estudio resultó mucho más económico y rápido que un estudio de cohorte.
- El estudio no permite determinar si la EVP produce tabaquismo o si el tabaquismo produce la EVP.
- Los resultados del estudio podrían estar sesgados porque sólo tienen en cuenta a los que sobrevivieron a la EVP.



Auto-evaluación (P3.7)

Para cada una de las siguientes características, seleccione la opción que aplique a los *estudios transversales*:

1. Rol del investigador con respecto a la exposición: . . **???**
   a. asignar b. observar

2. Distribución de individuos en grupos: . . . . **???**
   a. selección personal b. aleatorización

3. Direccionalidad: . . . . . **???**
   a. hacia atrás b. hacia adelante c. no direccional

4. Secuencia temporal: . . . . . **???**
   a. prospectiva b. retrospectiva c. indistinta

Determine si cada una de las siguientes afirmaciones es **verdadera** o **falsa**:

5. Los estudios transversales son más adecuados para generar hipótesis sobre las relaciones entre exposiciones y enfermedades que para comprobar dichas relaciones. . . . **???**

6. Como la exposición y la enfermedad se evalúan al mismo tiempo, los estudios transversales no están sujetos al sesgo de supervivencia. . . . . **???**

7. Como la exposición y la enfermedad se evalúan al mismo tiempo, los estudios transversales podrían ser incapaces de determinar si la exposición es anterior al inicio del proceso de la enfermedad. **???**

8. Los estudios transversales permiten examinar varias exposiciones y diversas enfermedades. **???**

3-4 Diseños híbridos

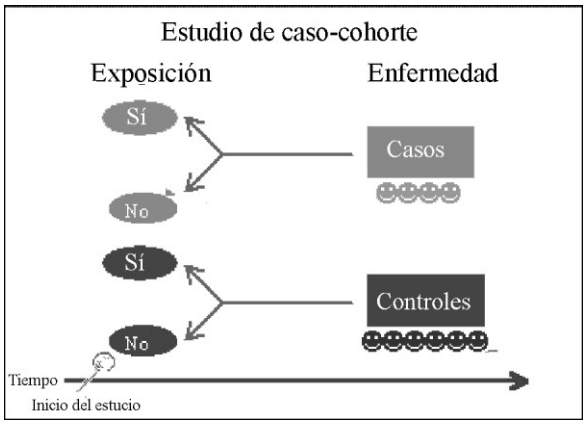
Diseños híbridos

Los diseños híbridos combinan los elementos de, al menos, dos diseños básicos o extienden la estrategia de un único diseño básico a través de la repetición. Dos de los diseños híbridos más difundidos son el ***estudio de caso-cohorte*** *y el* ***estudio de casos y controles anidados****. Ambos diseños combinan elementos de un estudio de cohorte y un estudio de casos y controles. Aquí también se describe un diseño híbrido de desarrollo más reciente, denominado diseño de* ***casos cruzados****.*

Estudio de caso-cohorte

Debido al método de selección de los controles, el estudio de caso-cohorte es un diseño híbrido menos propenso a sesgos que el diseño de casos y controles convencional.

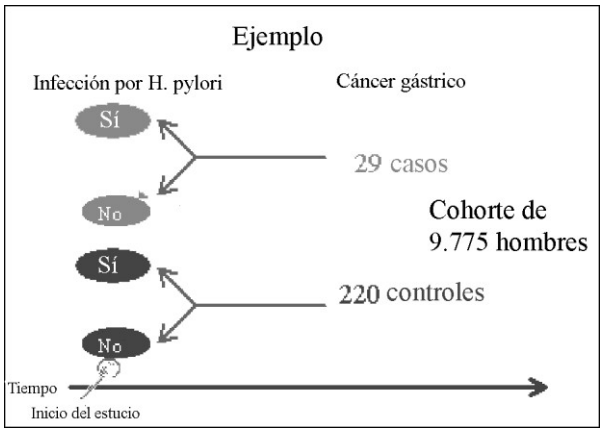
Un estudio de caso-cohorte utiliza un diseño híbrido que combina elementos de un estudio de cohorte y un estudio de casos y controles. A una población de caso-cohorte se le hace un seguimiento a lo largo del tiempo para identificar casos nuevos o incidentes de una enfermedad. El grupo de control consta de una muestra aleatoria de individuos obtenida a partir de la cohorte original de personas que no padecen la enfermedad. Nótese que esto permite muestrear como individuos de control a los individuos que, con el tiempo, se convierten en casos. Luego, se determina la condición de exposición previa tanto de los casos como de los controles.



Preguntas de estudio (P3.8)

1. ¿Cuál es la direccionalidad de este estudio? ¿**Hacia adelante** o **hacia atrás**?
2. ¿Cuál es la secuencia temporal de este estudio? ¿**Prospectiva** o **retrospectiva**?

Un ejemplo de esto es un estudio de 1995 que analizó los factores de riesgo de cáncer gástrico con una cohorte de 9.775 hombres en Taiwán, de los cuales se obtuvieron y congelaron muestras de sangre al momento de reclutarlos para el estudio. Durante el seguimiento posterior, que se basaba en datos de registros del cáncer, se identificaron 29 casos de cáncer gástrico. Se obtuvo un grupo de control de 220 individuos provenientes de la cohorte original completa. Es posible que algunos de estos controles incluso hayan desarrollado cáncer gástrico a lo largo del seguimiento. Una variable de exposición de interés era la presencia o ausencia de una infección por Helicobacter pylori, lo cual podía determinarse descongelando y analizando las muestras de sangre de los casos y los controles. Esto permitió ahorrar mucho tiempo y dinero en trabajo de laboratorio para determinar la condición de exposición, ya que no hacía falta considerar toda la cohorte.



En general, un diseño de caso-cohorte brinda la seguridad de que los controles provienen de la población fuente en la que se desarrollaron los casos. Además, puesto que los casos se determinan durante el seguimiento, se pueden estudiar varias enfermedades, lo cual no siempre es posible en los estudios de casos y controles típicos. Al mismo tiempo, un estudio de caso-cohorte es más eficiente que un estudio de cohorte en términos de costo y tiempo porque la cantidad de no casos que se observa es mucho menor. Una última ventaja de un estudio de caso-cohorte es que, a diferencia de los estudios de casos y controles más tradicionales, permite calcular una razón de riesgos (esto se describe en la Lección 5 en el análisis de las medidas de efecto para un estudio de casos y controles).

Sin embargo, un diseño de caso-cohorte es más propenso a errores de medición que un estudio de cohorte si la condición de exposición se determina de manera retrospectiva después de seleccionar los casos y controles. Este tipo de estudio puede requerir una inversión de tiempo y dinero mucho mayor que un estudio de casos y controles, puesto que el diseño tradicional no requiere la identificación de una cohorte original para la selección de controles.

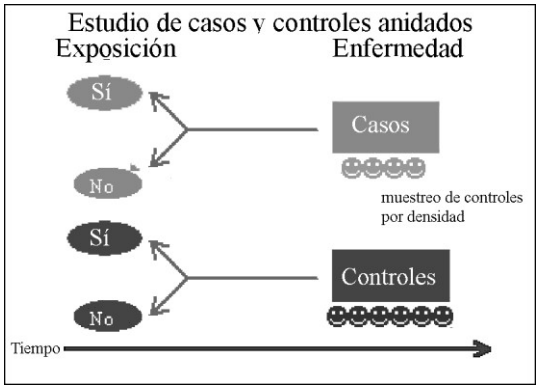
Resumen

- El estudio de caso-cohorte utiliza un diseño híbrido que combina elementos de un estudio de cohorte y un estudio de casos y controles.
- En un diseño de caso-cohorte, el muestreo de los controles se realiza a partir de la cohorte original.
- Los casos de una enfermedad son nuevos o incidentes.
- Los controles se seleccionan de la población fuente que proporciona los casos.
- A diferencia de un estudio de casos y controles, permite estudiar varias enfermedades.
- La cantidad de no casos es menor que la de un estudio de cohorte.
- Es más propenso a errores de medición que un estudio de cohorte.
- Es más costoso que un estudio de casos y controles.

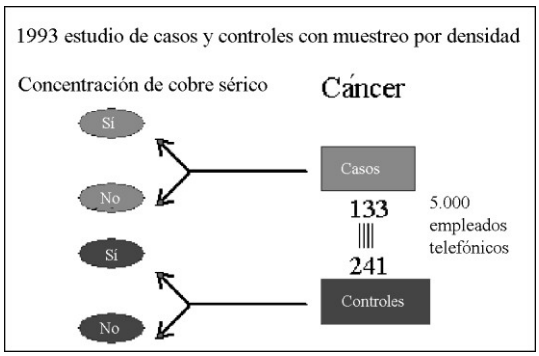
Estudio de casos y controles anidados

El **diseño de casos y controles anidados**, también denominado **estudio de casos y controles con muestreo por densidad**, es una variación del estudio de caso-cohorte. Este tipo de estudio puede utilizarse si se conoce el momento en que los individuos se convirtieron en casos. En este diseño, los controles se equiparan con los casos *al momento del diagnóstico de los casos*. Se seleccionan uno o más controles para cada caso de entre los individuos pertenecientes a la cohorte original que siguen en riesgo al momento en que se identifica un caso. Esta selección a menudo se denomina **muestreo de controles por densidad**. Por supuesto, los controles para un caso determinado pueden convertirse en casos después de que seleccionarlos como controles. Una ventaja de emplear el diseño de casos y controles anidados en lugar del diseño de caso-cohorte es que los controles por muestreo de densidad estuvieron en riesgo de convertirse en casos durante el mismo tiempo que el caso con el que se les

equipara.



En 1993 se realizó un estudio de casos y controles con muestreo por densidad para evaluar la relación entre el riesgo de desarrollar cáncer y la concentración de cobre sérico. Los investigadores obtuvieron muestras basales de sangre e información sobre los factores de riesgo de 5.000 empleados telefónicos. Un sistema de vigilancia de cáncer identificó 133 casos de cáncer que se desarrollaron de esta cohorte. Luego, se determinó el momento de diagnóstico de los casos para realizar un muestreo por densidad de 241 controles, los cuales se compararon con los casos con respecto a la concentración de cobre sérico y otras covariables de interés.



Pregunta: en este ejemplo, la cantidad de controles es mucho mayor que la cantidad de casos. ¿Cómo es posible esto mediante el uso de un muestreo por densidad? Para algunos casos se seleccionaron dos o más controles.

Al igual que un estudio de caso-cohorte, el diseño de casos y controles anidados brinda la seguridad de que los controles provienen de la población fuente en la que se desarrollaron los casos. Al mismo tiempo, el estudio de casos y controles anidados es más eficiente que un estudio de cohorte en términos de costo y tiempo porque la cantidad de no casos que se observa es mucho menor. Otra ventaja del estudio de casos y controles anidados es que, a diferencia de los estudios de caso-cohorte o los estudios de casos y controles más tradicionales, permite calcular una razón de ***tasas*** (esto se describe en la lección 5 en el análisis de las medidas de efecto para un estudio de casos y controles).

Al igual que el estudio de caso-cohorte, el estudio de casos y controles anidados es más propenso a errores de medición que un estudio de cohorte si la condición de exposición se determina de manera retrospectiva después de seleccionar los casos y controles. Aun así, esto no representa ningún problema si la información sobre la exposición se obtiene al momento de la inclusión de los individuos en la cohorte inicial, tal como sucede en el estudio de casos y controles anidados que se acaba de describir.

Resumen

- El diseño de casos y controles anidados es una variación del diseño de caso-cohorte en el que los controles se seleccionan mediante un muestreo por densidad.
- Para realizar el muestreo por densidad, los controles deben equipararse con los casos al momento del diagnóstico de los casos.
- Ventaja: los controles por muestreo de densidad están en riesgo durante el mismo tiempo que el caso con el que se les equipara.
- Desventaja: es más propenso a errores de medición que un estudio de cohorte.



Auto-evaluación (P3.9)

Indique si cada una de las siguientes características corresponde a un **estudio de caso-cohorte**, a un **estudio de casos y controles anidados**, a **ambos** o a **ninguno**.

1. Suele ser más económico que un estudio de cohorte prospectivo: . . **???**
2. El grupo comparativo proviene de la misma población que los casos: . **???**
3. Suele ser más económico que un estudio de casos y controles: . . **???**
4. La secuencia temporal es retrospectiva: . . . . **???**
5. La direccionalidad es hacia atrás: . . . . . **???**
6. Los controles se equiparan con los casos al momento del diagnóstico: . **???**
7. Compara la experiencia de exposición de los casos con la de los controles: **???**
8. Utiliza un muestreo por densidad: . . . . . **???**
9. Debe tener la misma cantidad de controles que de casos: . . **???**
10. Supongamos que, al momento de alistarse en el servicio militar, se obtuvo y almacenó una muestra de sangre de cada recluta. Después de la Guerra del Golfo Pérsico, algunos soldados desarrollaron una serie de síntomas que se conocieron con el nombre de síndrome de la Guerra del Golfo. Si los investigadores luego analizaran la sangre de todos los soldados que padecen el síndrome y la sangre del doble de soldados que no lo desarrollaron, esto sería un ejemplo de: **???**

Opciones
estudio de caso-cohorte **estudio de casos y controles anidados** **estudio de cohorte prospectivo** **estudio de cohorte retrospectivo**

Diseño de casos cruzados

El estudio de casos cruzados es una variante del estudio de casos y controles equiparados (consulte la sección sobre Equiparamiento en la lección 15) diseñado para reducir las probabilidades de sesgo del diseño de casos y controles convencional mediante el método de selección de los controles. El diseño incorpora elementos tanto de un estudio de casos y controles equiparados como de un estudio cruzado retrospectivo no experimental (nota: en un diseño cruzado, cada individuo recibe, al menos, dos exposiciones/tratamientos diferentes en distintas ocasiones). El aspecto fundamental del diseño de casos cruzados es que cada caso sirve como su propio control. Las exposiciones que varían en términos de tiempo se comparan entre los intervalos en que se produjo el resultado (intervalos de casos) y los intervalos en que no se produjo el resultado en el mismo individuo.

El diseño de casos cruzados se creó para evaluar el efecto de exposiciones breves con efectos transitorios en resultados de salud agudos cuando no se dispone de un grupo de control tradicional. La principal ventaja del diseño de casos cruzados radica en su capacidad para reducir el sesgo de confusión de los controles. Los individuos que se equiparan consigo mismos automáticamente eliminan la confusión entre individuos y la confusión de las covariables fijas medidas y no medidas.

Un ejemplo de diseño de casos cruzados es el estudio de Redelmeier y Tibshirani, quienes intentaron determinar si el uso del teléfono celular al conducir un vehículo motorizado incrementa el riesgo de sufrir un accidente de tránsito. Dice el resumen del estudio: "Estudiamos a 699 conductores que tenían teléfonos celulares y participaron en accidentes de tránsito que produjeron daños materiales importantes aunque no lesiones personales. Mediante el uso de registros de facturación detallados, se analizaron las llamadas por celular de cada persona correspondientes al día del accidente y a la semana anterior. El riesgo de sufrir un accidente cuando se empleó el celular resultó cuatro veces mayor que cuando no se utilizó (riesgo relativo de 4,3; intervalo de confianza del 95%, de 3,0 a 6,5). Las llamadas realizadas cerca de la hora en que se produjo el accidente resultaron particularmente peligrosas (riesgo relativo de 4,8 para llamadas realizadas con una diferencia de 5 minutos con respecto al accidente)…".

Diseños incompletos

*Los **diseños incompletos** son estudios en los que falta información sobre uno o más factores relevantes.*

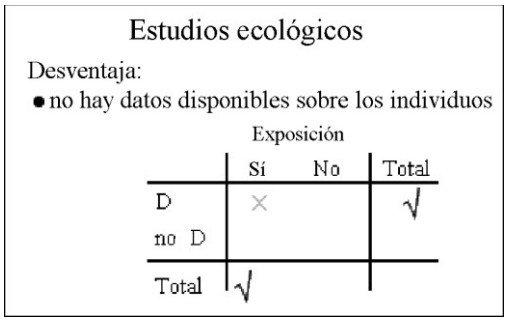
Estudios ecológicos

*Un **estudio ecológico** es un diseño incompleto cuya unidad de análisis es un grupo, a menudo definido en términos geográficos, como un sector censal, un estado o un país.*

En un estudio ecológico, la unidad de análisis es un *grupo*, a menudo definido en términos geográficos, en lugar de un individuo. Esto significa que los datos básicos suelen ser porcentajes u otros resúmenes estadísticos para cada grupo, en lugar de mediciones de características de individuos. Los grupos pueden ser sectores censales, estados o países.

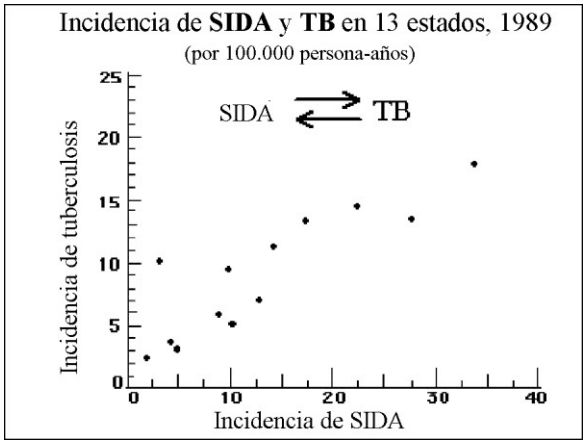
La ventaja de un estudio ecológico es que a menudo se puede realizar de un modo rápido y económico mediante el uso de datos existentes tales como datos de mortalidad. Los estudios ecológicos suelen utilizarse para generar hipótesis sobre relaciones entre exposiciones y enfermedades. También se utilizan para evaluar el impacto de los programas de intervención en la condición de salud de las poblaciones meta.

La principal crítica que se le puede hacer a un estudio ecológico es que no hay datos disponibles sobre los individuos. En particular, un estudio ecológico tiene información sobre la cantidad de personas expuestas y la cantidad de casos dentro de cada grupo, pero no incluye la cantidad de casos expuestos.



Por consiguiente, las conclusiones de los estudios ecológicos sobre los determinantes de un resultado de salud no pueden extenderse a los individuos. Este problema se denomina **falacia ecológica**.

El diagrama de dispersión que aparece a continuación muestra una comparación entre las tasas de incidencia del SIDA en 13 estados de EE. UU. durante 1989 y las tasas de incidencia de tuberculosis correspondientes al mismo año. El gráfico demuestra que los estados con una alta incidencia del SIDA también tenían una alta incidencia de tuberculosis. Además, indica que los estados con una baja incidencia del SIDA tendían a presentar una baja incidencia de tuberculosis. Por consiguiente, la relación que se ilustra aquí sugiere que el desarrollo del SIDA puede influir en el desarrollo de la tuberculosis o viceversa.



No obstante, estos datos muestran una relación que utiliza estados como unidad de análisis. Es posible que sean pocos los individuos que tengan tanto SIDA como tuberculosis, aun cuando las incidencias de tales enfermedades sean altas por separado. Dicha información sólo puede obtenerse de los datos sobre individuos.

Resumen

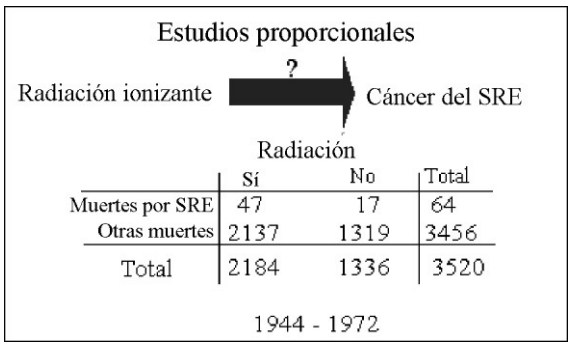
- En un estudio ecológico, la unidad de análisis es un grupo, a menudo definido en términos geográficos.
- Las conclusiones de los estudios ecológicos no pueden extenderse a los individuos (falacia ecológica).
- Un estudio ecológico dispone de información sobre la cantidad de personas expuestas y la cantidad de casos dentro de cada grupo, pero no sobre la cantidad de casos expuestos.

Estudios proporcionales

Un estudio de ***morbilidad*** **proporcional** o ***mortalidad*** **proporcional** sólo incluye observaciones sobre casos, pero carece de información sobre la población candidata en riesgo de desarrollar el resultado de salud. Si el diseño comprende casos incidentes, se trata de un estudio de morbilidad proporcional. Si se utilizan muertes, entonces es un estudio de mortalidad proporcional.

Los estudios de mortalidad proporcional se utilizan para generar hipótesis nuevas o para realizar pruebas preliminares de hipótesis etiológicas sin recabar una cantidad significativa de datos adicionales. Puesto que los estudios de morbilidad o mortalidad proporcionales carecen de no casos y dado que la población candidata en riesgo no está disponible, resulta imposible computar las medidas de efecto tradicionales tales como las razones de riesgo o los *odds ratios* que se utilizan para evaluar las relaciones entre exposiciones y enfermedades.

Un estudio de trabajadores de plantas nucleares realizado en 1997 analizó la posible relación entre niveles bajos de radiación ionizante y el cáncer a partir de la muerte de 3.500 empleados de una central, entre 1944 y 1972.



	Sí	No	Total
Muertes por SRE	47	17	64
Otras muertes	2137	1319	3456
Total	2184	1336	3520

1944 - 1972

Los datos de este estudio eran exclusivamente de mortalidad y no incluían información sobre la magnitud de la población candidata en riesgo de muerte. Sin esta información, era imposible demostrar una relación irrefutable entre la exposición y el riesgo de desarrollar cáncer del sistema reticuloendotelial (SRE). No obstante, estos datos sí demostraron que la proporción de muertes por cáncer del SRE había sido significativamente mayor entre los trabajadores expuestos a radiación en la planta que entre los no expuestos. Sin duda, esto fue un paso hacia la implicación del daño potencial de la exposición a la radiación en el riesgo de desarrollar cáncer del SRE.

Resumen

- Un estudio de morbilidad o mortalidad proporcional sólo incluye observaciones sobre casos, pero carece de información sobre la población candidata en riesgo.
- Los estudios proporcionales sirven para generar hipótesis.
- Los estudios proporcionales son útiles para realizar pruebas preliminares de hipótesis etiológicas.
- En los estudios proporcionales, no se pueden computar las medidas de efecto tradicionales tales como las razones de riesgo.

Auto-evaluación (P3.10)



Determine si cada una de las siguientes afirmaciones es **verdadera** o **falsa**:

1. El estudio ecológico es aquel en el que se comparan poblaciones con individuos. . . **???**

2. El análisis de la distribución de causas de muerte según se describe en el acta de defunción de todas las personas que murieron en Georgia en 1999 es un ejemplo de estudio de mortalidad proporcional. **???**

3. Tanto los estudios ecológicos como los estudios de mortalidad proporcional son más adecuados para generar hipótesis causales que para comprobar hipótesis. . . **???**

4. Tanto los estudios ecológicos como los estudios de mortalidad proporcional tienen información sobre poblaciones, pero no sobre individuos. . . . . . . **???**

5. Los estudios ecológicos y los estudios de mortalidad proporcional a menudo pueden realizarse en poco tiempo porque los dos utilizan mayormente datos existentes fáciles de conseguir. . **???**

6. Ni los estudios ecológicos ni los estudios de morbilidad/mortalidad proporcionales permiten calcular el riesgo que corre un individuo de enfermarse o morir. . . . . **???**

Referencias

Referencias sobre eficacia, efectividad y eficiencia

Detsky, A S. *Evidence of effectiveness: evaluating its quality*. En F. A. Sloan (Ed.), Valuing health care: costs, benefits, and effectiveness of pharmaceuticals and other medical technologies (pp. 15-29). Cambridge, Reino Unido: Cambridge University Press, 1995.

Detsky AS, Naglie IG. *A Clinician Guide to Cost-Effectiveness Analysis*. Ann Intern Medicine 1990; 113(2), 147-54.

Grimes DA, Schulz KF. An overview of clinical research: the lay of the land. Lancet 2002; 359(9300), 57-61.

Kocher MS, Henley MB. *It is money that matters: Decision analysis and cost-effectiveness analysis*. Clinical Orthopaedics and Related Research 2003; (413), 106-16.

Referencias sobre el estudio del Grupo de investigación sobre la diabetes (ensayo clínico)

The Diabetes Control and Complications Trial Research Group. *The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus*. N Engl J Med 1993; 329(14):977-86.

Referencias sobre el estudio de Framingham (estudio de cohorte prospectivo)

Feinleib M, *The Framingham study: sample selection, follow-up, and methods of analysis*, en la monografía N.° 67 del National Cancer Institute, Greenwald P (editor), US Department of Health and Human Services, 1985.

Dorgan JF, Brown C, Barrett M, et al. *Physical activity and risk of breast cancer in the Framingham Heart Study*, Am J Epidemiol 1994; 139(7): 662-9.

Margolis JR, Gillum RF, Feinleib M, Brasch RC, Fabsitz RR. *Community surveillance for coronary heart disease: the Framingham Cardiovascular Disease Survey. Methods and preliminary results*. Am J Epidemiol 1974; 100(6):425-36.

Referencias sobre el uso de VDT y los abortos espontáneos (estudio de cohorte retrospectivo)

Schnorr TM, Grajewski BA, Hornung RW, Thun MJ, Egeland GM, Murray WE, Conover DL, Halperin WE. *Video display terminals and the risk of spontaneous abortion*. N Engl J Med 1991; 324(11):727-33.

Referencias sobre los medicamentos antiinflamatorios no esteroideos y el mal de Alzheimer (estudio de cohorte prospectivo)

in t' Veld BA, Ruitenberg A, Hofman A, Launer LJ, van Duijn CM, Stijnen T, Breteler MM, Stricker BH. *Nonsteroidal antiinflammatory drugs and the risk of Alzheimer's disease*. N Engl J Med 2001;

345(21):1515-21.

Referencias sobre el síndrome de Reye (estudios de casos y controles)

Waldman RJ, Hall WN, McGee H, Van Amburg G. *Aspirin as a risk factor in Reye's syndrome*. JAMA 1982; 247(22):3089-94.

Halpin TJ, Holtzhauer FJ, Campbell RJ, Hall LJ, Correa-Villasenor A, Lanese R, Rice J, Hurwitz ES. *Reye's syndrome and medication use*. JAMA 1982; 248(6):687-91.

Daniels SR, Greenberg RS, Ibrahim MA. *Scientific uncertainties in the studies of salicylate use and Reye's syndrome*. JAMA 1983; 249(10):1311-6.

Hurwitz ES, Barrett MJ, Bregman D, Gunn WJ, Pinsky P, Schonberger LB, Drage JS, Kaslow RA, Burlington DB, Quinnan GV, et al. *Public Health Service study of Reye's syndrome and medications*. Informe del estudio principal. JAMA 1987; 257(14):1905-11.

Forsyth BW, Horwitz RI, Acampora D, Shapiro ED, Viscoli CM, Feinstein AR, Henner R, Holabird NB, Jones BA, Karabelas AD, et al. *New epidemiologic evidence confirming that bias does not explain the aspirin/Reye's syndrome association*. JAMA 1989; 261(17):2517-24

Referencias sobre la enfermedad de Creutzfeldt-Jakob (estudios de casos y controles)

Van Duijn CM, Delasnerie-Laupretre N, Masullo C, Zerr I, de Silva R, Wientjens DP, Brandel JP, Weber T, Bonavita V, Zeidler M, Alperovitch A, Poser S, Granieri E, Hofman A, Will RG. *Case-control study of risk factors of Creutzfeldt-Jakob disease in Europe during 1993-95*. European Union (EU) Collaborative Study Group of Creutzfeldt-Jakob disease (CJD). Lancet 1998; 351(9109):1081-5.

Will RG, Ironside JW, Zeidler M, Cousens SN, Estibeiro K, Alperovitch A, Poser S, Pocchiari M, Hofman A, Smith PG. *A new variant of Creutzfeldt-Jakob disease in the UK*. Lancet 1996; 347(9006):921-5.

Diseño epidemiológico general

Checkoway H, Pearce N, Dement JM. *Design and conduct of occupational epidemiology studies: II. Analysis of cohort data*. Am J Ind Med 1989; (15(4):375-94.

Greenberg RS, Daniels SR, Flanders WD, Eley JW, Boring JR. *Medical Epidemiology* (3.° edición). Lange Medical Books, Nueva York, 2001.

Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic Research: Principles and Quantitative Methods*. John Wiley and Sons Publishers, Nueva York, 1982.

Steenland K (ed.). *Case studies in occupational epidemiology*. Oxford University Press, Nueva York, 1993.

Ejemplo de estudios transversales

Smith WCS, Woodward M, Tunstall-Pedoe H. *Intermittent claudication in Scotland, in Epidemiology of Peripheral Vascular Disease*. (ed FGR Fowkes.), Springer-Verlag, Berlín, 1991.

Diseños híbridos

Coates RJ, Weiss NS, Daling JR, Rettmer RL, Warnick GR. *Cancer risk in relation to serum copper levels*. Cancer Res 1989; 49(15): 4353-6.

Linn JT, Wang LY, Wang JT, Wang TH, Yang CS, Chen CJ. *A nested case-control study on the association between Helicobacter pylori infection and gastric cancer risk in a cohort of 9775 men in Taiwan*. Anticancer Res 1995; 15:603-6.

Maclure M, Mittleman MA. *Should we use a case-crossover design?* Annu Rev Public Health 2000; 21:193-221.

Maclure M. *The case-crossover design: a method for studying transient effects on the risk of acute events*. Am J Epidemiol 1991; 133(2):144-53.

Redelmeier DA, Tibshirani RJ. *Association between cellular-telephone calls and motor vehicle collisions*. N Eng J Med 1997; 336(7):453-8.

Ecológico

CDC. *Summary of notifiable diseases in the United States*. MMWR Morb Mortal Wkly Rep 1990; 38(54):1-59.

Morgenstern H. *Ecologic studies in epidemiology: concepts, principles, and methods*. Annu Rev Public Health 1995; 16:61-81.

Mortalidad proporcional

Mancuso TF, Stewart A, Kneale G. *Radiation exposures of Hanford workers dying from cancer and other causes*. Health Phys 1977; 33:369-85.

Causalidad y metaanálisis

Blalock HM, ***Causal Inferences in Nonexperimental Research***, Capítulo 1, Norton Publishing, 1964.

Chalmers I, Altman DG (editores), ***Systematic Reviews***, BMJ Publishing Group, Londres, 1995.
Chalmers TC. *Problems induced by meta-analyses*. Stat Med 1991; 10(6):971-80.
Hill AB, *Principles of Medical Statistics*, 9.° edición, Capítulo 24, Oxford University Press, 1971.
Lipsey MW, Wilson DB. *Practical meta-analysis*. Applied Social Research Methods Series; Vol. 49. Sage Publications, Inc., Thousand Oaks, CA: 2001.
Mosteller F, Colditz GA. *Understanding research synthesis (meta-analysis)*. Annu Rev Public Health 1996; 17:1-23.
Petitti DB, *Meta-analysis Decision Analysis and Cost-Effectiveness Analysis; Methods for Quantitative Synthesis in Medicine*, Oxford University Press, 1994.
Popper KR, *The Logic of Scientific Discovery*, Harper and Row Publishers, 1968.
Rothman KJ. *Causes*. Am J Epidemiol 1976; 104(6): 587-92.
Susser M, *Causal Thinking in the Health Sciences*, Oxford University Press, 1973.
U.S. Department of Health, Education, and Welfare, *Smoking and Health*, PHS Publ. No. 1103, Government Printing, Washington DC, 1964.
Weiss NS. *Inferring causal relationships: elaboration of the criterion of 'dose-response'*. Am J Epidemiol 1981; 113(5):487-90.

Ejercicios

EPI-1. Tipo de estudio. Indique el tipo de estudio que se describe en cada uno de los siguientes párrafos:

a. Con el objeto de investigar la relación entre el consumo de huevo y las enfermedades cardiacas se interrogó a un grupo de pacientes con infarto de miocardio internados en un hospital para determinar su consumo de huevo. Luego, se utilizó un protocolo idéntico para determinar el consumo de huevo de un grupo de pacientes internados en una clínica por fracturas, que fueron equiparados en edad y sexo con el primer grupo.
b. A fin de investigar la relación entre ciertos solventes y el cáncer, se interrogó a todos los empleados de una fábrica sobre su exposición a un solvente industrial y se midió la cantidad y la duración de las exposiciones. Se realizó un seguimiento periódico de estos individuos y, 10 años más tarde, se obtuvo una copia del acta de defunción de todos los que habían muerto.
c. Se realizó una encuesta a todas las enfermeras que trabajaban en un hospital determinado. Además de otras preguntas, el cuestionario solicitaba información sobre la categoría de la enfermera y el nivel de satisfacción con sus perspectivas laborales.
d. A fin de evaluar una nueva escuela (es decir, un enfoque) para el tratamiento del dolor lumbar, los pacientes que padecían dolores en la región inferior de la espalda fueron asignados de manera aleatoria a la nueva escuela o a una terapia ocupacional convencional. Después de 3 meses, se les preguntó por el dolor lumbar y se les examinó levantando peso mediante monitores independientes.
e. En la unidad local de accidentes y emergencias se implementó un nuevo sistema de triage. Para evaluar este sistema, se midieron los tiempos de espera de los pacientes durante 6 meses y se compararon con los tiempos de espera en un período cercano equivalente.
f. El Registro del Tumor de un estado particular de EE. UU. se utilizó para identificar todos los casos primarios de cáncer de vejiga que se habían producido en el estado durante un período determinado. Estos casos se compararon con una muestra de no casos del mismo estado que se equipararon en edad y momento del diagnóstico. Todos los individuos o familiares sobrevivientes de ambos grupos fueron entrevistados para recabar información sobre el consumo de sacarina y otros factores de riesgo conocidos del cáncer de vejiga.
g. En el estudio de trabajadores nucleares de Hanford (1977), se registraron 3.500 muertes de empleados de la central entre 1944 y 1972. De estas muertes, la proporción de muertes por cáncer del SRE (sistema reticuloendotelial) había sido significativamente mayor entre los trabajadores expuestos a niveles bajos de radiación ionizante que entre los no expuestos.

EPI-2. Casos y controles vs. cohorte prospectivo. ¿Cuáles de las siguientes opciones **no** son una ventaja del uso de un estudio de casos y controles por sobre un estudio de cohorte prospectivo (es posible que haya más de una respuesta correcta)?

a. Más económico.
b. Puede completarse en menos tiempo.
c. Es más adecuado para el estudio de enfermedades poco frecuentes.
d. Es más adecuado para el estudio de enfermedades de evolución lenta.
e. Es más adecuado para el estudio de varias exposiciones.
f. Es más adecuado para el estudio de varias enfermedades.
g. Permite una evaluación más precisa de la exposición.

EPI-3. Aleatorización. Se diseñó un ensayo clínico aleatorizado a fin de comparar dos enfoques diferentes para el tratamiento del síndrome de colon irritable. El fin de la aleatorización en este estudio era:

a. aumentar el nivel de compromiso del paciente con el tratamiento;
b. obtener grupos comparativos similares en otras variables que pudieran influir en la enfermedad;
c. obtener grupos comparativos similares en cualquier otra variable medida en el estudio;
d. incrementar las probabilidades de encontrar un efecto significativo del tratamiento;
e. obtener una muestra representativa en el estudio.

EPI-4. Ensayo clínico. En un ensayo clínico aleatorizado diseñado para comparar dos tratamientos para el asma, los clínicos sabían qué tratamiento recibían los pacientes, pero los propios pacientes, no. Éste es un ejemplo de:

a. cumplimiento
b. intención de tratar
c. doble ciego
d. efecto placebo
e. ninguna de las anteriores

EPI-5. Estudio de casos y controles: tuberculosis. (Principalmente para estudiantes de medicina y clínicos) Las siguientes preguntas corresponden al artículo *Variations in the NRAMPI gene and susceptibility to tuberculosis in West Africans* (Bellamy R, Rowende et al., **New Eng J of Med, 38 (10)**, pp. 640-643, marzo de 1998).

a. ¿Quiénes eran los pacientes en este estudio?
b. ¿Quiénes eran los controles?
c. ¿Los controles y los casos diferían en algún aspecto importante además de la condición de enfermedad?
d. ¿Cuál era el diseño de este estudio?
e. Exprese la hipótesis "nula", ya sea con símbolos o palabras.
f. ¿A qué conclusión llegan los investigadores con respecto a la hipótesis nula?

EPI-6. Estudio de cohorte prospectivo. En una conocida investigación de cohorte prospectiva, la población del estudio comprendía a todos los médicos que figuraban en el **Registro Médico Británico** y vivían en Inglaterra y Gales desde octubre de 1951. Mediante un cuestionario, se obtuvo información sobre los antecedentes de tabaquismo de los individuos. La información sobre cáncer de pulmón se obtuvo a partir de actas de defunción y otros datos de mortalidad registrados durante los años posteriores.

a. ¿Qué hace que este estudio sea prospectivo? Indique dos ventajas y dos desventajas de este enfoque.
b. ¿Cuáles son las ventajas y desventajas que ocasiona la selección de médicos como cohorte para el seguimiento?

EPI-7. Intención de tratar. Se realizó un ensayo clínico para comparar una nueva terapia contra la hipertensión con la terapia convencional. Al final del período de seguimiento, los investigadores realizaron dos análisis por separado. Para el primer análisis, siguieron la regla de "intención de tratar". Para el segundo análisis, incluyeron únicamente a los pacientes de los que se sabía que habían recibido la terapia prescrita a lo largo de todo el período de estudio. Los resultados de los dos análisis demostraron una diferencia importante. La explicación más probable para la discrepancia es que: (seleccione la mejor respuesta)

a. la aleatorización no tuvo éxito
b. la nueva terapia era ineficaz
c. había un grado significativo de sesgo de memoria
d. había un problema con el cumplimiento de los pacientes

EPI-8. Enmascaramiento. Se realizó un ensayo clínico para comparar el desempeño de dos tratamientos. Describa una situación en la que el ensayo NO pudiera ser **ciego**. (Asegúrese de que su respuesta demuestre una comprensión del significado de un ensayo ciego).

EPI-9. Aleatorización: ensayos clínicos. ¿Cuál es la finalidad de la aleatorización en un ensayo clínico? (Seleccione la mejor respuesta).

a. Para que los enfermos y no enfermos sean lo más parecidos posible con respecto a todas las variables, a excepción de la exposición de interés.
b. Para reducir la cantidad de individuos que se pierden durante el seguimiento.
c. Para aislar el efecto de la exposición de interés.
d. Para fomentar el cumplimiento con el régimen de tratamiento asignado.

EPI-10. Diseño del estudio: adiposidad y enfermedad coronaria. (Principalmente para estudiantes de medicina y clínicos). Las siguientes preguntas corresponden al artículo *Abdominal adiposity and coronary heart disease in women* (Rexrode KM, Carey VJ et al., **JAMA, 280 (21)**, pp. 1643-1646, diciembre de 1998).

a. ¿Cuál es el diseño de este estudio?
b. ¿Cuáles son las dos hipótesis nulas principales para este estudio?
c. ¿Cuántas mujeres se incluyeron en el análisis final? ¿Qué porcentaje de la cohorte total representa esto?
d. El estudio utilizó informes personales sobre el peso. ¿Cuál es el grado de precisión que ofrece esta información?
e. Dados los resultados de la Tabla 2, ¿cual de las dos variables, la relación cintura-cadera o la circunferencia de la cintura, parece ser el mejor predictor del riesgo de padecer una enfermedad coronaria? ¿Por qué?

EPI-11. Muestreo por densidad. ¿Cuál de las siguientes afirmaciones NO se aplica al muestreo de controles "por densidad" en un estudio de casos y controles? (Seleccione la mejor respuesta).

a. Un individuo identificado como control más tarde puede identificarse como caso.
b. Un individuo identificado como caso más tarde puede identificarse como control.
c. Es probable que el *odds ratio* obtenido a partir de un estudio de esta naturaleza sea una buena estimación de una medida de incidencia de la asociación.
d. Un individuo particular puede servir de control para más de un caso.

EPI-12. Grupo de control. ¿Cuál es la finalidad del grupo de control en un estudio de casos y controles? (Seleccione la mejor respuesta).

a. Proporcionar una estimación del riesgo o la tasa basal de una enfermedad.
b. Proporcionar una estimación de la frecuencia de exposición entre la población que produjo los casos.
c. Proporcionar una estimación de la magnitud del efecto placebo.

d. Proporcionar una estimación de la cantidad de casos previstos entre los no expuestos.

EPI-13. Diseño de estudio: cáncer de mama. En un número reciente de la publicación *American Journal of Epidemiology* se incluyó un trabajo intitulado *Electric Blanket Use and Breast Cancer Risk among Younger Women*. La sección sobre métodos incluía la siguiente información:

Los casos fueron mujeres a las que por primera vez se les había diagnosticado un cáncer de mama in situ o invasivo entre el 1.º de mayo de 1990 y el 31 de diciembre de 1992, y que vivían en tres áreas geográficas determinadas de EE. UU. Los controles fueron mujeres que se identificaron mediante el discado de números telefónicos aleatorios y se equipararon en frecuencia con los casos según la edad, distribuida en grupos de 5 años, y el área geográfica. A todas las mujeres se les preguntó si alguna vez habían usado en forma periódica mantas eléctricas, cobertores eléctricos para colchones o camas de agua calefaccionadas. Una respuesta positiva hacía referencia al uso conjunto de todos o cualquiera de los dispositivos en algún momento de la vida de la encuestada antes de su inclusión en el estudio.

¿Cuál de las siguientes opciones describe mejor el diseño de este estudio?

a. Transversal
b. De cohorte
c. De casos y controles anidados
d. De casos y controles de base poblacional
e. Descriptivo

Respuestas a las preguntas de estudio y auto-evaluaciones

P3.1

1. Experimental
2. Observacional
3. Observacional
4. Experimental
5. Experimental
6. Observacional

P3.2

1. Terapéutico
2. Preventivo

1. a
2. b
3. b
4. a
5. c
6. c
7. a

P3.3

1. Para las usuarias de VDT el porcentaje es (54/366) x 100 = 14,8%; mientras que para las que no usaron VDT el porcentaje es (82/516) x 100 = 15,9%. Estos porcentajes difieren entre sí en apenas un 1%.

P3.4

1. En un estudio de cohorte, la direccionalidad es siempre hacia adelante.
2. La secuencia temporal es prospectiva, puesto que el resultado de salud, aquí el mal de Alzheimer, se produce después del inicio del estudio.

P3.5

1. b
2. a
3. b
4. c
5. b
6. Retrospectivo
7. Retrospectivo
8. Prospectivo
9. Ambos
10. Ninguno
11. Ambos

12. Retrospectivo

P3.6

1. b
2. a
3. a
4. c
5. b
6. De casos y controles
7. De casos y controles
8. De cohorte prospectivo
9. De cohorte prospectivo
10. De casos y controles
11. De cohorte prospectivo
12. De casos y controles
13. De cohorte prospectivo
14. V: Si los controles se seleccionan a partir de una población diferente de la que provienen los casos, existe el riesgo de que haya un sesgo de selección.
15. F: Los controles de base hospitalaria padecen una enfermedad, por lo que no suelen ser representativos de la comunidad que dio origen a los casos.
16. F: Los controles de base poblacional pueden obtenerse mediante el discado de números telefónicos aleatorios correspondientes a la comunidad de la que se obtuvieron los casos. No se garantiza que se seleccionen vecinos de los casos.

P3.7

1. b
2. a
3. c
4. b
5. V
6. F
7. V: Un estudio transversal sólo incluye casos que sobreviven lo suficiente como para estar disponibles para el estudio. Por ende, la exclusión de los que no sobrevivieron podría dar lugar a una conclusión errónea sobre la relación entre una exposición y la enfermedad.
8. V

P3.8

1. Hacia atrás
2. Prospectivo

P3.9

1. Ambos
2. Ambos
3. Ninguno
4. Ninguno
5. Ambos
6. Estudio de casos y controles anidados
7. Ambos
8. Estudio de casos y controles anidados
9. Ninguno
10. Estudio de caso-cohorte

P3.10

1. F: La unidad de análisis en un estudio ecológico es un grupo (por ejemplo, un sector censal, un estado o un país) y no se obtiene información sobre la exposición y enfermedad de los individuos de manera simultánea.
2. V: Un estudio de mortalidad proporcional incluye observaciones sobre muertes, pero carece de información sobre la población candidata (es decir, los denominadores).
3. V: Tanto los estudios ecológicos como los estudios de mortalidad proporcional utilizan diseños “incompletos”.
4. F: Los estudios de mortalidad proporcional emplean información sobre muertes de individuos.
5. V
6. V

LECCIÓN 4

MEDIDAS DE FRECUENCIA DE ENFERMEDAD

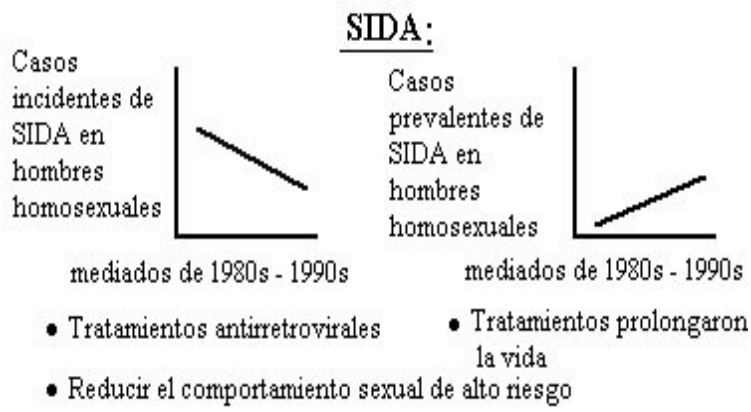
En los estudios epidemiológicos se utiliza una medida de frecuencia de enfermedad para determinar qué tan a menudo ocurre la enfermedad u otro resultado de salud en varios subgrupos de interés. En este capítulo se describen dos tipos básicos de medidas de frecuencia de enfermedad: las medidas de incidencia y las de prevalencia. La elección normalmente depende del diseño de estudio empleado y sus objetivos.

4-1

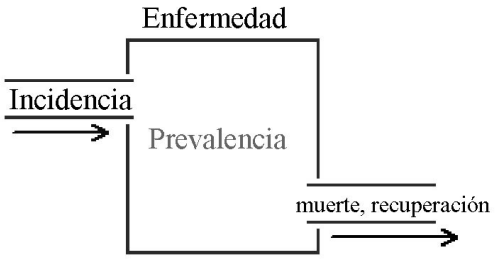
Incidencia versus prevalencia

Existen dos tipos generales de medidas de frecuencia de enfermedad: **incidencia (I)** y **prevalencia (P)**. La incidencia mide casos ***nuevos*** de una enfermedad los cuales se desarrollan durante un período de tiempo. La prevalencia mide casos ***existentes*** de una enfermedad en un punto específico del tiempo o en un período.

Para ejemplificar cómo difieren incidencia y prevalencia, consideremos nuestra experiencia con el SIDA. La cantidad de casos incidentes anuales de SIDA en hombres homosexuales disminuyó en los Estados Unidos desde mediados de la década de los ochenta hasta fines de la década de los noventa. Principalmente, esto fue el resultado tanto de los enfoques recientes de tratamientos antirretrovirales como de las estrategias de prevención para reducir el comportamiento sexual de alto riesgo. Por el contrario, los casos prevalentes anuales de SIDA en hombres homosexuales han aumentado enormemente en los Estados Unidos durante el mismo período porque los enfoques recientes en cuanto a tratamientos del SIDA tuvieron éxito en prolongar la vida de las personas infectadas con el virus del VIH y/o el SIDA.



La prevalencia también puede entenderse como la descripción de un grupo de enfermedades en una población, mientras que la incidencia representa el flujo de entrada de casos de ese grupo, al igual que la muerte y recuperación reflejan el flujo de salida del mismo.



Las medidas de incidencia son útiles para identificar los factores de riesgo y para evaluar la etiología de la enfermedad. Normalmente, las medidas de incidencia se calculan a partir de ensayos clínicos y estudios de cohortes, que implican el seguimiento de los individuos en el tiempo.

Las medidas de prevalencia no son tan útiles como las de incidencia para evaluar la etiología, porque la prevalencia no considera a las personas que mueren debido a la enfermedad antes de que comience el estudio de prevalencia. Generalmente, las medidas de prevalencia se calculan a partir de estudios transversales y estudios de casos y controles que usan casos prevalentes, en lugar de incidentes. Ya que la cantidad de casos prevalentes indica la demanda de atención médica, las medidas de prevalencia son las más útiles para planificar los servicios de salud.

Resumen

- La incidencia se refiere a casos nuevos de una enfermedad u otro resultado de salud durante un período de seguimiento.
- La prevalencia se refiere a casos existentes de una enfermedad en un punto del tiempo.
- Las medidas de incidencia son útiles para identificar los factores de riesgo y para evaluar la etiología de la enfermedad.
- Las medidas de prevalencia son las más útiles para la planificación de los servicios de salud.

La mortalidad podría utilizarse en lugar de incidencia de la enfermedad

Se analiza la incidencia y la prevalencia en función de casos nuevos o existentes de una **enfermedad**, independientemente de si estos casos provocan o no la muerte después de un período de estudio. Sin embargo, existen muchas situaciones en las que también vale la pena utilizar información estrictamente sobre **mortalidad**.

Las medidas de mortalidad son una herramienta importante para la **vigilancia epidemiológica**. En la actualidad, dichos programas de vigilancia se han aplicado para monitorear la presencia de una amplia variedad de eventos relacionados con la salud, incluso la muerte, en poblaciones grandes. Los datos estadísticos sobre mortalidad también son prácticos para **evaluar hipótesis etiológicas**, especialmente cuando los datos de incidencia no se encuentran disponibles. En especial, para las enfermedades con una baja tasa de cura o recuperación, tales como el cáncer de pulmón, las medidas de mortalidad brindan una aproximación razonable a las medidas de incidencia.

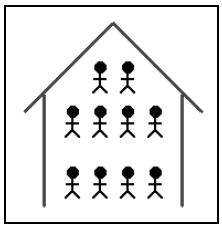
La utilización de la información sobre mortalidad para cualquiera de los propósitos mencionados anteriormente tiene varias ventajas pragmáticas:

- Los datos sobre mortalidad son fáciles de recabar y están virtualmente completos, ya que el registro de las defunciones es obligatorio en la mayoría de los países industrializados y son muy pocas las muertes que no se informan.
- Los datos de mortalidad se definen empleando nomenclatura estandarizada. En particular, la Clasificación Internacional de Enfermedades (ICD, por sus siglas en inglés) se usa para promover la uniformidad cuando se informan las causas de muerte.
- El registro de datos de mortalidad es relativamente económico.

Ejemplo de los invitados

Supongamos que llegan invitados a su casa a una tasa de dos por día y se quedan exactamente cinco días. ¿Cuántas personas habrá en su casa después de una semana?

Veamos lo que sucede día por día. El primer día llegan dos invitados y ninguno se va, por lo tanto, hay 2 invitados en su casa al final del primer día. El segundo día llegan dos invitados más y ninguno se va, por lo tanto, ahora hay 4 invitados en su casa después de 2 días. De la misma forma, hay 6 invitados después de 3 días, 8 después de 4 días y 10 invitados después de cinco días, y ninguno se ha ido hasta el momento. El sexto día llegan dos nuevos invitados, pero los dos invitados que llegaron el primer día, y se quedaron durante cinco días, ahora se van, dejándolo nuevamente con 10 invitados en la casa. Al final del séptimo día, habrá todavía 10 invitados, lo cual responde a la pregunta formulada al comienzo de todo esto.



Este escenario ejemplifica la diferencia fundamental entre incidencia y prevalencia. En el ejemplo, después de 5 días, se llega a un "estado estacionario" en el que habrá 10 invitados siempre que la tasa de llegada sea de 2 por día. Este estado estacionario de 10 invitados es una prevalencia, que describe la cantidad existente de invitados en cualquier momento después de que se haya alcanzado el estado estacionario. La tasa de llegada de 2 invitados por día es una incidencia, que describe la velocidad con la que llegan nuevos invitados. La duración de los cinco días durante los que se quedan los invitados en su casa es la información necesaria para vincular la incidencia con la prevalencia.

La prevalencia puede vincularse con la incidencia mediante la siguiente fórmula:

P = I x D

En nuestro ejemplo, **P** es la cantidad de invitados en la casa cualquier día después de cinco días, **I** es la tasa de llegada de 2 invitados por día y **D** es la duración de 5 días para cada invitado. La fórmula funciona en este ejemplo, ya que 2 multiplicado por 5 es igual a 10.

De esta fórmula podemos ver que para una incidencia dada, la prevalencia aumentará o disminuirá a medida que aumente o disminuya la duración. Por ejemplo, si los invitados se quedaran 8 días en lugar de 5, con la misma tasa de llegada, la cantidad de invitados que hay en la casa en estado estacionario sería 2 multiplicado por 8, lo que es igual a 16, en vez de 10.

Para una duración dada, la prevalencia aumentará o disminuirá a medida que aumente o disminuya la incidencia. Por lo tanto, si los invitados llegan a la tasa de sólo 1 individuo por día, en lugar de 2, y se quedan 8 días, la prevalencia será 1 multiplicado por 8, que es igual a 8, en vez de 16.

Resumen

- Un escenario en el que llegan 2 invitados por día y se quedan 5 días ejemplifica la diferencia fundamental entre incidencia y prevalencia.
- Un estado estacionario de 10 invitados es un ejemplo de *prevalencia*, que describe la cantidad existente de invitados en cualquier momento después de que se haya alcanzado dicho estado estacionario.
- La llegada de 2 invitados por día es un ejemplo de *incidencia*, que describe la velocidad con la que llegan nuevos invitados.
- La *duración* de 5 días es la información necesaria para vincular la forma en que la incidencia lleva a la prevalencia.
- La prevalencia se obtiene multiplicando la incidencia por la duración (P = I x D).

Relación entre prevalencia e incidencia

En el ejemplo de los "invitados", la fórmula

$$P = (I \square D)$$

se utilizó para demostrar que la cantidad de invitados en estado estacionario que estaban en la casa después de 7 días era igual al producto de la cantidad de invitados que llegaban cada día por la duración de la estadía de cada invitado en la casa.

Los términos **P**, **I** y **D** en esta fórmula representan los conceptos de prevalencia, incidencia y duración, respectivamente; pero, tal como se utilizan en el ejemplo, cada uno de ellos no se ajusta estrictamente a las definiciones epidemiológicas de estos términos. Tal como se describe en actividades posteriores en esta lección (es decir, capítulo) sobre medidas de frecuencia de enfermedad, las definiciones estrictas de prevalencia e incidencia requieren denominadores, mientras que el escenario de "los invitados" analizado aquí utiliza sólo la información del numerador.

Específicamente, la ***prevalencia*** se calcula utilizando la fórmula:

$$P = \frac{C}{N}$$

Además, la ***incidencia*** emplea una de las siguientes dos fórmulas posibles, ya sea que la medida de incidencia elegida sea el riesgo o la tasa:

$$IA = I/N \quad o \quad TI = I/PT$$

En las fórmulas anteriores, **P**, **C** y **N** denotan la prevalencia, la cantidad de casos existentes y el tamaño de la población en estado estacionario, respectivamente. Asimismo, **IA** denota la incidencia acumulada, que calcula el riesgo, **I** indica la cantidad de nuevos casos (incidentes) y **N** denota el tamaño de una cohorte sin enfermedad seguida durante todo el período de estudio. Además, **TI** indica la tasa de incidencia y **PT** la información persona-tiempo acumulada. Todas estas fórmulas se describen y ejemplifican en actividades subsiguientes.

El punto importante que se quiere explicar aquí es que las tres fórmulas anteriores tienen denominadores, los cuales no se utilizaron en el ejemplo de los invitados, pero sí se requieren para calcular la prevalencia y la incidencia en epidemiología.

El término **D** en la fórmula al inicio de esta página fue utilizado en el ejemplo de los invitados para definir la duración de la estadía que se estimó para cada invitado. En el uso epidemiológico de esta fórmula, **D** en verdad denota la duración promedio de la enfermedad para todos los individuos de la población bajo estudio, en lugar de suponer que es la misma para cada persona de la población.

Sin embargo, usando las definiciones más estrictas de las medidas de prevalencia e incidencia y la duración promedio, la fórmula anterior -que relaciona la prevalencia con la incidencia y la duración- sigue teniendo validez, siempre y cuando la población esté en estado estacionario y la enfermedad sea poco frecuente. La expresión "estado estacionario" implica que aunque la población puede ser dinámica, la cantidad de personas que entran y salen de ella por cualquier razón es esencialmente la misma durante el período de estudio, de forma que la población no cambia. Si la enfermedad no es poco frecuente, se requiere, entonces, una fórmula de relación de la prevalencia con la incidencia modificada, concretamente:

$$P = \frac{I \square D}{(I \square D) + 1}$$



Auto-evaluación (P4.1)

Para cada uno de los siguientes escenarios, determine si está más estrechamente relacionado con **incidencia** o con **prevalencia**.

1. Cantidad de campistas que desarrollaron gastroenteritis a los pocos días de comer ensalada de papa en el salón comedor. . . . . **???**
2. Cantidad de personas que informaron padecer diabetes en la Encuesta Nacional de Salud. . **???**
3. Infarto de miocardio agudo (ataque cardíaco) entre los participantes durante los primeros 10 años de seguimiento del estudio Framingham. . . . . **???**
4. Cantidad de personas fallecidas cuyas muertes se atribuyeron al huracán Floyd en Carolina del Norte en 1999. . . . . **???**
5. Cantidad de niños que son inmunes al sarampión, ya sea porque tuvieron la enfermedad o porque fueron vacunados. . . . . **???**

Suponga que un sistema de vigilancia pudo capturar con precisión y en su totalidad todos los casos nuevos de una enfermedad dentro de una comunidad. Presuma también que se realizó una encuesta el 1.º de julio en donde se preguntó a todos los miembros de la comunidad si padecían la enfermedad en ese momento. Para cada una de las siguientes afecciones, determine si es *probable que sea más alta* la **incidencia** (por 1.000 personas por año) o la **prevalencia** (por 1.000 personas el 1.º de julio).

6. Rabia (poco frecuente y de duración corta, por ejemplo, muerte dentro de la primera semana). **???**
7. Esclerosis múltiple (poco frecuente, duración larga [muchos años]). . . . **???**
8. Gripe (frecuente pero en invierno, duración corta). . . . . **???**
9. Dermatitis por hiedra venenosa (frecuente en primavera/verano/otoño, duración de 2 semanas). **???**
10. Hipertensión arterial (relativamente frecuente, duración de por vida). . . . **???**

4-2

Riesgo

El término *riesgo* se utiliza comúnmente en la vida cotidiana para describir la posibilidad, o probabilidad, de que ocurra algún evento de interés. Nos podríamos preguntar, por ejemplo, cuál es el riesgo de que el mercado de valores se desplome o de que estemos involucrados en un choque automovilístico grave. Podríamos preocuparnos sobre el riesgo de desarrollar un problema de salud no deseado como, por ejemplo, una enfermedad que ponga en peligro nuestra vida, o incluso que corramos el riesgo de morir.

En epidemiología, el riesgo es la probabilidad de que un individuo con ciertas características, por ejemplo, edad, raza, sexo y tabaquismo, desarrolle o muera por una enfermedad o, incluso, en forma más general, experimente un cambio de interés en el estatus de salud durante un período de seguimiento determinado. Cuando

el resultado de salud es una enfermedad, esta definición supone que el individuo no la padece al comienzo del seguimiento y no muere por ninguna otra causa durante dicho período. Debido a que el riesgo es una probabilidad, es un número entre 0 y 1, o un porcentaje correspondiente.

$$0 \leq \text{RIESGO} \leq 1$$

$$(0 \leq \text{Porcentaje} \leq 100)$$

Cuando se describe el riesgo, es necesario especificar un plazo de seguimiento, que se denomina **período de riesgo**. Por ejemplo, para describir el riesgo de que un hombre de 45 años desarrolle cáncer de próstata, debemos detallar el período de riesgo como, por ejemplo, 10 años de seguimiento, durante el cual queremos predecir este riesgo. Si el período de riesgo fuera, por ejemplo, 20 años en lugar de 10, esperaríamos que nuestro cálculo del riesgo fuera mayor que el riesgo en 10 años, ya que se está dando más tiempo para que se desarrolle la enfermedad.

Preguntas de estudio (P4.2)

1. ¿Cuál es el significado de la siguiente afirmación? El riesgo de que un hombre de 45 años desarrolle cáncer de próstata en 10 años es del 5%. (Formule su respuesta en términos de probabilidad y sea lo más específico posible sobre los supuestos requeridos).

2. ¿El riesgo en 5 años para la misma persona descrita en la pregunta anterior será mayor o menor que el riesgo en 10 años? Desarrolle su respuesta brevemente.

Resumen

- El riesgo es la probabilidad de que un individuo desarrolle o muera por una enfermedad determinada, o en forma más general, experimente un cambio en el estatus de salud durante un período de seguimiento determinado.
- El riesgo supone que el individuo no padece la enfermedad al comienzo del seguimiento y no muere por ninguna otra causa durante el mismo.
- El riesgo debe ser un valor entre 0 y 1, o el porcentaje correspondiente.
- Es necesario especificar un período de seguimiento durante el cual se predecirá el riesgo de un determinado resultado de salud.

Confusión entre riesgo y tasa

A menudo, el término ***tasa*** se ha empleado en forma incorrecta para describir una medida de ***riesgo***. Por ejemplo, el término ***tasa de ataque*** se utiliza frecuentemente en los estudios del brote de alguna afección para describir el cálculo de la probabilidad de desarrollar una enfermedad infecciosa cuando, en verdad, se hace un cálculo del ***riesgo***. Asimismo, el término ***tasa de mortalidad*** se ha confundido con ***riesgo de muerte*** en los estudios de mortalidad. En particular, el término ***tasa de letalidad*** se ha usado muchas veces incorrectamente para describir la proporción de casos que mueren, es decir, dicha proporción, en verdad, está calculando un ***riesgo***.

Los términos riesgo y tasa tienen significados muy diferentes, como se describe en otras actividades en esta lección. Lo ideal es aplicar el término correcto a la medida real que se está utilizando. Esto no siempre sucede cuando se publican resultados epidemiológicos. En consecuencia, cuando se lee material sobre epidemiología, se debe tener cuidado al determinar la medida real que se está informando.

Incidencia acumulada

La forma más común de calcular el riesgo es dividir la cantidad de casos nuevos detectados que se desarrollaron durante el seguimiento por la cantidad de individuos disponibles que no padecían la enfermedad al comienzo del seguimiento. Dicho cálculo se denomina muchas veces ***incidencia acumulada*** o **IA**. Cuando se describe la incidencia acumulada, es necesario especificar el período de seguimiento durante el cual se calcula el riesgo.

$$IA = \frac{I}{N} = \frac{\text{\# de casos nuevos durante seguimiento}}{\text{\# de individuos que no padecían la enfermedad al comienzo del seguimiento}}$$

En términos técnicos, la incidencia acumulada no es equivalente al riesgo individual, sino que más bien es una estimación del riesgo individual calculado, ya sea a partir de una población entera o de una muestra de ella. Sin embargo, a menudo utilizamos los términos riesgo e incidencia acumulada en forma indistinta, como lo hacemos en todo este curso.

Generalmente colocamos un acento circunflejo ("^") sobre la IA cuando el cálculo de la incidencia acumulada se basa en una muestra; no colocamos dicho acento si tenemos los datos para toda la población.

$\hat{I}A =$ IA "acento circunflejo"

La fórmula de incidencia acumulada, con o sin "acento circunflejo" siempre es una proporción, por lo que sus valores pueden variar entre 0 y 1. Si la incidencia acumulada es alta, como sucede en un brote, la IA a veces se expresa como un porcentaje.

Como ejemplo simple, supongamos que hacemos un seguimiento de 1.000 hombres de 45 años y descubrimos que 50 desarrollaron cáncer de próstata dentro de los 10 años de seguimiento y que no se perdió el seguimiento de ningún individuo o ninguno abandonó el estudio. Entonces nuestro cálculo de incidencia acumulada es 50 sobre 1.000, ó 0,05 ó 5%.

$$\hat{I}A = \frac{I}{N} = \frac{50}{1000} = 0{,}05 = 5\%$$

En otras palabras, se calcula que el riesgo en 10 años o, en términos técnicos, la incidencia acumulada para un hombre de 45 años es del 5%. La fórmula incluida para calcular el riesgo se denomina, muchas veces, **incidencia acumulada simple** porque es una proporción simple que supone una cohorte fija. Sin embargo, el uso de la incidencia acumulada simple no es siempre adecuado en todos los tipos de estudios de seguimiento. En las actividades siguientes se discuten problemas con incidencia acumulada simple y los métodos para solucionarlos.

Resumen

- La incidencia acumulada (IA) es un cálculo del riesgo individual basado en la población.
- La incidencia acumulada siempre es una proporción.
- Cuando se describe la incidencia acumulada, es necesario especificar el período de seguimiento durante el cual se calcula el riesgo.
- La fórmula de incidencia acumulada es IA=I/N, donde I denota la cantidad de casos nuevos de la enfermedad que se desarrollan durante el período de seguimiento y N indica el tamaño de la población sin enfermedad al inicio del seguimiento.
- Los términos incidencia acumulada y riesgo se utilizan en forma indistinta en este curso, aunque técnicamente, son distintos.

Utilización de datos de población para calcular el riesgo

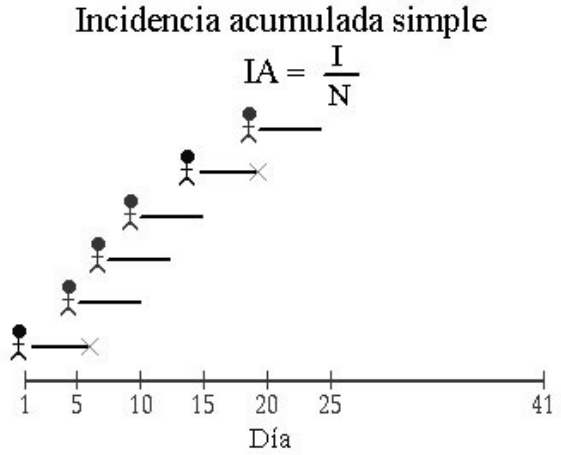
Suponga que hacemos un seguimiento de 1.000 hombres de 45 años para calcular que el riesgo en 10 años de que desarrollen cáncer de próstata sea 0,05 ó 5%. Para obtener este cálculo, debemos utilizar la información de un grupo de individuos, todos de la misma edad, a fin de predecir el riesgo de un solo individuo. Podríamos obtener un cálculo mucho mejor si supiéramos sus características específicas, por ejemplo, su dieta, si es fumador o no o si consume alcohol. Sin embargo, incluso si conociéramos más características, todavía tendríamos que depender de un cálculo de riesgo basado en datos obtenidos de un grupo de individuos en los que se midan u observen estas características adicionales. Así es como trabajan los epidemiólogos para calcular el riesgo individual; es decir, deben depender de la acumulación de pruebas basadas en datos demográficos.

Cambio o restructuración de la cohorte

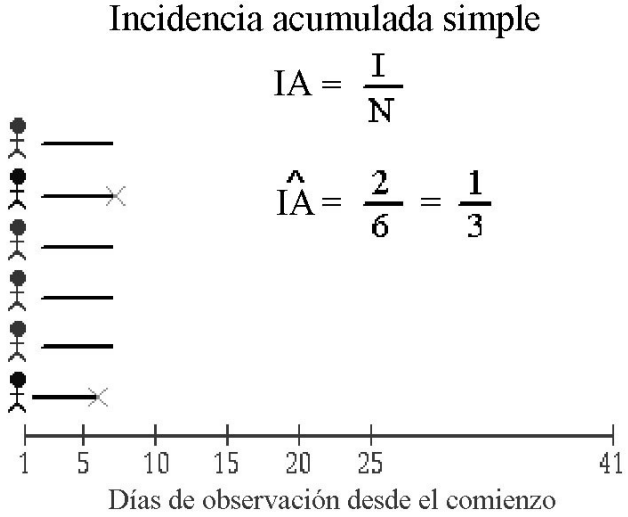
La fórmula de *incidencia acumulada simple* supone implícitamente que la cohorte es "**fija**", en el sentido de que no se permiten entradas en ella durante el período de seguimiento. ¿Qué deberíamos hacer si permitimos nuevas incorporaciones en la cohorte?

Por ejemplo, en el estudio de usuarios de las playas de Sydney descrito en la lección 2, los individuos se seleccionaron de 12 playas concurridas de Sydney durante 41 días de muestreo en los meses de verano de 1989-90. Los individuos pudieron incorporarse progresivamente a la cohorte en distintos días durante el verano, y una semana después se obtuvieron informes personales sobre la exposición y las enfermedades.

Como ejemplo, considere estos seis individuos. Se hace un seguimiento de cada uno de ellos durante los 7 días requeridos. Los individuos 1 y 5 (contando de abajo hacia arriba) son los únicos que manifestaron haberse enfermado.



Podemos reestructurar estos datos cambiando la línea de seguimiento para cada persona del margen izquierdo, de forma que el eje horizontal que describe el tiempo ahora refleje los días de observación desde el comienzo del examen de cada individuo, en lugar de los días calendario reales en los que las observaciones ocurrieron. Esto se ajusta al seguimiento de una cohorte fija, para la cual se calcula una incidencia acumulada de 2/6 o un tercio.



A menudo tenemos una cohorte que permite la incorporación progresiva de individuos al estudio en distintos días calendario. Podemos cambiar o reestructurar la cohorte para que sea fija cambiando los datos de cada individuo, a fin de reflejar el tiempo de observación desde la entrada inicial al estudio, en lugar de los días calendario. Entonces, podemos utilizar la fórmula de incidencia acumulada simple para estimar el riesgo.

Preguntas de estudio (P4.3)

1. Después de cambiar o reestructurar el tiempo de observación de la entrada inicial al estudio de la cohorte, ¿debemos suponer que todos los individuos que se convirtieron en casos de enfermedad fueron acompañados u observados durante la misma cantidad de tiempo que los individuos que permanecieron sin la enfermedad?

Suponga que después de cambiar la cohorte, un individuo no padece la enfermedad durante 4 años de seguimiento, mientras que otro individuo de la cohorte no padece la enfermedad, pero fue observado sólo durante 2 años.

2. ¿Existe algún problema en calcular la incidencia acumulada (IA) que incluye a estos dos individuos en el denominador de la fórmula de IA?
3. Después de que hayamos cambiado la cohorte, ¿tenemos que suponer que para calcular la incidencia acumulada (IA) se hizo un seguimiento de TODOS los individuos, incluso de aquellos que se convirtieron en casos, durante la misma cantidad de tiempo?

Resumen

- Si los individuos se incorporan al estudio progresivamente en diferentes días calendario, los datos pueden ***cambiarse*** para reflejar el tiempo de observación desde la entrada inicial.
- La incidencia acumulada simple puede utilizarse para estimar el riesgo para una cohorte ***cambiada***.
- Después de cambiar la cohorte, podemos calcular la incidencia acumulada, siempre y cuando, durante el período de seguimiento completo, se haga un control de todos los individuos que no padecieron la enfermedad durante el mismo.

Incidencia acumulada simple: limitaciones



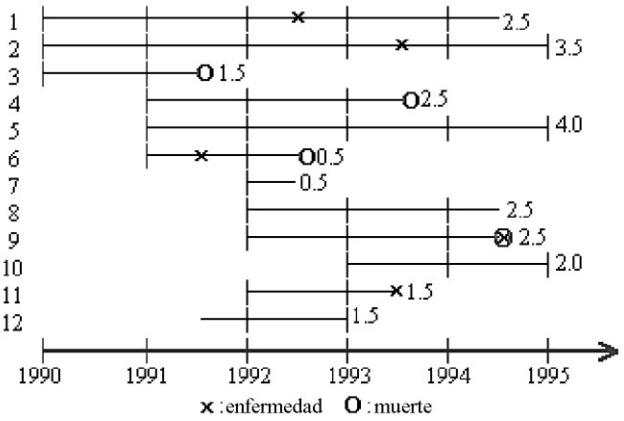
Existen varias limitaciones potenciales al estimar una cohorte fija cuando se utiliza la fórmula de incidencia acumulada simple para calcular el riesgo. Por ejemplo, es probable que se reduzca el tamaño de una cohorte fija durante el período de seguimiento, como resultado de muertes u otras fuentes de desgaste tales como la pérdida de seguimiento o el abandono del estudio. No sabemos si un individuo del cual se perdió el seguimiento desarrolló la enfermedad de interés.

Otro problema surge si la población en estudio es una **población dinámica**, en lugar de una **cohorte fija**. Una cohorte fija es un grupo de individuos identificados en un momento y seguidos durante un período dado para la detección de casos nuevos. La cohorte es "fija" en el sentido de que no se permiten incorporaciones al estudio después del comienzo del seguimiento, aunque pueden ocurrir posteriores pérdidas de individuos por varias razones como, por ejemplo, abandono del estudio, migración o muerte. Pero una población dinámica está en constante cambio, lo que permite tanto la incorporación de miembros nuevos como la pérdida de aquellos que se habían incorporado anteriormente durante el período de seguimiento.

El denominador de la fórmula de incidencia acumulada simple no refleja el tamaño continuamente cambiante de una población dinámica. Además, el numerador en la fórmula de incidencia acumulada simple no cuenta casos nuevos que pueden surgir de esas personas que se incorporaron a la población dinámica después del inicio del seguimiento.

Otra dificultad tanto para una cohorte fija como para una dinámica es que se puede efectuar el seguimiento de los individuos durante *diferentes períodos de tiempo,* de forma que para un cálculo de incidencia acumulada no será útil diferir los períodos de seguimiento. Este problema puede presentarse cuando se pierde el seguimiento de individuos o alguien abandona el estudio. También podría surgir cuando los individuos se incorporan al estudio después de que éste comienza y no padecen la enfermedad hasta que el estudio finaliza, o si el tiempo de seguimiento en el que el individuo desarrolla la enfermedad varía para diferentes individuos.

Para ilustrar estos problemas, consideremos un ejemplo hipotético que involucra a 12 individuos que inicialmente no padecen la enfermedad y a quienes se efectúa un seguimiento durante un período de 5 años, de 1990 a 1995.



Una **X** denota el momento en el que se diagnosticó la enfermedad al individuo y un **círculo** (**O**) indica el momento del fallecimiento, *que podría deberse a la enfermedad (círculo con una* **X** *dentro) o a otra causa (círculo sin una* **X** *dentro).* Aquellos individuos que no tienen X o círculo en su línea de tiempo abandonaron el estudio o se perdió su seguimiento, o fueron seguidos hasta el final del estudio sin desarrollar la enfermedad. El valor a la derecha de la línea de tiempo de cada individuo denota su período de seguimiento hasta que se le diagnosticó la enfermedad, abandonó el estudio o se perdió su seguimiento, o hasta que terminó el estudio. Basándose en esta información, responda las siguientes preguntas:

Preguntas de estudio (P4.4)

Las siguientes preguntas se refieren a la figura anterior:

1. ¿Qué tipo de cohorte se está estudiando, **fija** o **dinámica**?

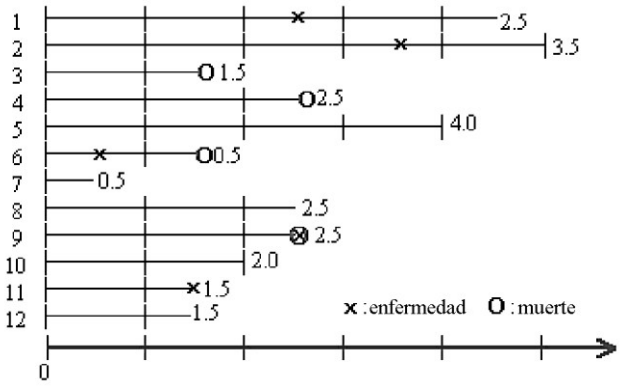
2a. ¿A cuál de estos individuos se le diagnosticó la enfermedad?
Individuo 2 **Individuo 3** **Individuo 5** **Individuo 7**

2b. ¿De cuál de estos individuos se perdió el seguimiento o cuál abandonó el estudio?
Individuo 2 **Individuo 3** **Individuo 5** **Individuo 7**

2c. ¿Cuál de estos individuos murió a raíz de la enfermedad?
Individuo 3 **Individuo 5** **Individuo 7** **Individuo 9**

2d. ¿Cuál de estos individuos murió *sin* contraer la enfermedad?
Individuo 3 **Individuo 5** **Individuo 7** **Individuo 9**

2e. ¿Cuál *no padecía* la enfermedad y estaba vivo al final del estudio?
Individuo 3 **Individuo 5** **Individuo 7** **Individuo 9**



3. Si pudiésemos cambiar la cohorte, ¿cuál es el valor de la incidencia acumulada simple del diagnóstico de la enfermedad, expresado *en porcentaje*?

4. ¿Cuál es el valor de la incidencia acumulada simple de muerte, expresado *en porcentaje sin decimales?*

5. *Utilizando el gráfico "sin cambios" de la página anterior*, ¿cuál de las siguientes características tienen en común los individuos 5, 8, 10 y 12?

   A. La misma cantidad de tiempo de seguimiento.
   B. Fueron incorporados al estudio el mismo día calendario.
   C. Abandonaron el estudio.
   D. No desarrollaron la enfermedad durante el seguimiento.

El cálculo de incidencia acumulada simple para los datos mostrados anteriormente es un problema porque...

6a. no todos los individuos desarrollaron la enfermedad. **Sí** **No**

6b. no todos los individuos murieron. **Sí** **No**

6c. la cohorte es dinámica. **Sí** **No**

6d. algunos individuos murieron por otra causa.
Sí No

6e. se perdió el seguimiento de algunos individuos o algunos abandonaron el estudio.
Sí No

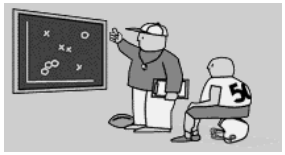
6f. algunos individuos desarrollaron la enfermedad en diferentes momentos del seguimiento.
Sí No

6g. algunos individuos que no desarrollaron la enfermedad tuvieron distintos tiempos de seguimiento.
Sí No

Resumen

- Existen problemas al estimar una cohorte fija cuando se utiliza la fórmula de incidencia acumulada simple para calcular el riesgo.
- Si una cohorte fija se desgasta, no sabremos si el individuo del cual se perdió seguimiento desarrolló la enfermedad.
- Para una cohorte dinámica, el denominador en la fórmula de incidencia acumulada simple no refleja el tamaño de la población en constante cambio.
- La incidencia acumulada simple no permite que se realice un seguimiento de los individuos durante períodos de tiempo distintos.

Auto-evaluación (P4.5)



Después del segundo juego de la temporada de fútbol universitario, 60 miembros del equipo -compuesto por 97 personas- desarrollaron fiebre, malestar general, pérdida de apetito y molestia abdominal. En pocos días, 30 jugadores desarrollaron ictericia. Se tomaron muestras de sangre de todos los miembros del equipo para efectuar la prueba de anticuerpos para la hepatitis A (diagnóstico presunto) y para determinar si presentaban un aumento de las enzimas hepáticas.

1. ¿Cuál es la incidencia acumulada de la ictericia? . . . . . **???**

2. Si supone que todas las personas con síntomas tenían hepatitis A, incluso aquellas que no desarrollaron ictericia, ¿cuál es la incidencia acumulada presunta de la hepatitis A? . . . **???**

3. Las pruebas de laboratorio revelaron que 91 personas tenían niveles elevados de enzimas hepáticas, de las cuales 90 tenían anticuerpos IgM, lo que indica una infección aguda de hepatitis A. Dos jugadores con niveles normales de enzimas hepáticas tenían anticuerpos IgG, lo que indica que ya habían estado expuestos a la hepatitis A y que ahora eran inmunes. ¿Cuál es la incidencia acumulada de la hepatitis A? . . . . . **???**

Opciones
30/60 30/97 60/97 90/91 90/95 90/97 91/95 91/97

Coloque **verdadero** o **falso** en cada una de las siguientes afirmaciones:

4. La incidencia acumulada siempre es una proporción, incluso para una cohorte con incorporación escalonada ("cohorte cambiada"). . . . . **???**

5. La incidencia acumulada es una medida útil para las enfermedades con períodos de incubación cortos en poblaciones bien definidas. . . . . **???**

6. La incidencia acumulada no es una medida tan ideal para las enfermedades con períodos de incubación prolongados en poblaciones dinámicas. . . . . **???**

7. Si una población fija tiene una pérdida sustancial de seguimiento, la incidencia acumulada sobreestimará el verdadero riesgo de la enfermedad. . . . . **???**

4-3

Tasa

*La **tasa** es una medida de frecuencia de la enfermedad que describe la velocidad con que se presentan eventos relacionados con la salud tales como nuevos diagnósticos de casos o muertes en la población de interés. Sinónimos: **riesgo instantáneo, densidad de incidencia**.*

Concepto de tasa

El concepto de tasa no es tan fácil de entender como el concepto de riesgo, y a menudo se le confunde con éste. En términos generales, la tasa es una medida de la rapidez con que sucede un hecho de interés. Cuando queremos saber a qué velocidad estamos manejando, o qué tan rápido aumentan los precios en el mercado de valores, o a qué ritmo disminuye la tasa de criminalidad, estamos buscando una tasa.

Supongamos que viajamos en auto. Estamos manejando por una autopista, miramos nuestro velocímetro y vemos que la velocidad es de 65 millas por hora. ¿Esto significa que recorreremos exactamente 65 millas en la próxima hora? Por supuesto que no. La lectura del velocímetro nos dice a qué velocidad estamos manejando en el momento en que leemos el velocímetro. Si pudiésemos manejar exactamente de esta forma durante la próxima hora sin detenernos a cargar gasolina o a descansar, o sin disminuir la velocidad debido al tránsito pesado, recorreríamos exactamente 65 millas en la próxima hora. La lectura de 65 millas por hora en nuestro velocímetro es la velocidad vectorial a la cual estamos viajando, y ésta es un ejemplo de tasa.

En verdad, la velocidad vectorial es un ejemplo de una **tasa instantánea**, ya que describe la rapidez con la que viajamos en un instante específico del tiempo. Hay otro tipo de tasa, llamada **tasa promedio**, que también podemos ejemplificar continuando con nuestro viaje en automóvil. Si en verdad manejamos por la autopista durante la próxima hora y cubrimos 55 millas en ese tiempo, la tasa promedio, llamada muchas veces *velocidad escalar por hora*, sería 55.

En epidemiología se utiliza una tasa para medir la rapidez con la que se desarrollan nuevos casos de una enfermedad o, en forma alternativa, qué tan rápido mueren las personas que padecen una enfermedad de interés. Al igual que con la velocidad vectorial o la velocidad escalar, podríamos querer conocer la tasa instantánea o la tasa promedio. Con los datos epidemiológicos, normalmente es más fácil determinar una tasa promedio que una tasa instantánea. Difícilmente podríamos esperar tener un dispositivo tipo velocímetro que mida con qué rapidez se presenta una enfermedad en un momento específico en una cohorte de individuos. En consecuencia, en los

estudios epidemiológicos, normalmente medimos la tasa promedio a la que surge una enfermedad durante un período de tiempo.

Debido a que la tasa es una medida de la rapidez con que algo ocurre, siempre se mide en unidades de tiempo, por ejemplo, días, semanas, meses o años. Esto clarifica su interpretación. Si describimos una tasa de 50 casos nuevos por 10.000 personas-años, significa que en una cohorte de individuos se presenta un promedio de 50 casos por cada 10.000 años de seguimiento sin la enfermedad. La cifra de 10.000 se obtiene sumando los tiempos de seguimiento para todos los individuos de la cohorte.

Si la unidad de tiempo fuera *meses* en lugar de *años,* la interpretación de la tasa puede ser bastante diferente. Una tasa de 50 casos nuevos por 10.000 personas-meses indica una tasa mucho más rápida que 50 casos nuevos por 10.000 personas-años.

Preguntas de estudio (P4.6)

1. ¿Cuál de las siguientes tasas *no* es equivalente a una tasa de 50 casos nuevos por 10.000 personas-años?

    A. 100 casos nuevos por 20.000 personas-años
    B. 50 casos nuevos por 120.000 personas-meses
    C. 50 casos nuevos por 52.000 personas-semanas

2. Determine si cada una de las siguientes afirmaciones describe una tasa:

A. 5 casos nuevos por 100 personas-días	**Sí**	**No**
B. 40 millas por hora	**Sí**	**No**
C. 10 casos nuevos en 100 personas que no padecían la enfermedad	**Sí**	**No**
D. 60 asesinatos nuevos por año	**Sí**	**No**
E. 60 muertes en 200 pacientes con cáncer de pulmón	**Sí**	**No**

Resumen

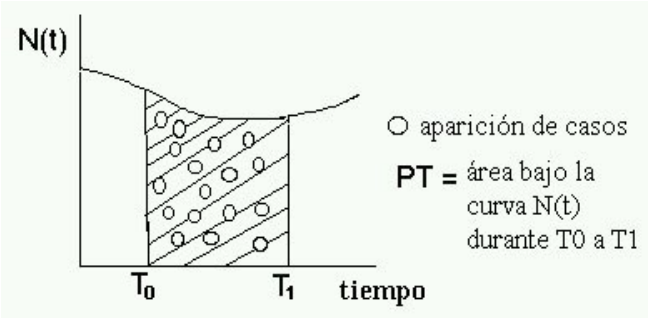
- En términos generales, la tasa es una medida de la rapidez con que sucede algo de interés.
- En epidemiología, la tasa es una medida de la rapidez con la que se desarrollan nuevos casos de una enfermedad o, en forma alternativa, qué tan rápido mueren las personas que padecen una enfermedad de interés.
- Una **tasa instantánea**, como la velocidad vectorial, describe con qué rapidez se presenta la enfermedad o la muerte en un momento del tiempo.
- Una **tasa promedio**, como la velocidad escalar, describe la rapidez con que se presenta la enfermedad o la muerte en promedio durante un período de tiempo.
- En epidemiología, normalmente utilizamos tasas promedio en lugar de tasas instantáneas.
- Las tasas deben medirse en unidades de tiempo.

Densidad de incidencia – Concepto

El término *densidad de incidencia* (DI) ha sido propuesto (Miettinen OS, Am J Epidemiol 1976; 103(2):226-35) para brindar una interpretación intuitiva del concepto de **tasa de incidencia promedio**.

El siguiente diagrama ilustra la densidad de incidencia como la concentración (es decir, la **densidad**) de la aparición de nuevos casos en una acumulación (o *conjunto*) de persona-tiempo. El valor **persona-tiempo (PT)** está representado por el área bajo la curva **N(t)** que describe la cantidad de personas sin la enfermedad en el momento t, durante el período de seguimiento desde el momento T_0 a T_1.

Cada caso nuevo está señalado mediante un pequeño círculo ubicado dentro del conjunto de persona-tiempo al momento de la aparición de la enfermedad. La concentración de círculos dentro del conjunto representa la **densidad de casos**. Cuanto más alta es la concentración, más elevada es la tasa promedio durante el período de seguimiento.



La tasa instantánea en epidemiología

Se ha utilizado una variedad de definiciones matemáticas para definir una **tasa**. En epidemiología, cuando la incidencia de una condición de salud resulta de interés, se utiliza comúnmente la siguiente definición:

$$Tasa = \lim_{\Delta t \to 0} \frac{\left[\#\text{ de casos nuevos } (t,\ t + \Delta t)\right] / \Delta t}{N(t)}$$

donde **t** es el tiempo, **Δt** denota un pequeño cambio del tiempo, y **N(t)** indica el tamaño de la **población de riesgo** (por ejemplo, la cohorte que no padece la enfermedad) en cualquier tiempo dado del momento **t**. Ésta es una definición general de una **tasa instantánea** y se aplica tanto a una cohorte fija como a las poblaciones dinámicas.

Esta definición puede interpretarse como el **potencial instantáneo en el momento t** (tal como lo define la declaración de límite) para la cantidad de casos nuevos que se desarrollarían entre los momentos t y t + Δt **por tiempo unitario relativo** a la población de riesgo en el momento **t**.

Una característica especial de esta definición es la participación de **N(t)**, que no se requiere en otros usos más conocidos del término *tasa* fuera de la epidemiología como, por ejemplo, cuando se describe la velocidad vectorial que se observa en el velocímetro de un automóvil. La velocidad vectorial implica cuánta distancia uno recorrería potencialmente por período de tiempo de viaje en un momento específico del tiempo. La distancia recorrida corresponde a la cantidad de casos nuevos desarrollados durante un período de tiempo Δt. Pero la interpretación de velocidad vectorial no involucra un término similar a N(t).

La tasa promedio en epidemiología

En epidemiología, cuando la incidencia de una condición de salud resulta de interés, se utiliza comúnmente la siguiente definición general de tasa promedio:

$$\text{Tasa Promedio} = \frac{\left[\#\text{ de casos nuevos en } (\mathbf{T_0}, \mathbf{T_1})\right]}{\mathbf{PT}}$$

donde $\mathbf{T_0}$ y $\mathbf{T_1}$ denotan los momentos de inicio y finalización del seguimiento y **PT** indica la cantidad de persona-tiempo sin la enfermedad acumulada durante el intervalo de tiempo desde $\mathbf{T_0}$ a $\mathbf{T_1}$. Matemáticamente, PT origina el área bajo la curva que describe cómo cambia N(t) en el tiempo entre los momentos $\mathbf{T_0}$ y $\mathbf{T_1}$, donde N(t) es el tamaño de la población de riesgo en el momento **t**. Técnicamente, la fórmula de PT está dada por:

$$PT = \int_{T_0}^{T_1} N(t)dt$$

Como un ejemplo simple del cálculo de PT, si estamos estudiando una población dinámica y estable del tamaño N desde el momento $\mathbf{T_0}$ a $\mathbf{T_1}$, entonces:

$$PT=N(T_1 - T_0)$$

En forma alternativa, si estamos estudiando una cohorte fija de tamaño N que no padece la enfermedad y conocemos los tiempos de seguimiento individual de cada sujeto en la cohorte, entonces:

$$PT = \sum_{i=1}^{N} \Delta T_1$$

a medida que i aumenta de 1 a N.

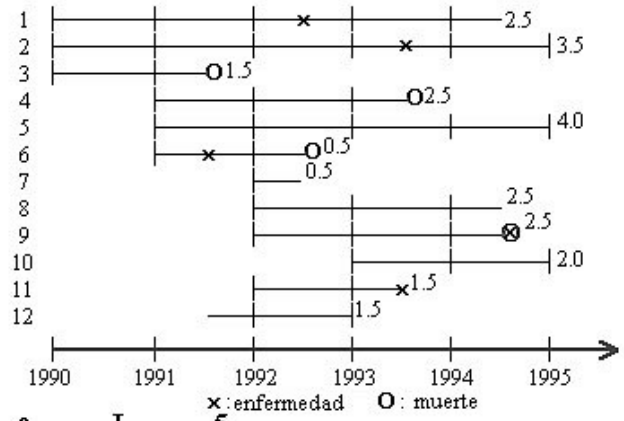
Cálculo de tasa

Para calcular una tasa, debemos hacer un seguimiento de una cohorte de individuos, contar la cantidad de casos nuevos (o incidentes), **I**, de una enfermedad en esa cohorte, y calcular el tiempo total, llamado persona-tiempo o **PT**, durante el cual se observa a los individuos sin la enfermedad en la cohorte durante el período de estudio. La tasa de incidencia estimada ($\hat{\text{TI}}$) se obtiene dividiendo **I** por **PT**:

$$\hat{\text{TI}} = \frac{\text{I}}{\text{PT}}$$

Esta fórmula brinda una tasa promedio, en lugar de una tasa instantánea, que es más difícil de calcular. La fórmula es lo suficientemente general como para ser utilizada para cualquier resultado de interés, incluso la muerte. Si el resultado es la muerte en lugar de la incidencia de la enfermedad, la fórmula brinda la **tasa de incidencia de la mortalidad** en lugar de la **tasa de incidencia de la enfermedad**.

Por ejemplo, consideremos nuevamente la siguiente cohorte hipotética de 12 individuos que inicialmente no padecen la enfermedad y a quienes se les hace un seguimiento durante un período de 5 años, de 1990 a 1995.



$$\hat{TI} = \frac{I}{PT} = \frac{5}{25\ PA} = 0.20$$

Para tener todos los datos, la tasa de incidencia estimada es 0,2 por persona-año.

De estos datos, la cantidad de casos nuevos es 5. El total de persona-tiempo, en este caso persona-años, se obtiene sumando los tiempos de seguimiento individual observados sin la enfermedad, que da un total de 25 personas-años. Por lo tanto, la tasa es 5 dividido 25 ó 0,20, que puede traducirse como 20 casos nuevos por 100 personas-años de seguimiento.

Preguntas de estudio (P4.7)

1. En este ejemplo, ¿el valor 0,20 es una proporción?
2. En este ejemplo, ¿el valor 0,20 representa el riesgo de desarrollar la enfermedad?
3. ¿Cuál de las siguientes tasas no equivale a una tasa de 20 casos nuevos por 100 personas-años?
   A. 5 casos nuevos por 25 años
   B. 40 casos nuevos por 200 personas-años
   C. 480 casos nuevos por 2.400 personas-meses
   D. 20 casos nuevos por 1.200 personas-meses

Resumen

- Una tasa se calcula utilizando la fórmula **I/PT**, donde **I** es la cantidad de casos incidentes y **PT** denota el valor persona-tiempo acumulado de seguimiento durante el período de estudio.
- Esta fórmula brinda una tasa promedio, en lugar de una tasa instantánea, que es más difícil de calcular.
- Una tasa es siempre mayor que cero y no tiene límite superior.
- La tasa siempre se expresa en unidades de persona-tiempo.
- Una tasa de 0,20 casos por persona-año es equivalente a 20 casos por 100 personas-años, como también a 20 casos por 1.200 personas-meses.

El supuesto del Big-Mac sobre persona-tiempo

Hemos visto que la fórmula general para calcular una **tasa promedio (R)** es:

$$R = \frac{I}{PT}$$

donde **I** es la cantidad de casos nuevos y **PT** es el valor persona-tiempo acumulado durante un período especificado de seguimiento. Cuando se dispone de los tiempos de seguimiento individual observados, PT se determina sumando estos tiempos individuales para todos los individuos N en la cohorte sin la enfermedad.

Por ejemplo, si se realiza un seguimiento individual de 100 personas durante 10 años, entonces PT=1.000 personas-años. Asimismo, si se realiza un seguimiento individual de 1.000 personas durante 1 año, entonces obtenemos que PT=1.000.

Un supuesto clave sobre PT es que ambas de estas situaciones brindan información de persona-tiempo equivalente. En otras palabras, la tasa correspondiente a un valor especificado de PT no debería estar afectada por la forma en que se obtiene el total de persona-tiempo. A este supuesto lo llamamos el **supuesto del Big-Mac** porque es similar a asumir que comer 50 hamburguesas en un restaurante de comidas rápidas, que cuestan $2,00 cada una, es equivalente a comer comida gourmet por $100 en el mejor restaurante de la ciudad.

Sin embargo, el supuesto del Big-Mac para PT no se cumplirá si el tiempo promedio entre la primera exposición y la detección de la enfermedad (es decir, la latencia) es mayor que el promedio de tiempo de seguimiento individual. En dicho caso, esperaríamos que la tasa sea menor en una cohorte grande que acumula la misma cantidad de PT que en una cohorte más pequeña con tiempos de seguimiento individual más largos.

Por ejemplo, si la latencia fuera de 2 años, esperaríamos una tasa extremadamente baja para 1.000 personas seguidas en forma individual durante un año, pero una tasa mucho más alta para 100 personas seguidas en forma individual durante 2 años. Los individuos en la cohorte más grande no serían seguidos durante un tiempo suficientemente prolongado como para que se den muchos casos nuevos.

Determinación de la información de persona-tiempo

Existen varias formas de determinar el denominador de persona-tiempo en la fórmula de tasa. Como se ejemplificó en la actividad anterior, cuando se dispone de tiempos de seguimiento individual para cada persona en la cohorte, el valor de persona-tiempo se calcula con la sumatoria (Σ) de los tiempos de seguimiento individual de toda la cohorte que no padece la enfermedad.

$$\hat{I}R = \frac{I}{PT}$$

Cuando no se dispone de los tiempos de seguimiento individual, un método para calcular la información de persona-tiempo utiliza la fórmula:

$$PT = N^* \times \Delta t$$

donde N* es el tamaño promedio de la cohorte que no padece la enfermedad durante el período de tiempo de estudio y Δt es la duración del período de estudio. Esta fórmula es especialmente útil si la cohorte de estudio es una población grande como, por ejemplo, una ciudad, donde la información de persona-tiempo individual sería muy difícil de obtener. Para dicha cohorte grande, también sería difícil excluir los casos existentes de la enfermedad al comienzo del período del estudio, así como determinar la cantidad de personas sin la enfermedad que no son seguidas durante todo el período del estudio.

Sin embargo, podría suceder que relativamente pocas personas en la población desarrollen la enfermedad.

Y podríamos suponer que la población es una cohorte dinámica y estable, es decir, que dicha población no sufre cambios demográficos importantes durante el período de interés. De ser así, el tamaño promedio de la cohorte sin la enfermedad puede calcularse por el tamaño de toda la población sobre la base de datos de censos disponibles cercanos al período del estudio, que es lo que hemos denotado como N* en nuestra fórmula de persona-tiempo.

Como ejemplo, suponga que se hace un seguimiento de una población estable de 100.000 hombres durante un período de 5 años, durante el cual se detectan 500 casos nuevos de cáncer de vejiga. El valor de personas-años acumulado para esta cohorte puede calcularse, entonces, como 100.000 multiplicado por 5, ó 500.000 personas-años. En consecuencia, la tasa de incidencia promedio para el período de 5 años está dada por 500 dividido por 500.000, ó 0,001 por año, o lo que equivale a 1 caso nuevo por 1.000 personas-años.

Ejemplo

$$\hat{IA} = \frac{I}{PT} = \frac{500}{500{,}000} = 0.001 \text{ por año}$$

$$PT = N^* \times \Delta t$$
$$= 100{,}000 \times 5 = 500{,}000 \text{ personas-años}$$

$N^* = 100{,}000$ hombres

$\Delta t = 5$ años

$I = 500$ casos

Resumen

- Existen formas alternativas para determinar la información de personas-años requerida en el denominador de una tasa cuando no se encuentran disponibles los tiempos de seguimiento individual.
- Un método utiliza la fórmula PT = N* x Δt, donde N* es el tamaño promedio de una cohorte dinámica y estable basado en datos de censos disponibles cercanos al tiempo cronológico del estudio, y Δt es el período de tiempo del estudio.
- Esta fórmula es útil si la cohorte de estudio es una población grande, para la cual sería muy difícil obtener la información de persona-tiempo individual.

Un tercer método para determinar PT

Aquí describimos otro método para determinar la información acumulada de persona-tiempo (**PT**) cuando no se encuentra disponible la información sobre los tiempos de seguimiento individual. Este enfoque permite cambiar el tiempo de incorporación de las personas que ingresan progresivamente en la cohorte después del comienzo del estudio. Suponga que conoce:

N, la cantidad de personas en riesgo al inicio del seguimiento,
W, la cantidad de abandonos durante el período de estudio,
D, la cantidad de muertes por otras enfermedades durante el período de estudio, e
I, la cantidad de casos nuevos de la enfermedad durante el período de estudio.

Entonces, la información de persona-tiempo se calcula mediante la fórmula:

$$PT = \left(N - \frac{W}{2} - \frac{D}{2} - \frac{I}{2} \right) \times \Delta t$$

donde **Δt** es la duración del estudio. Esta fórmula muestra el tamaño de la cohorte inicial menos la mitad de los

individuos a los que no se les efectuó seguimiento durante todo el período de riesgo.

Básicamente, esta fórmula muestra la **cantidad efectiva** de individuos en riesgo que producirían **I** casos nuevos de la enfermedad, si se pudiese hacer un seguimiento de todos los individuos durante todo el período. Los valores **W/2**, **D/2** y **I/2** se utilizan para suponer que el tiempo de seguimiento promedio para aquellos que no se siguieron durante todo el estudio ocurre en el punto medio del período de seguimiento.

Para ejemplificar este enfoque, considere otra vez el ejemplo hipotético descrito en la actividad anterior que involucra a 12 individuos inicialmente sin la enfermedad, a quienes se les hace un seguimiento durante un período de 5 años, de 1990 a 1995: suponga que no conocen los tiempos de seguimiento individual, sino que de los 12 individuos sin la enfermedad, 5 se retiraron del estudio, 2 murieron y a 5 se les diagnosticó la enfermedad. Entonces:

$N=12 \quad W=5 \quad D=2 \quad I=5 \quad y \quad \Delta t=5$

Sustituyendo estos valores en la fórmula, calculamos que PT será 30. Ya que hubo 5 casos nuevos, la tasa estimada es, entonces, 5 sobre 30 ó 0,17 por persona-año. Esto no se aleja tanto de la tasa estimada de 0,20 por persona-año que se obtiene si utilizamos los tiempos de seguimiento individual para calcular PT igual al valor correcto de 25.

Tasa de incidencia de la enfermedad de Parkinson

La enfermedad de Parkinson es una afección discapacitante que se caracteriza por temblor en reposo, rigidez, lentitud en los movimientos y alteración de los reflejos. Se hizo un seguimiento durante seis años de una cohorte de más de 6.500 ancianos holandeses que no padecían la enfermedad de Parkinson al comienzo del estudio, a fin de determinar la tasa de incidencia a la que se desarrollan casos nuevos de esta enfermedad. Durante el período de seguimiento, se les diagnosticó la enfermedad de Parkinson a 66 participantes.

Debido a que esta enfermedad tiene un comienzo sutil, fue difícil determinar exactamente cuándo había comenzado el proceso. Por lo tanto, los investigadores calcularon el momento de inicio como el punto medio entre el momento del diagnóstico y el último momento en el que se determinó que el participante no padecía Parkinson. Entonces, se pudo calcular la cantidad total de personas-años sin la enfermedad en este estudio mediante la suma de la cantidad de personas-años que cada uno de los 6.500 participantes había aportado al estudio hasta que:

1. desarrolló la enfermedad de Parkinson;
2. murió; o
3. llegó al final del período del estudio con vida y sin haber desarrollado la enfermedad de Parkinson.

Esto dio como resultado un total de 38.458 personas-años sin la enfermedad. En este estudio, la tasa de incidencia promedio de la enfermedad de Parkinson para el período de estudio de 6 años es:

66 / 38.458 = 0,0017 casos por persona-año

Esto significa que se desarrollan 1,7 casos nuevos de enfermedad de Parkinson por 1.000 personas-años.

Preguntas de estudio (P4.8)

1. Utilizando la fórmula PT = N* x (Δt), ¿qué cantidad de personas-años se habría computado para esta población de estudio si no se hubiera contado con la información detallada sobre el aporte de cada individuo a la cantidad total de personas-años?
2. Empleando la cantidad de personas-años de la pregunta anterior, ¿cuál es la tasa de incidencia?

Resumen

- Se hizo un seguimiento durante seis años de una cohorte de más de 6.500 ancianos holandeses que no padecían la enfermedad de Parkinson al comienzo del estudio, a fin de determinar la tasa a la que se desarrollan casos nuevos.
- Los resultados indican que se presentan 1,7 casos nuevos de enfermedad de Parkinson por cada 1.000 personas-años de seguimiento.
- Para el cálculo de personas-años, se utilizó la fórmula PT = N* x Δt, ya que no había información detallada de la cantidad de personas-años de cada individuo.

Auto-evaluación (P4.9)

Coloque **verdadero** o **falso** en cada una de las siguientes afirmaciones:

1. La tasa no es una proporción. . . . . **???**
2. La tasa tiene unidades de 1/persona-tiempo y varía de cero a uno. . . . **???**
3. Una tasa sólo puede calcularse si se hace un seguimiento individual de cada persona en una cohorte para contar y sumar la cantidad de persona-tiempo. . . . . **???**
4. Se puede calcular una tasa para una cohorte dinámica, pero no para una fija y estable. . **???**

Riesgo versus tasa

*La incidencia puede medirse como un **riesgo** o una **tasa**. El empleo de uno de estos tipos de medida es una elección importante cuando se planea un estudio epidemiológico.*

Hemos visto dos medidas distintas para cuantificar la frecuencia de la enfermedad: riesgo y tasa. El **riesgo** es una probabilidad, entre 0 y 1, que brinda la eventualidad de un cambio en el estatus de salud de un individuo durante un período determinado de seguimiento.

$0 \leq \text{Riesgo} \leq 1$

La **tasa** describe la rapidez con que se presentan eventos nuevos en una población. La **tasa instantánea**, que se calcula en raras ocasiones, se aplica a un punto fijo del tiempo mientras que la **tasa promedio** se aplica a un período de tiempo. Una tasa *no* es una probabilidad, nunca es negativa, pero no tiene límite superior, y se define en unidades de tiempo como, por ejemplo, años, meses o días.

$0 \leq \text{Tasa} \leq \times$

Cuando se planea un estudio epidemiológico, ¿qué medida queremos utilizar? ¿Riesgo o tasa? La elección depende del objetivo del estudio, el tipo de enfermedad que se considera, la naturaleza de la población de interés y la información disponible.

Si el objetivo del estudio es predecir un cambio en el estatus de salud de un individuo, entonces se necesita el riesgo. En especial, el riesgo es relevante para evaluar la prognosis de un paciente, seleccionar una estrategia de tratamiento apropiada y tomar decisiones personales sobre comportamientos relacionados con la salud tales como el tabaquismo, el ejercicio y la alimentación. Por el contrario, una tasa no tiene una interpretación útil a nivel individual.

Si el objetivo del estudio es probar una hipótesis específica sobre la etiología de la enfermedad, la opción puede ser tanto el riesgo como la tasa, según la naturaleza de la enfermedad y la forma en que observamos casos nuevos. Si la enfermedad es crónica y requiere un período largo de seguimiento para obtener una cantidad suficiente de casos, normalmente existirá una cantidad considerable de pérdidas en el seguimiento o abandonos del estudio. En consecuencia, los tiempos de seguimiento individual tienden a variar considerablemente. La tasa, en lugar del riesgo, puede solucionar este problema.

Sin embargo, si se considera una enfermedad aguda como, por ejemplo, un brote debido a un agente infeccioso, probablemente habrá pérdidas de seguimiento mínimas, y el riesgo, por lo tanto, puede calcularse directamente. Con una enfermedad aguda, no estamos tan interesados en la rapidez con que surge la enfermedad, ya que el período del estudio es relativamente corto. Más bien, estamos interesados en identificar el factor de origen que es el principal responsable del aumento del riesgo individual.

Si la población que se estudia es dinámica y grande, los tiempos de seguimiento individual, más allá de si se pueden obtener o no, variarán considerablemente para individuos diferentes, de forma que la tasa debe preferirse antes que el riesgo. Sin embargo, si no se dispone de tiempos de seguimiento individual, no se puede calcular ni siquiera una tasa, a menos que se suponga que el tamaño de la población es estable, la enfermedad es poco frecuente y se cuente con un cálculo de la población basado en un censo reciente.

Generalmente se prefiere el riesgo a la tasa porque es más fácil de interpretar. Sin embargo, muchas veces la tasa debe ser la medida de elección a causa de los problemas asociados al cálculo del riesgo.

Resumen

- El riesgo es la probabilidad de que un individuo desarrolle una enfermedad dada durante un período especificado de seguimiento.
- La tasa describe la rapidez con que se presentan eventos nuevos en una población.
- El riesgo debe estar entre 0 y 1 mientras que la tasa nunca es negativa, no tiene límite superior y se define en unidades de tiempo.
- Generalmente se prefiere el riesgo a la tasa porque es más fácil de interpretar.

- Muchas veces la tasa debe ser la medida de elección a causa de los problemas asociados al cálculo del riesgo.
- La elección de riesgo versus tasa depende del objetivo del estudio, el tipo de enfermedad, el tipo de población y la información disponible.
-

La mortalidad podría utilizarse en lugar de la incidencia de la enfermedad

Al igual que en el caso de las medidas de incidencia de enfermedad, las medidas de incidencia de mortalidad pueden tomar la forma de riesgo o tasa, según el diseño y los objetivos del estudio. Las medidas de mortalidad se describen en una actividad posterior (en la sección 4-4).

Auto-evaluación (P4.10)

Determine si las siguientes afirmaciones definen mejor una **tasa, riesgo** o **ambos**:

1. Más útil para la toma de decisiones individual. . . . . **???**
2. El numerador es la cantidad de casos nuevos durante un período de seguimiento. . **???**
3. El valor más bajo posible es cero. . . . . **???**
4. No hay límite superior. . . . . **???**
5. Puede expresarse como un porcentaje. . . . . **???**
6. Es mejor para estudios con períodos variables de seguimiento. . . . **???**
7. Se calcula tradicionalmente en la situación de brote agudo (seguimiento corto). . **???**
8. Mide la rapidez con que se presenta la enfermedad o muerte en una población. . **???**
9. Incidencia acumulada. . . . . **???**
10. Mide la aparición de la enfermedad en una población. . . . . **???**

4-4 Prevalencia y mortalidad

Prevalencia

La prevalencia mide los casos existentes de una enfermedad y es la característica principal del diseño de ***estudio transversal****. Hay dos tipos de prevalencia, la* ***prevalencia puntual****, que es la que se usa más comúnmente, y la* ***prevalencia de período****.*

En epidemiología, la prevalencia normalmente se refiere a la identificación de casos existentes de una enfermedad en una población y es la característica principal del diseño de estudios transversales. La prevalencia también puede referirse, en forma más amplia, a la identificación de personas con cualquier característica de interés, no necesariamente una enfermedad. Por ejemplo, podríamos querer considerar la prevalencia del tabaquismo, el estatus de inmunidad o el colesterol alto en una población.

La medida más común de la prevalencia es la **prevalencia puntual**, que se define como la probabilidad de que un individuo en una población sea un caso en el momento **t**.

$$\hat{P} = \frac{C}{N} = \frac{(\text{\# de casos observados en el momento t})}{(\text{Población en el momento t})}$$

*La prevalencia puntual se calcula como la proporción de personas en una población de estudio que padecen una enfermedad en un punto específico del tiempo (**C**). Por ejemplo, si hay 150 individuos en una población y, cierto día, 15 están enfermos con gripe, la prevalencia estimada para esta población es de 10%.*

$$\hat{P} = \frac{15}{150} = 10\%$$

Preguntas de estudio (P4.11)

1. ¿Es la prevalencia puntual una proporción?

2. Un estudio con un denominador grande o uno que involucre eventos poco frecuentes puede arrojar una prevalencia muy baja. Por ejemplo, suponga que 13 personas de una población de 406.245 personas tuvieron una enfermedad específica en el momento **t**. ¿Cuál es la prevalencia puntual de esta enfermedad en el momento **t**?

   A. 0,0032
   B. 32%
   C. 0,000032
   D. 0,0000032

3. ¿Cuál de las siguientes expresiones es equivalente a un valor de la prevalencia puntual de 0,000032?

   A. 3,2 por 1.000
   B. 3,2 por 100.000
   C. 32 por 100.000

Cuando se mide la prevalencia puntual, es esencial indicar *cuándo* se enumeraron los casos, especificando un día del calendario o un punto fijo en una secuencia de tiempo como, por ejemplo, el tercer día postoperatorio. Las medidas de prevalencia son muy útiles para evaluar el estatus de salud de una población y para planificar los servicios de salud. Esto se debe a que la cantidad de casos existentes en cualquier momento es un determinante de la demanda de tales servicios.

Sin embargo, para identificar los factores de riesgo, las medidas de prevalencia no son tan adecuadas como las medidas de incidencia como, por ejemplo, el riesgo o la tasa. Esto se debe a que la prevalencia se refiere sólo a los sobrevivientes, de forma que se ignoran los casos que murieron antes del momento en que se mide la prevalencia.

Prevalencia

Útil para:
Evaluar el estado de salud de una población.
Planificar los servicios de salud.

No es útil para:
Identificar factores de riesgo.

Resumen

- La prevalencia se refiere a casos existentes de una enfermedad en un punto o período de tiempo.
- Las mediciones de prevalencia se calculan principalmente a partir de encuestas transversales.
- La prevalencia puntual es la probabilidad de que un individuo en una población sea un caso en el momento **t**.
- La prevalencia puntual se calcula mediante la fórmula **P** = **C**/**N**, donde **C** es la cantidad existente de casos en el momento **t** y **N** es el tamaño de la población en el momento **t**.
- Las medidas de prevalencia son útiles para evaluar el estatus de salud de una población y para planificar los servicios de salud.
- Las medidas de prevalencia se refieren a los sobrevivientes, por lo que no son tan adecuadas para identificar factores de riesgo.

Prevalencia de período

Una medida alternativa a la **prevalencia puntual** es la **prevalencia de período (PP)**, que requiere el supuesto de una población dinámica y estable para el cálculo. La **PP** se calcula como la razón de la cantidad de personas **C*** en las que se observó la condición de salud (por ejemplo, una enfermedad) en cualquier momento durante un período de seguimiento determinado, por ejemplo, desde el momento $\mathbf{T_0}$ a $\mathbf{T_1}$, sobre el tamaño **N** de la población para este mismo período, es decir, la fórmula de prevalencia de período es:

$$PP = \frac{C*}{N} = \frac{C + I}{N}$$

donde **C** denota la cantidad de casos prevalentes en el momento $\mathbf{T_0}$ e **I** es la cantidad de casos incidentes que se desarrollan durante el período. Por ejemplo, si hiciéramos un seguimiento de una población de 150 personas durante un año, y 25 tuviesen una enfermedad de interés al comienzo del seguimiento y otros 15 casos nuevos se desarrollaran durante el año, la prevalencia de período para el año sería:

PP = (25 + 15)/ 150 = 0,27 ó 27%,

mientras que la prevalencia puntual estimada al inicio del período es:

P = 25/150 = 0,17 ó 17%

Y la incidencia acumulada estimada para el período de un año es:

IA = 15/125 = 0,12 ó 12%

Auto-evaluación (P4.12)



Coloque **verdadero** o **falso** en cada una de las siguientes afirmaciones:

1. La prevalencia es una medida más útil para la planificación de la salud que para la investigación etiológica. . . . . **???**

2. Al igual que la incidencia acumulada, la prevalencia es una proporción que puede oscilar entre cero y uno. . **???**

3. Las medidas de prevalencia derivan más comúnmente de estudios de seguimiento. . **???**

4. Mientras que la incidencia generalmente se refiere a la aparición de la enfermedad, lesión o muerte; la prevalencia puede referirse a la enfermedad, discapacidad, conductas, exposiciones y factores de riesgo genéticos. . . . . **???**

Seleccione la respuesta correcta:

5. La fórmula de la prevalencia puntual es:

    a. Cantidad de casos nuevos / cantidad de personas en la población
    b. Cantidad de casos nuevos / cantidad de personas que no padecían la enfermedad en el punto inicial de la observación
    c. Cantidad de casos nuevos / cantidad de personas-tiempo de seguimiento
    d. Cantidad de casos actuales / cantidad de personas en la población
    e. Cantidad de casos actuales / cantidad de personas que no padecían la enfermedad en el punto inicial de la observación
    f. Cantidad de casos actuales / cantidad de personas-tiempo de seguimiento

Mortalidad

Al igual que en el caso de las medidas de incidencia de la enfermedad, las medidas de incidencia de la mortalidad pueden tomar la forma de **riesgo** o **tasa**, según el diseño y los objetivos del estudio. El riesgo de mortalidad puede medirse de distintas formas, lo que incluye **el riesgo de mortalidad por enfermedad específica, el riesgo de mortalidad por cualquier causa** y **el riesgo de letalidad**. Para cada medida, puede utilizarse la fórmula de incidencia acumulada simple. Aquí, **I** denota la cantidad de muertes observadas durante un período específico de estudio en una cohorte inicial de tamaño **N**.

Riesgo de Mortalidad

Riesgo de mortalidad por enfermedad específica
Riesgo de mortalidad por cualquier causa
Riesgo de letalidad

$$IA = \frac{I \text{ (\# de muertes)}}{N \text{ (tamaño de la cohorte)}}$$

Preguntas de estudio (P4.13)

1. Para el **riesgo de mortalidad por enfermedad específica**, ¿qué representa **I** en la fórmula **IA=I/N**?
   ???

    A. La cantidad de muertes por cualquier causa.
    B. La cantidad de muertes a raíz de una enfermedad específica de interés.
    C. La cantidad de personas con una enfermedad específica.
    D. El tamaño de la cohorte inicial, independientemente de la condición de enfermedad.

Para calcular el riesgo de mortalidad por enfermedad específica, **I** es la cantidad de muertes a causa de la enfermedad específica de interés, y **N** es el tamaño de la cohorte inicial, independientemente de la condición de enfermedad.

Preguntas de estudio (P4.13) (continuación)

2. Para el **riesgo de mortalidad por cualquier causa**, ¿qué representa **I** en la fórmula **IA=I/N**? **???**

   A. La cantidad de muertes por cualquier causa.
   B. La cantidad de muertes a raíz de una enfermedad específica de interés.
   C. La cantidad de personas con una enfermedad específica.
   D. El tamaño de la cohorte inicial, independientemente de la condición de enfermedad.

I es la cantidad de muertes por cualquier causa, y **N** es el tamaño de la cohorte inicial, independientemente de la condición de enfermedad.

El **riesgo de letalidad** es la proporción de personas con una enfermedad específica que mueren a causa de ella durante el período de estudio.

Preguntas de estudio (P4.13) (continuación)

3. Para el riesgo de *letalidad*, ¿qué representa **I** en la fórmula **IA=I/N**? **???**

   A. La cantidad de muertes por cualquier causa.
   B. La cantidad de muertes a raíz de una enfermedad específica de interés.
   C. La cantidad de personas con una enfermedad específica.
   D. El tamaño de la cohorte inicial, independientemente de la condición de enfermedad.

I es la cantidad de personas que mueren a causa de una enfermedad dada, y **N** es la cantidad de personas con esta enfermedad en la cohorte inicial.

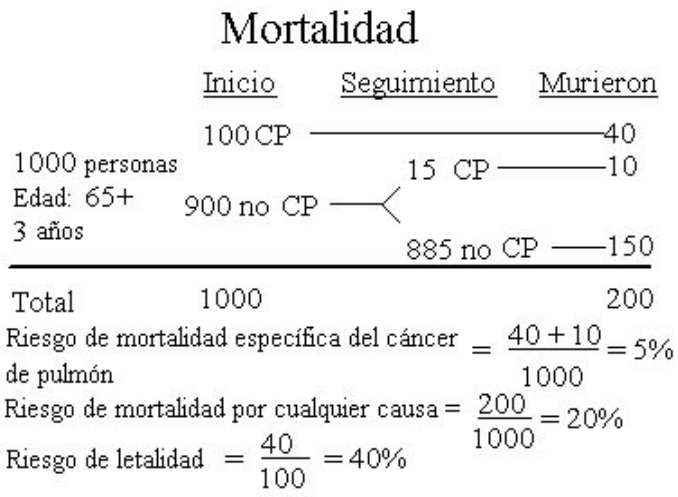
De forma similar, **la tasa de mortalidad** puede medirse mediante la fórmula general para la tasa promedio.

$$\text{IR} = \frac{\text{I}}{\text{PT}}$$

Aquí, **I** denota la cantidad de muertes observadas durante un período especificado de estudio en una cohorte inicial que acumula persona-tiempo **PT**. Para calcular la tasa de mortalidad por enfermedad específica, **PT** es la cantidad de persona-tiempo para la cohorte inicial, independientemente de la condición de enfermedad. Para estimar la **tasa de mortalidad por cualquier causa, PT** nuevamente es la cantidad de persona-tiempo para la cohorte inicial, independientemente de la condición de enfermedad.

Para calcular la **tasa de letalidad, PT** representa la cantidad de persona-tiempo para una cohorte inicial de personas con la enfermedad específica de interés *a quienes se efectúa el seguimiento para observar el estatus de mortalidad.*

Como ejemplo del cálculo de riesgo de mortalidad, suponga que observa una cohorte inicial de 1.000 personas de 65 años o mayor durante tres años. Cien de los 1.000 tenían cáncer de pulmón al inicio del seguimiento, y 40 de estos 100 murieron a causa del cáncer de pulmón. Además, 15 personas desarrollaron cáncer de pulmón durante el período de seguimiento y 10 murieron. De las restantes 885 personas sin cáncer de pulmón, 150 también murieron.



- El riesgo de mortalidad específica del cáncer de pulmón para esta cohorte es 50/1.000 ó 5%.
- El riesgo de mortalidad por cualquier causa es 200/1.000 ó 20%.
- El riesgo de letalidad para los 100 pacientes con cáncer de pulmón en la cohorte inicial es 40/100 ó 40%.

Preguntas de estudio (P4.14)

Para el ejemplo del cáncer de pulmón que se planteó recién, conteste las siguientes preguntas:

1. A partir de los datos, ¿cuál es el riesgo estimado para la incidencia del cáncer de pulmón durante el período de tres años?
2. ¿Por qué la incidencia estimada del cáncer de pulmón (CP) difiere de la mortalidad estimada del CP de 5%?
3. ¿Bajo qué circunstancias esperaría que la incidencia del CP y el riesgo de mortalidad del CP sean aproximadamente iguales?

Resumen

- Las medidas de incidencia de mortalidad pueden tomar la forma de riesgo o tasa, según el diseño y los objetivos del estudio.
- El riesgo o la tasa de mortalidad pueden medirse de distintas formas, lo que incluye el riesgo o tasa de mortalidad por enfermedad específica, el riesgo o tasa de mortalidad por cualquier causa y el riesgo o tasa de letalidad.
- Para medir el riesgo de mortalidad, puede utilizarse la fórmula de incidencia acumulada simple, es decir IA = I / N.
- De forma similar, la tasa de mortalidad puede medirse mediante la fórmula general de la tasa promedio, es decir, TI = I / PT.

Auto-evaluación (P4.15)

Durante los últimos dos años, murieron exactamente 2.000 residentes en una comunidad para jubilados con una población dinámica y estable de 10.000 personas.

1. Dados estos datos, la mejor opción para medir la mortalidad es la tasa de mortalidad. . **???**

2. Ya que la mortalidad muchas veces se expresa por 1.000, esta medida de mortalidad podría expresarse como 200 por 1.000 por año. . . . . **???**

3. El riesgo de mortalidad por enfermedad específica es la cantidad de muertes que se pueden atribuir a una enfermedad particular, dividido por la cantidad de personas con esa enfermedad. . . **???**

4. El denominador para el riesgo de mortalidad por cualquier causa y el riesgo de mortalidad por causa específica es el mismo. . . . . **???**

5. El denominador para el riesgo de letalidad es el numerador de la prevalencia de la enfermedad. **???**

Tasa ajustada por edad

La mayoría de los estudios epidemiológicos implican una comparación entre medidas de frecuencia de enfermedad para dos o más grupos. Por ejemplo, para estudiar el efecto de las condiciones climáticas sobre la mortalidad, podríamos comparar los riesgos o las tasas de mortalidad en dos o más lugares con climas distintos. Concentrémonos en dos estados de los Estados Unidos: Arizona y Alaska. Esto permitiría efectuar una comparación de la mortalidad en un clima frío y húmedo con la mortalidad en un clima caluroso y seco.

Las tasas crudas de mortalidad correspondientes a estos dos estados para el año 1996 fueron:

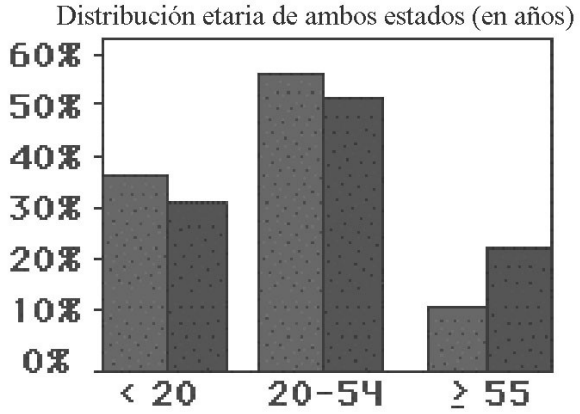
Alaska	426,57 muertes por 100.000 habitantes
Arizona	824,21 muertes por 100.000 habitantes

Le sorprendería saber, especialmente si considera los climas de los dos estados, que la tasa de mortalidad en Arizona casi duplica la de Alaska. ¿Eso significa que es mucho más peligroso vivir en Arizona que en Alaska?

Preguntas de estudio (P4.16)

1. ¿Qué cree usted? ¿Es mucho más peligroso vivir en Arizona que en Alaska?

Algunos conocimientos sobre la composición demográfica de estos dos estados podrían hacerle cuestionar dicha interpretación. Observe la distribución etaria de ambos estados:



(Nota: Alaska es la barra de la izquierda en cada grupo de barras, Arizona es la barra de la derecha).

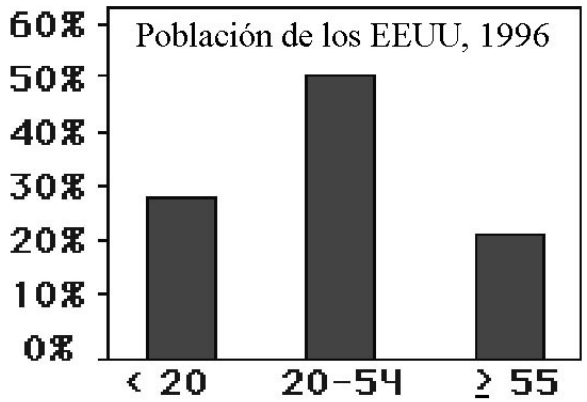
Preguntas de estudio (P4.16) (continuación)

2. ¿Qué población es más vieja?
3. ¿Por qué deberíamos esperar relativamente más muertes en Arizona que en Alaska?

La edad variable en esta situación se llama **factor de confusión** porque distorsiona la comparación de interés. Deberíamos corregir dicho efecto que potencialmente puede llevar a confusión. Un método popular para proceder a dicha corrección es el **ajuste de tasa**. Si el factor de confusión es la edad, este método se denomina generalmente **ajuste por edad**, y las tasas corregidas se denominan **tasas ajustadas por edad**.

El objetivo del ajuste por edad es modificar las tasas crudas de forma que cualquier diferencia entre las tasas de mortalidad de Alaska y Arizona no pueda explicarse por las diferentes distribuciones de edad en los dos estados. El método más popular de ajuste de tasa es el **método directo**. Este método obliga a realizar una comparación de las dos poblaciones sobre la base de una **distribución etaria común**. El factor de confusión edad se elimina recalculando las tasas mediante la sustitución de las distribuciones etarias separadas por una distribución etaria común. Luego se comparan las dos poblaciones como si tuvieran la misma estructura etaria.

*La distribución etaria común se determina identificando una **población estándar.** Una elección lógica aquí sería la población total de los Estados Unidos en 1996. También son posibles otras alternativas para la población estándar y, generalmente, no implicarán una diferencia muy importante en la comparación de tasas ajustadas.*



(Nota: Aquí no se muestra el cálculo real de las tasas ajustadas por edad. Para ver los detalles del cálculo de tasas ajustadas por edad para este ejemplo, en el CD, presione el botón de suplemento en la página de la lección del CD-ROM o vea el ejemplo al final de esta actividad).

Aquí se muestran las tasas de mortalidad ajustadas por edad, obtenidas mediante el método de ajuste directo con la población de los Estados Unidos en 1996 como estándar, junto con las tasas crudas.

Alaska	Arizona
Tasas de mortalidad ajustadas por edad	
856.00 / 100,000	832.21 / 100,000
Tasas crudas	
426.57 / 100,000	824.21 / 100,000

Cuando sacamos la edad como factor, la tasa de mortalidad ajustada por edad en Arizona es, en verdad, más baja que en Alaska.

Preguntas de estudio (P4.17)

1. ¿Cómo interpretamos estos nuevos resultados ajustados por edad?
2. Sobre la base de estos resultados, ¿cómo cree que es la distribución etaria de Alaska comparada con la de la población de los Estados Unidos en 1996?
3. ¿Cómo cree que es la distribución etaria de Arizona comparada con la de la población de los Estados Unidos en 1996?

Resumen

- La comparación de tasas crudas para dos o más grupos puede llevar a conclusiones erróneas, ya que dichas tasas no explican los efectos de los factores de confusión.
- Si el factor de confusión es la edad, este método se denomina generalmente ajuste por edad, y las tasas corregidas se denominan tasas ajustadas por edad.
- El objetivo del ajuste por edad es modificar las tasas crudas para que cualquier diferencia entre las tasas no pueda explicarse por la distribución etaria de los grupos comparativos.
- El método directo de ajuste por edad recalcula las tasas mediante la sustitución de las distribuciones etarias separadas de los grupos que se están comparando por una distribución etaria común.
- La distribución etaria común se determina identificando una población estándar.

Terminología sobre el ajuste

Las tasas descritas para Alaska y Arizona son, en verdad, **riesgos**. Intencionalmente, hemos utilizado el término **tasas** en este ejemplo para coincidir con la terminología utilizada normalmente en los informes/trabajos publicados que realizan **ajuste por edad**. En cualquier caso, el procedimiento empleado para el ajuste (por edad) puede aplicarse a cualquier medida de frecuencia de enfermedad: riesgo, tasa y/o prevalencia.

Además, también pueden ajustarse, tanto en forma individual como simultánea, los factores de confusión de interés potenciales distintos de la edad, por ejemplo, raza y sexo. Generalmente, utilizamos el término **ajuste de tasa** para describir un ajuste que implique cualquier tipo o cantidad de factores de confusión y cualquier tipo de medida de frecuencia de enfermedad, ya sea riesgo, tasa o prevalencia.

Ajuste por edad: ejemplo ya analizado

El método de **ajuste por edad** directo involucra los siguientes pasos:

1. Seleccione una **población estándar** cuya estructura etaria sea conocida. Por convención, la distribución estándar utilizada para el ajuste por edad de las tasas de mortalidad en los Estados Unidos es la distribución etaria de dicho país correspondiente al año más cercano al año de las tasas que se comparan.
2. Multiplique las tasas de mortalidad por edad específica para cada grupo que se está comparando, por la cantidad correspondiente de personas en la población estándar con la edad específica. El resultado es la cantidad de muertes esperadas en cada grupo.
3. Sume la cantidad de muertes esperadas dentro de cada grupo etario para obtener la cantidad total de muertes esperadas para cada grupo que se compara.
4. Divida la cantidad total de muertes esperadas en cada grupo por el tamaño total de la población estándar para obtener el resumen de las **tasas de mortalidad ajustadas por edad**.

Ejemplificamos el ajuste directo mediante la comparación de las tasas de mortalidad para Alaska y Arizona en 1996. Las tasas de edad específica y las tasas crudas generales para estos dos estados son las siguientes:

Tamaño de la población, tasas de mortalidad por cualquier causa y tasas de mortalidad por edad, Alaska y Arizona, 1996.

	Alaska 1996			Arizona 1996		
Edad	**n**	**m**	**t**	**n**	**m**	**t**
<1	10.037	72	717,34	75.322	575	763,38
1-4	40.445	18	44,50	290.256	127	43,75
5-9	54.359	12	22,07	344.886	67	19,42
10-14	52.437	14	26,69	328.220	95	28,94
15-19	49.475	53	107,12	313.422	322	102,73
20-24	44.690	60	134,25	294.762	372	126,20
25-34	84.864	137	161,43	657.439	1.022	155,45
35-44	116.015	238	205,14	684.967	1.700	248,18
45-54	81.857	306	373,82	509.569	2.271	445,67
55-64	40.162	359	893,87	347.841	3.632	1.044,15
65-74	20.668	518	2506,28	333.235	7.639	2.292,67
75-84	8.337	509	6.105,31	199.416	10.494	5.262,36
85+	1.947	286	14.689,26	55.929	8.240	14.731,96
Total/crudas	605.285	2.582	426,57	4.435.264	36.556	824,21

n = cantidad de personas, **m** = cantidad de muertes, y **t** = (d/n) x 100.000; es decir, muertes por 100.000 personas.

Paso 1. A continuación se brinda la distribución etaria de la población estándar (Estados Unidos en 1996), junto con los porcentajes de la distribución etaria para Alaska y Arizona. Observe que la distribución etaria de Arizona es bastante similar a la distribución etaria de los Estados Unidos, mientras que la de Alaska es un poco distinta, con un porcentaje pequeño en las edades más avanzadas.

Ajuste por edad: ejemplo ya analizado (continuación)
Distribución etaria de los Estados Unidos, Alaska y Arizona, 1996.

	Estados Unidos 1996		Alaska 1996	Arizona 1996
Edad	**n**	**%**	**%**	**%**
<1	3.891.494	1,5	1,6	1,6
1-4	15.516.482	5,8	6,7	6,5
5-9	19.441.182	7,3	9,9	7,8
10-14	18.981.045	7,2	8,7	7,4
15-19	18.662.151	7,0	8,1	7,1
20-24	17.559.730	6,6	7,4	6,6
25-34	40.368.234	15,2	14,0	14,8
35-44	43.393.341	16,4	19,2	15,4
45-54	32.369.791	12,2	13,5	11,4
55-64	21.361.460	8,5	6,6	7,8
65-74	18.669.337	7,0	3,4	7,5
75-84	11.429.984	4,3	1,4	4,5
85+	3.761.561	1,4	0,3	1,3
Crudas/total	265.405.792	100,0	100,0	100,0

La idea básica al calcular una tasa ajustada directa para un estado dado, por ejemplo, Alaska, es calcular cuál sería la tasa cruda hipotética para Alaska si tuviese la misma estructura etaria que la población estándar (Estados Unidos en 1996). Ya que ni Alaska ni Arizona tienen realmente la misma estructura etaria que los Estados Unidos, sus tasas ajustadas (tomando a los Estados Unidos como estándar) son, en verdad, hipotéticas, pero al menos ahora pueden compararse porque se utiliza el mismo estándar para ambos estados.

Paso 2. La cantidad de muertes esperadas para un grupo etario dado en Alaska se obtiene multiplicando el tamaño de la población estándar para ese grupo etario por la tasa de mortalidad según edad específica para Alaska. Por ejemplo, para el grupo etario 25-34, debemos multiplicar 40.368.234 (es decir, la población de los Estados Unidos para las edades 25-34) por 161,43/100.000 (es decir, la tasa de mortalidad en Alaska para las edades 25-34), lo cual da 65.166,4 muertes esperadas en Alaska para este grupo etario. La respectiva cantidad de muertes esperadas para Arizona se calcula multiplicando 40.368.234 por 155,45/100.000 (es decir, la tasa de mortalidad en Arizona para las edades 25-34), lo cual da como resultado 62.752,4 muertes esperadas en Arizona para este grupo etario.

Paso 3. Luego debemos sumar la cantidad de muertes esperadas de todos los grupos etarios para Alaska y Arizona en forma separada, a fin de obtener la cantidad total de muertes esperadas para cada estado. Sin mostrar los cálculos, la suma de estos valores esperados es igual a 2.271.873,6 para Alaska y 2.208.441,6 para Arizona.

Paso 4. Finalmente, para cada estado en forma separada, dividimos la cantidad total de muertes esperadas por el tamaño total de la población estándar (es decir, 265.405.792) para llegar a las tasas ajustadas para cada estado. Así obtenemos (2.271.873,6/265.405.792) = 856,0/100.000 para Alaska y (2.208.441/265.405.792) = 832,1/100.000 para Arizona.

Resumiendo las tasas crudas y ajustadas (por 100.000 personas) para cada estado, observamos lo siguiente:

	Cruda	**Ajustada**
Alaska	426,6	856,0
Arizona	824,2	832,1

La tasa ajustada para Alaska es más alta que la tasa ajustada para Arizona, mientras que la tasa cruda para Alaska es menor que la tasa cruda para Arizona.

El método indirecto de ajuste y la SMR

Una **tasa ajustada (por edad) indirecta** es:

- un promedio ponderado de tasas específicas (por edad) correspondientes a una población estándar seleccionada utilizando la distribución de la población del estudio como ponderaciones;
- se utilizan normalmente cuando alguna de las tasas específicas (por edad) correspondientes a la población del estudio no está disponible o no es confiable.

Para el ajuste indirecto se necesitan:

1. las tasas específicas para la población estándar seleccionada;
2. la distribución de la población del estudio en los mismos estratos utilizados para calcular las tasas específicas de la población estándar;
3. la tasa cruda para la población del estudio;
4. la tasa cruda para la población estándar.

La **SMR** (sigla inglesa que significa **razón de morbilidad/mortalidad estandarizada**) es una medida de razón definida por:

- O/E = cantidad observada (O) de casos de enfermedad en la población del estudio dividida por la cantidad esperada (E); generalmente, esta razón se multiplica por 100.
- La cantidad esperada se calcula aplicando tasas de enfermedad estándar por edad específica basadas en una población estándar (por ejemplo, las tasas nacionales) a la distribución etaria de la cohorte.

Cálculo de una SMR: ejemplo hipotético

	Población del estudio:		**Población estándar:**		
Edad	**Casos observados (O) en la cohorte**	**Personas-años**	**Tasa por 1.000 personas-años**	**Cantidad de casos esperados (E)**	**(O/E) x 100**
30-39	2	350	1,4	0,5	400
40-49	8	2.375	2,5	5,9	153
50-59	88	4.535	11,4	51,7	170
60-69	111	1.349	23,5	31,7	350
Total	**210**	8.609		**89,8**	**234**

Ventajas de la SMR

- Utilizada ampliamente en los estudios ocupacionales.
- No es necesario identificar, reclutar y hacer un seguimiento de un grupo sin exposición (de referencia). (Simplemente, se pueden utilizar las tasas de enfermedad por edad específica nacionales o estatales que se encuentren disponibles).
- Las tasas de enfermedad por edad específica nacionales o estatales son estables, porque se basan en poblaciones grandes.

Desventajas de la SMR

- Quizás no puedan compararse las SMR de cohortes diferentes (efecto de confusión residual por edad debido a que las tasas de enfermedad estándar se aplican a cohortes con diferentes distribuciones etarias).
- La SMR (como cualquier medida de resumen) puede ocultar los efectos de edad específica. En el ejemplo hipotético, el efecto en el grupo de edad 60-69 era aproximadamente el doble que el efecto en los grupos de edad 40-49 y 50-59.

Auto-evaluación (P4.18)

Coloque **verdadero** o **falso** en cada una de las siguientes afirmaciones:

1. El ajuste por edad es un método para eliminar las disparidades de edad entre dos poblaciones. **???**
2. El ajuste por edad siempre acerca dos tasas dispares. . . . . **???**
3. Para el ajuste por edad, deberíamos emplear la población de los Estados Unidos como población estándar. . . . . **???**
4. El ajuste por edad puede utilizarse para una, dos o muchas tasas. . . . . **???**
5. Si las distribuciones etarias de dos poblaciones son muy similares, sus tasas ajustadas por edad también lo serán. . . . . **???**
6. Si las distribuciones etarias de dos poblaciones son muy similares, la comparación de las tasas ajustadas por edad no será muy diferente de la comparación de las tasas crudas. . . . **???**

A principios de la década de los noventa, se incluyeron a 7.983 ancianos holandeses en un estudio de cohorte prospectivo. Los investigadores calcularon la cantidad de personas-años que cada participante había aportado al estudio hasta enero de 2000. El total fue de 52.137 personas-años. Durante el seguimiento, 2.294 de los participantes murieron, y de éstos, 477 fallecieron por enfermedad coronaria.

7. ¿Cuál es la tasa de mortalidad por cualquier causa en esta población? **???** por 1.000 personas-años.
8. ¿Cuál es la tasa de mortalidad específica para la enfermedad coronaria? **???** por 1.000 personas-años.

Opciones
2294 **44** **477** **9,1**

La tasa cruda de mortalidad por cualquier causa para los hombres fue de 47,4 por 1.000 personas-años (PA) y para las mujeres fue de 41,9 por 1.000 personas-años. Después de hacer que la distribución etaria de las mujeres sea comparable con la distribución etaria de los hombres (estandarizando las tasas mediante el uso de la distribución etaria de los hombres), la tasa de mortalidad para las mujeres fue de sólo 27,8 por 1.000 PA.

9. Sobre la base de estas cifras, las mujeres deben ser considerablemente **???** que los hombres en esta población.

Opciones
mayores **menores**

Nomenclatura

C	Cantidad de casos prevalentes en el momento T
C*	C + I (cantidad de casos prevalentes en el momento T más casos incidentes durante el período del estudio)
IA	Incidencia acumulada (“riesgo”): IA=I/N
D	Duración de la enfermedad
I	Incidencia
TI	Tasa de incidencia (“tasa”): TI=I/PT
N	Tamaño de la población en estudio
P	Prevalencia: P=C/N
PP	Prevalencia de período: PP=C*/N

PT	Persona-tiempo
R	Tasa promedio
T o **t**	Tiempo

Referencias

Referencias generales

Greenberg RS, Daniels SR, Flanders WD, Eley JW, Boring JR. *Medical Epidemiology* (3° edición). Lange Medical Books, Nueva York, 2001.

Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic Research: Principles and Quantitative Methods*. John Wiley and Sons Publishers, Nueva York, 1982.

Ulm K. *A simple method to calculate the confidence interval of a standardized mortality ratio (SMR)*. Am J Epidemiol 1990; 131(2):373-5.

Referencias sobre tasas

Miettinen O. *Estimability and estimation in case-referent studies*. Am J Epidemiol 1976; 103(2):226-35.

Giesbergen PCLM, de Rijk MC, van Swieten JC, et al. *Incidence of parkinsonism and parkinson's disease in a general population: the Rotterdam Study*. Am J Epidemiol (en prensa).

Referencias sobre ajuste por edad

Dawson-Saunders B, Trapp RG. *Basic and Clinical Biostatistics*, 2nd ed., Appleton and Lange, Stamford, CN, 1994.

Woodward M. *Epidemiology: Study Design and Analysis*, Capítulo 4, págs. 157-167, Chapman and Hall, Boca Raton, FL, 1999.

Referencia para prevalencia e incidencia del VIH:

Horsburgh CR, Jarvis JQ, McArthur T, Ignacio T, Stock P. *Serconversion to human immunodeficiency virus in prison inmates*. Am J Public Hlth 1990; 80(2):209-210.

Ejercicios

EPI-1. Medidas de frecuencia de enfermedad

1. ¿Cuál es el propósito de la medida de frecuencia de enfermedad?
2. ¿Cuál es la diferencia entre incidencia y prevalencia?
3. ¿Cómo están interrelacionadas la incidencia y la prevalencia?
4. ¿Cuál es la diferencia entre incidencia acumulada y densidad de incidencia?
5. ¿Qué significa decir que el riesgo en 2 años de una persona es de 0,03?
6. ¿Qué significa afirmar que la tasa en determinada población es de 0,03/año?
7. ¿En qué circunstancias (de diseño) sería conveniente medir el riesgo?
8. ¿En qué circunstancias (de diseño) sería ventajoso medir una tasa?
9. ¿Por qué realizamos ajustes por edad de los riesgos o tasas?
10. ¿Cómo funciona el método directo de ajuste por edad?

EPI-2. Persona-tiempo

¿Cuáles son dos formas de calcular persona-tiempo en el cálculo de una tasa (es decir, la densidad de incidencia)? ¿En qué circunstancias usaría cada fórmula? Describa un ejemplo del empleo de cada fórmula.

EPI-3. Incidencia versus prevalencia

Determine si cada una de las siguientes afirmaciones requiere la medición de la INCIDENCIA o PREVALENCIA.

a. Una nueva vacuna oral, que supuestamente previene el cólera, fue incorporada en cierto distrito de salud. El funcionario de salud del distrito quiere monitorear una medida adecuada para determinar si la vacuna está funcionando.
b. Una psicóloga escolar quiere determinar si hay una asociación entre la lectura de material pornográfico y la violencia sexual en los adolescentes. La psicóloga logra recabar datos de entrevistas sobre la cantidad de pornografía que se lee regularmente y la cantidad de encuentros sexuales violentos que experimentaron los estudiantes.
c. Una organización para el mantenimiento de la salud está considerando ofrecer una clínica diabética orientada a la comunidad. Será necesario determinar cuántos pacientes estarían interesados en utilizar el servicio.
d. Una empresa farmacéutica ha desarrollado una nueva droga que supuestamente cura el asma. La empresa quiere controlar la efectividad del producto.
e. Una enfermera partera decide examinar la relación que existe entre los partos en el hogar y la infección posparto. Logra hacer un seguimiento de un grupo de mujeres durante el embarazo y la primera semana después del nacimiento de sus hijos.
f. *Quaker Oats* tiene una campaña publicitaria que afirma que una dieta rica en cereales ayuda a prevenir el cáncer de colon. Un epidemiólogo quiere evaluar la validez de esta afirmación.
g. Una empresa está considerando un nuevo programa para dejar de fumar en el lugar de trabajo. Se distribuye un cuestionario entre los empleados para determinar cuántas personas estarían interesadas en formar parte de dicho programa.
h. Se informa a los administradores escolares que el sistema escolar de determinado estado está obligado por ley a brindar clases de educación especial para todos los niños de escuelas públicas con discapacidades de aprendizaje. La junta quiere calcular cuántos profesores de educación especial necesitarán contratar para cumplir con esta obligación.
i. Un investigador tiene interés en evaluar si las mujeres embarazadas expuestas al humo de tabaco ambiental tienen más probabilidad de dar a luz bebés con bajo peso.

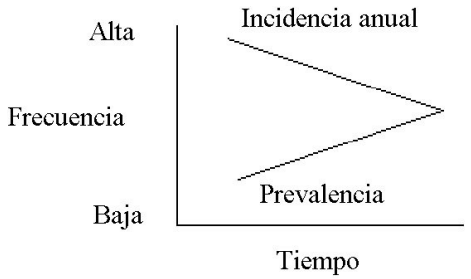
EPI-4. Incidencia y prevalencia: VIH

Un estudio publicado en 1990 (**Amer. J. Pub. Health 80**: págs. 209-10) investigó la presencia de infección por VIH entre los presos de Nevada. De 1.105 presos a los que se les hizo el examen de VIH cuando entraron en el sistema penitenciario, 36 resultaron estar infectados. Se hizo un seguimiento de todos los presos no infectados durante un total de 1.207 personas-años y se les volvió a realizar el examen de VIH cuando eran liberados de prisión. Se determinó que dos de los presos no infectados anteriormente contrajeron infección por VIH. Suponiendo que los dos presos se infectaron durante el período que pasaron en prisión:

a. Sobre la base de la información anterior, calcule la tasa de incidencia de la infección por VIH entre los presos de las cárceles de Nevada.
b. Exprese la tasa de incidencia calculada en el punto anterior, en términos de casos por 1.000 personas-años.
c. ¿Por qué no puede obtener un cálculo del riesgo basándose en la información suministrada?
d. ¿Por qué calcular el riesgo para estos datos podría ser inadecuado?
e. Calcule la prevalencia de la infección por VIH entre los presos que se incorporan al estudio efectuado en Nevada.
f. ¿Por qué la prevalencia calculada en el punto anterior no es necesariamente igual a una medida adecuada de riesgo que podría calcularse para estos datos?

EPI-5. Interpretación de la incidencia y la prevalencia

El siguiente gráfico indica las tasas variables de incidencia y prevalencia para la enfermedad "X" durante un período de tiempo:



Indique si cada una de las siguientes afirmaciones es coherente (sí o no) con la información representada en el gráfico.

a. Las personas que contraen esta enfermedad se curan más rápidamente.
b. Los esfuerzos para prevenir esta enfermedad parecen tener éxito.
c. La enfermedad se vuelve más crónica con el tiempo.

EPI-6. Calcule medidas de frecuencia de enfermedad

Los siguientes datos fueron obtenidos en un estudio en el que se realizó un seguimiento de 1.000 enfermeras durante 20 años para examinar la hipótesis de que el uso de ciertas píldoras para adelgazar era un factor de riesgo de ataque cardíaco.

	Uso de píldoras para adelgazar		
Ataque cardíaco	Sí	No	Total
Sí	30	11	41
No	470	489	959
Total	500	500	1.000

a. Calcule la cantidad de mujeres-años contribuidos por el grupo no expuesto.
b. ¿Qué información necesitaría para obtener un mejor cálculo de la cantidad de mujeres-años?
c. ¿Cuál es el riesgo en 20 años de sufrir un ataque cardíaco entre las que utilizaron píldoras para adelgazar?
d. ¿Cuál es la prevalencia del uso de píldoras para adelgazar entre las que no tuvieron ataques cardíacos?
e. ¿Sugieren los datos que el uso de píldoras para adelgazar es un factor de ataque cardíaco? Desarrolle su respuesta.

EPI-7. Ejercicio físico versus enfermedad coronaria

Un grupo de epidemiólogos estaba interesado en investigar la relación entre el ejercicio físico y el desarrollo de enfermedad coronaria (EC) entre mujeres. Se realizó una encuesta en una población sana de mujeres de entre 35 y 75 años, a fin de evaluar sus hábitos en relación con el ejercicio físico. Luego se les hizo un seguimiento durante 15 años para determinar la incidencia de la EC. Estos son los resultados:

	Frecuencia de ejercicio – veces por semana		
	Dos veces	Una vez	No realiza ejercicio
Enfermedad coronaria	4	40	23
Personas-años	25.111	117.205	32.843
Tasa por 10.000 personas-años	_____	_____	_____

a. ¿Qué proporción de mujeres que desarrollaron EC hacían ejercicio una vez por semana?
b. Complete el cuadro calculando las tasas por 10.000 y completando las tres celdas vacías. Exprese las respuestas con dos decimales.
c. ¿Puede sacar una conclusión a partir de estos datos sobre la relación entre el ejercicio físico y la EC?

EPI-8. Tasas estandarizadas: hipertensión

Un investigador está interesado en comparar las tasas de hipertensión de dos poblaciones. ¿Cuál de los siguientes aspectos deberían considerarse cuando se decide si es necesario estandarizar las tasas por raza? (Es posible que haya más de una respuesta correcta).

a. Si la tasa de hipertensión difiere según la raza.
b. Si la distribución racial difiere en las dos poblaciones.
c. Si la tasa de hipertensión difiere en las dos poblaciones.
d. La tasa de hipertensión en la población estándar.

EPI-9. Tasas y ajuste de tasas

Para cada una de las siguientes afirmaciones, indique si es verdadera o falsa.

a. Dos poblaciones con las mismas tasas de mortalidad por edad específica podrían tener diferentes tasas crudas o brutas de mortalidad.
b. Dos poblaciones con las mismas tasas crudas de mortalidad podrían tener diferentes tasas de mortalidad por edad específica.
c. El proceso de ajuste directo de tasas utiliza tasas por estrato específico basadas en la población estándar.
d. Una tasa cruda es un promedio ponderado de las tasas por estrato específico.

EPI-10. Ajuste de tasas: poblaciones estándar

Utilice los datos que se brindan a continuación y complete lo siguiente, expresando todos los cálculos con un decimal:

a. Obtenga las tasas totales de incidencia de la leucemia ajustadas por edad en los condados de Mesa y Weld, utilizando su población completa como estándar.
b. Calcule las tasas totales de incidencia de la leucemia ajustadas por edad en los condados de Mesa y Weld, utilizando la población de Colorado en 1970 (expresada en porcentajes) como estándar.
c. ¿Las tasas ajustadas por edad para cada condado en los ítems a y b anteriores son iguales?
d. ¿Podría elegirse una población estándar de forma que la tasa de incidencia ajustada por edad para el condado de Weld sea más alta que la tasa de incidencia ajustada por edad para el condado de Mesa?
e. Independientemente de la población estándar utilizada anteriormente, la tasa ajustada por edad para el condado de Weld es similar a la tasa sin ajuste (es decir, cruda). ¿Qué conclusión puede sacar de esto?
    - El ajuste por edad fue necesario sólo para el condado de Mesa, no para el condado de Weld.
    - Las tasas de incidencia de leucemia son similares en el condado de Weld y en la población estándar.
    - La estructura etaria es similar en el condado de Weld y la población estándar.

		Condado de Weld		Condado de Mesa	
Grupo etario	Colorado 1970		Leucemia		Leucemia
	Población (%)	Población de 1970	TIAP*	Población de 1970	TIAP*
< 5	8,4	7.491	9,5	3.754	7,6
5-19	30,6	28.452	3,0	16.852	4,2
20-34	22,5	20.382	1,4	9.253	4,6
35-49	17,2	13.859	1,0	9.329	3,1
50-64	12,8	11.219	6,4	8.685	11,5
65+	8,5	7.894	12,7	6.501	43,9
Total	100,0	89.297	4,2	54.374	10,2

* Tasas anuales promedio de incidencia de leucemia por población de 100.000 para el intervalo 1970-76, basadas en la población de 1970.

Respuestas a las preguntas de estudio y auto-evaluaciones

P4.1

1. Incidencia: aquí nos interesa la cantidad de casos nuevos después de comer ensalada de papa.
2. Prevalencia: aquí nos interesa la cantidad de casos existentes.
3. Incidencia: aquí nos interesa la cantidad de casos nuevos que se presentan durante el seguimiento.
4. Incidencia: aquí nos interesa la cantidad de muertes nuevas atribuidas al huracán.
5. Prevalencia: aquí nos interesa la cantidad real de niños inmunes al sarampión.
6. Incidencia: debido a que la rabia es de corta duración, se podría esperar que la prevalencia en un día en particular fuese baja con relación a la incidencia.
7. Prevalencia: la incidencia de esclerosis múltiple sería baja, pero como tiene una larga duración, se podría esperar que la prevalencia fuese más alta.
8. Incidencia: la incidencia de la gripe sería alta, pero ya que es de corta duración, la prevalencia sería baja.
9. Incidencia: ya que la duración de la dermatitis por hiedra venenosa es relativamente corta, la prevalencia sería baja y, como se presenta en forma frecuente, la incidencia sería alta.
10. Prevalencia: dado que la hipertensión arterial es frecuente y de larga duración, tanto la

incidencia como la prevalencia serían altas; sin embargo, la prevalencia sería más alta.

P4.2

1. La afirmación significa que un hombre de 45 años que no padece cáncer de próstata tiene una probabilidad de 0,05 de desarrollar ese tipo de cáncer durante los próximos 10 años, si no muere por cualquier otra causa durante el período de seguimiento.
2. Menor, porque el riesgo en 5 años implica un período más corto para que la misma persona desarrolle cáncer de próstata.

P4.3

1. No, los individuos deberían contarse como casos nuevos si estuviesen sanos al comienzo del seguimiento y se convirtiesen en un caso en algún momento durante el período de seguimiento especificado.
2. Sí, hay un problema, ya que un individuo seguido durante 2 años no tiene la misma oportunidad para desarrollar la enfermedad que un individuo al que se le hace seguimiento durante 4 años.
3. No, pero tenemos que suponer que aquellos individuos que no desarrollan la enfermedad tienen la misma cantidad de tiempo de seguimiento. De lo contrario, podemos obtener un cálculo erróneo de la IA porque no todos los individuos tendrán la misma oportunidad de desarrollar la enfermedad durante el período de seguimiento.

P4.4

1. Dinámica
2. Individuo 2
3. Individuo 7
4. Individuo 9
5. Individuo 3
6. Individuo 5
7. 5/12=42%
8. 4/12=33%
9. D
10. No
11. No
12. Sí
13. Sí
14. Sí
15. No
16. Sí

P4.5

1. 30/97
2. 60/97
3. 90/95
4. Verdadero. El numerador de la fórmula de IA es un subgrupo del denominador.
5. Verdadero. Debido a que el período de incubación es corto, no es probable que se pierda el seguimiento de los individuos.
6. Verdadero. El período de incubación prolongado significa que es probable que se pierda el seguimiento de los individuos y, por lo tanto, es posible que no se detecten casos. Para una cohorte dinámica, el denominador en la fórmula de IA no refleja el tamaño de la población en constante cambio.
7. Falso. La IA calculada subestimará el verdadero riesgo de la enfermedad.

P4.6

1. C
2.
    a. Sí
    b. Sí
    c. No
    d. Sí
    e. No

P4.7

1. No, el denominador 25 no describe 25 personas, sino el tiempo de seguimiento acumulado para 12 personas.
2. No, el riesgo en este ejemplo se calcularía como 5/12 ó 0,42. Sin embargo, utilizar el riesgo sería cuestionable aquí, ya que distintos individuos tienen diferentes tiempos de seguimiento.
3. C

P4.8

1. N* es el tamaño promedio de la cohorte que no padece la enfermedad y Δt es la duración del período del estudio. Por lo tanto, un cálculo aproximado de la cantidad total de personas-

años aportados por el estudio es 6.500 *6 = 39.000 personas-años.
2. La tasa de incidencia es 66/39.000 = 0,0017 ó 1,7 por 1.000 personas-años.

P4.9

1. Verdadero. Para las preguntas 1 y 2: la tasa puede oscilar entre 0 e infinito, mientras que el riesgo (que es una proporción) oscila entre 0 y 1 (ó 0% y 100%).
2. Falso
3. Falso. Hay formas alternativas de calcular la información de persona-tiempo cuando no está disponible el tiempo de seguimiento individual.
4. Falso. Una tasa puede calcularse para una cohorte dinámica o una fija, según la información de persona-tiempo disponible.

P4.10

1. Riesgo
2. Ambos
3. Ambos
4. Tasa
5. Riesgo
6. Tasa
7. Riesgo
8. Tasa
9. Riesgo
10. Ambos

P4.11

1. Sí, su valor puede oscilar entre 0 y 1 y generalmente se expresa como un porcentaje.
2. C. La prevalencia de la enfermedad es 13/406.245 = 0,000032
3. B. 3,2 por 100.000 es una expresión equivalente y es más fácil de interpretar.

P4.12

1. Verdadero. La prevalencia considera los casos existentes en lugar de los casos incidentes.
2. Verdadero. Ya que el numerador está contenido en el denominador, la prevalencia es una proporción y debe oscilar entre 0 y 1 (ó 0% y 100%).
3. Falso. Los estudios transversales se realizan esencialmente en un punto único (o corto) en el tiempo.
4. Verdadero. La prevalencia puede referirse a un resultado de salud o cualquier otra característica de un individuo.
5. d

P4.13

1. B
2. A
3. B

P4.14

1. El cálculo de la incidencia del CP se calcula como IA = 15/900 = 0,017 ó 1,7%.
2. El cálculo de mortalidad igual al 5% cuenta los 40 casos prevalentes de CP y no cuenta los 5 casos nuevos de CP que no murieron. Además, los denominadores son diferentes.
3. La incidencia del CP y los riesgos de mortalidad serían aproximadamente iguales si la enfermedad fuese rápidamente letal, de forma que se presentarían pocos casos prevalentes en la cohorte inicial, si los hubiera, y todos los casos nuevos habrían muerto antes del final del seguimiento.

P4.15

1. Verdadero
2. Verdadero
3. Falso. El denominador del riesgo de mortalidad por enfermedad específica es el tamaño de la cohorte inicial, independientemente de la condición de enfermedad.
4. Verdadero
5. Verdadero

P4.16

1. Ambas son tasas crudas porque representan la experiencia general de mortalidad en 1996 para toda la población de cada estado. Las tasas crudas no representan ninguna diferencia entre estas poblaciones con respecto a factores como, por ejemplo, edad, raza o sexo, que podrían tener alguna influencia en la

mortalidad. Sin considerar tales factores, sería prematuro llegar a dicha conclusión.
2. Arizona. El clima seco y cálido de Arizona atrae a muchas personas más viejas que el clima de Alaska.
3. Hay personas relativamente más viejas viviendo en Arizona, y las personas de mayor edad tienen un riesgo alto de morir.

P4.17

1. Controlando cualquier diferencia de edad en las dos poblaciones, la tasa de mortalidad general es más alta en Alaska con un clima frío y húmedo que en Arizona, donde el clima es cálido y seco.
2. La población de Alaska debe ser mucho más joven que la de los Estados Unidos, ya que la tasa ajustada por edad es mucho más alta que la tasa cruda.
3. La tasa para Arizona no cambió mucho de cruda a ajustada, porque la distribución etaria de Arizona era sólo apenas más joven que la de todo Estados Unidos en 1996.

P4.18

1. Verdadero. Si no se utiliza el ajuste por edad, entonces una diferencia de riesgo o tasas entre dos poblaciones puede deberse, principalmente, a las diferencias de edad entre las dos poblaciones.
2. Falso. No se garantiza que dos medidas ajustadas serán más cercanas o más alejadas una de otra de lo que lo fueron las medidas crudas correspondientes.
3. Falso. La elección de la población estándar depende de las características de las poblaciones que se consideren.
4. Verdadero. No hay límite con respecto a la cantidad de poblaciones que podrían ajustarse por edad.
5. Falso. Para las preguntas 5 y 6: si las tasas crudas son bastante diferentes mientras que las distribuciones etarias son similares, entonces es probable que las tasas ajustadas sean bastante distintas.
6. Verdadero
7. 44
8. 9,1
9. Mayores. Las mujeres deben ser mayores que los hombres en este caso. La tasa de mortalidad cae sustancialmente en las mujeres cuando estandarizamos la tasa mediante la distribución etaria de los hombres. En otras palabras, si sacamos la edad, las tasas para las mujeres disminuyen. Si las mujeres fueran más jóvenes, esperaríamos ver que la tasa ajustada aumenta cuando eliminamos el factor edad.

LECCIÓN 5

MEDIDAS DE EFECTO

5-1 La razón de riesgos versus el *odds ratio**

*En los estudios epidemiológicos, comparamos las frecuencias de enfermedad para dos o más grupos mediante una **medida de efecto**. En este capítulo, describiremos varios tipos de medidas de efecto. La elección de la medida depende normalmente del diseño del estudio que se está utilizando.*

Medidas de efecto mediante razones y diferencias

En la lección 5, nos concentramos en las medidas de efecto mediante razones, que son de la forma M_1/M_0, donde M_1 y M_0 son dos medidas de frecuencia de la enfermedad, por ejemplo, riesgos, tasas o prevalencias que se están comparando.

En la lección 6 sobre "Medidas de impacto potencial", consideramos diferentes medidas de efecto, que son de la forma M_1-M_0. Las medidas que utilizan diferencias también se llaman medidas de riesgo atribuible.

Las medidas que utilizan razones se usan normalmente en los estudios epidemiológicos que abordan la etiología de una enfermedad/resultado en la salud, mientras que las medidas que utilizan diferencias se usan para cuantificar la importancia en la salud pública de los factores que son determinantes de una enfermedad/resultado en la salud.

Tabaquismo y cáncer de pulmón

Después de la Primera Guerra Mundial, el fumar cigarrillos se tornó muy popular en Estados Unidos, ya que se repartían cigarrillos a los soldados para levantarles el ánimo. Pero junto con el aumento del tabaquismo, también aumentaron en forma preocupante los casos de cáncer de pulmón y algunas advertencias tempranas de unos pocos doctores sobre los posibles peligros de fumar. Los estudios iniciales de las décadas de los años treinta y cuarenta sobre la posible relación entre el tabaquismo y el cáncer de pulmón fueron **estudios de casos y controles**. Se tornó bastante evidente que los pacientes con cáncer de pulmón fumaban mucho más que los controles. En un estudio específico, los pacientes con cáncer de pulmón tenían una probabilidad de fumar dos paquetes por día 17 veces mayor a la de los controles.

A principios de la década de los años cincuenta, los doctores Horn y Hammond de la Sociedad Americana del Cáncer realizaron uno de los primeros **estudios de cohorte** sobre los efectos dañinos del tabaquismo. Aproximadamente 200.000 personas recibieron un cuestionario sobre tabaquismo y luego se realizó su seguimiento durante cuatro años. Se compararon las tasas de mortalidad y la causa de muerte para fumadores y no fumadores. El estudio preliminar publicado en 1958 causó conmoción. Fue el estudio sobre tabaquismo más

*Medida también conocida en la literatura epidemiológica como razón de disparidad, razón de productos cruzados, razón de posibilidades, razón de ventajas, razón de chances o razón de momios.

grande que se había realizado y demostró que los fumadores tenían diez veces más probabilidad de padecer cáncer de pulmón que los no fumadores.

Tanto el estudio de cohorte como el de casos y controles trataron de evaluar la relación propuesta entre tabaquismo y cáncer de pulmón mediante la derivación de una medida de efecto que cuantificara la magnitud de esta relación. La medida descrita en el **estudio de casos y controles** se llama ***odds ratio*, o razón de posibilidades**. La medida descrita en el **estudio de cohorte** se llama **razón de riesgos**. En las actividades siguientes se discuten estas dos medidas fundamentales de efecto.

Resumen

- El *odds ratio* y la razón de riesgos son dos medidas fundamentales de efecto.
- Estas medidas fueron utilizadas en los estudios de epidemiología sobre la relación entre tabaquismo y cáncer de pulmón.
- El *odds ratio* es la medida de efecto que normalmente se utiliza en los estudios de casos y controles.
- La razón de riesgos es la medida de efecto que normalmente se utiliza en los estudios de cohorte.

Razón de riesgos

El cuadro siguiente resume los resultados de un estudio de seguimiento de cinco años para determinar si los fumadores que han tenido un ataque cardíaco reducirán su riesgo de morir al dejar de fumar. Se estudió una cohorte de 156 pacientes que padecieron ataques cardíacos, todos ellos fumadores regulares hasta el momento de padecer el ataque. Setenta y cinco de estos pacientes continuaron fumando después del ataque cardíaco. Los otros 81 pacientes dejaron de fumar durante el período de recuperación. De los 75 pacientes que continuaron fumando, 27 murieron; por lo tanto, la proporción de estos pacientes que murieron es de 0,36. De los 81 pacientes que dejaron de fumar, 14 murieron; por lo tanto, la proporción correspondiente es de 0,17. Estas proporciones calculan el riesgo de morir en cinco años para estos dos grupos de pacientes. Podríamos preguntarnos si aquellos pacientes que sufrieron un ataque cardíaco y que continuaron fumando tienen más probabilidades de morir dentro de los 5 años después de sufrir su primer ataque cardíaco que aquellos que dejaron de fumar.

Pacientes de ataques cardíacos

	Fuman	Dejaron de fumar	**Total**
Murieron	27	14	41
Sobrevivieron	48	67	115
Total	75	81	156

Probabilidad de morir en 5 años

siguieron fumando: $27/75 = 0.36$

dejaron de fumar: $14/81 = 0.17$

Una medida de efecto brinda una respuesta numérica a esta pregunta. Dicha medida nos permite comparar dos o más grupos, en este caso, los fumadores que continuaron fumando y los que dejaron de fumar. Para los estudios de seguimiento como los que se describen aquí, la medida típica de efecto es una **razón de riesgos**. Para calcular una razón de riesgos, tomamos la razón de dos riesgos que se están comparando, es decir, simplemente dividimos un riesgo por el otro. En verdad, obtenemos un "cálculo", o estimación, de la razón de riesgos, lo cual indicamos colocando un "acento circunflejo" sobre la notación RR. $\hat{RR}$ es una estimación porque utilizamos dos

estimaciones de riesgo basadas en muestras de los dos grupos que se están comparando. En nuestro ejemplo, por lo tanto, dividimos 0,36 por 0,17 para obtener 2,1.

$$\hat{\text{RR}}\ \text{Estimado} = \frac{\text{Riesgo estimado para los que continúan fumando}}{\text{Riesgo estimado para los que dejaron de fumar}} = \frac{0{,}36}{0{,}17} = 2{,}1$$

La razón de riesgos del 2,1 nos dice que los fumadores que continúan fumando tienen aproximadamente el doble de probabilidades de morir que los que dejan de fumar. En otras palabras, para los pacientes que sufrieron ataques cardíacos, el riesgo en cinco años para los que continuaron fumando es de aproximadamente el doble que el riesgo correspondiente para los fumadores que dejaron de fumar.

Preguntas de estudio (P5.1):

Con el ejemplo del estudio de seguimiento de cinco años que compara la mortalidad entre los que continuaron fumando y los que dejaron de fumar:

1. ¿cómo interpretaría una razón de riesgos de 4,5?
2. ¿Qué sucedería si la razón de riesgos fuese de 1,1?
3. ¿Y si la razón de riesgos fuese menor que 1, por ejemplo, 0,5?
4. ¿Cómo interpretaría un valor de 0,25?

Si nuestra razón de riesgos hubiese sido de 1,1, habríamos comprobado que el riesgo para los fumadores que continuaron fumando era básicamente igual que el riesgo para los que dejaron de fumar. Llamamos **valor nulo** de la razón de riesgos a una razón de riesgos igual a 1. Éste es el valor que obtenemos para la razón de riesgos cuando no hay efecto, es decir, cuando el efecto es nulo.

Resumen

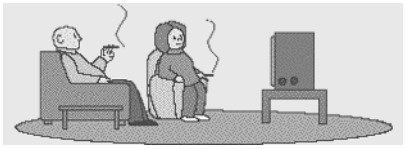
- La razón de riesgos (RR) es la razón del riesgo en un grupo, digamos grupo 1, con respecto al riesgo en otro grupo, digamos grupo 0.
- El valor de la RR puede ser mayor, igual o menor que 1.
- Si la RR es mayor que 1, el riesgo para el grupo 1 es más alto que el riesgo para el grupo 0.
- Si la RR es menor que 1, el riesgo para el grupo 1 es más bajo que el riesgo para el grupo 0.
- Y si la RR es igual a 1, los riesgos en los grupos 1 y 0 son iguales, de forma que resulta indistinto estar en un grupo o en el otro.

El numerador y el denominador de la razón de riesgos

En general, la razón de riesgos que compara dos grupos se define como el riesgo en un grupo dividido por el riesgo en el otro grupo. Es importante especificar claramente qué grupo es el numerador y qué grupo es el denominador.

Si, por ejemplo, los dos grupos se llaman grupo 1 y grupo 0, y el riesgo en el grupo 1 está en el numerador, entonces decimos que la razón de riesgos compara el grupo 1 con el grupo 0. Por otro lado, si el riesgo en el grupo 0 está en el numerador, entonces decimos que la razón de riesgos compara el grupo 0 con el grupo 1.

Auto-evaluación (P5.2)



Para los pacientes que sufrieron ataques cardíacos, la razón de riesgos es definida como el riesgo en los fumadores que siguen fumando dividido por el riesgo en los fumadores que dejaron de fumar. En los siguientes escenarios, ¿cuál sería la razón de riesgos?

1. Los fumadores que siguen fumando tienen el doble de probabilidades de morir que los fumadores que dejaron de fumar. . . . **???**

2. Los fumadores que siguen fumando tienen las mismas probabilidades de morir que los fumadores que dejaron de fumar. . . . **???**

3. Los fumadores que dejaron de fumar tienen el doble de probabilidades de morir que los fumadores que siguen fumando. . . . **???**

Opciones

0 **0,1** **0,2** **0,5** **1** **2**

Consideremos los datos de un ensayo clínico realizado en forma aleatoria para evaluar si tomar aspirina reduce el riesgo de padecer una enfermedad coronaria. El grupo expuesto recibió aspirina día de por medio mientras que el grupo de comparación recibió un placebo. A continuación se presenta un cuadro con los resultados.

		Aspirina		**Placebo**		**Total**
		n	Columna %	n	Columna %	
Desarrolló enfermedad coronaria	**Sí**	104	(1,04)	189	(2,36)	293
	No	9.896	(98,96)	7.811	(97,64)	17.707
	Total	10.000	(100,00)	8.000	(100,00)	18.000

La auto-evaluación continúa en la página siguiente

4. El riesgo calculado para el grupo con aspirina es **???**

5. El riesgo calculado para el grupo con placebo es **???**

6. La razón de riesgos calculada que compara el grupo con aspirina con el grupo con placebo es: **???**

Opciones

0,0104 **0,0236** **0,44** **104/189** **2,269** **98,96/97,64**

El *odds ratio*

Los epidemiólogos de la División de Enfermedades Bacterianas de los Centros para el Control y Prevención de Enfermedades de Atlanta (CDC, por sus siglas en inglés), investigan los orígenes de los brotes causados por consumir alimentos contaminados. Por ejemplo, se llevó a cabo un estudio de casos y controles para determinar el origen de un brote de diarrea en un club de vacaciones de Haití desde el 30 de noviembre hasta el 8 de diciembre de 1984.

Los investigadores se preguntaban si el origen principal del brote fue la ingesta de carne de hamburguesa cruda. Ya que éste es un **estudio de casos y controles**, y no un estudio de seguimiento, el estudio comienza con **casos** de personas en el club de vacaciones que tuvieron diarrea durante el período de interés. Los **controles** fueron una muestra al azar de 33 personas que se quedaron en el club de vacaciones durante el mismo período, pero que no tuvieron diarrea. Se registró un total de 37 casos durante el período del estudio. Un equipo de investigadores entrevistó a la totalidad de los 37 casos y los 33 controles para determinar qué comida consumieron durante su estadía en el club de vacaciones.

De los 37 casos, 17 personas comieron carne de hamburguesa cruda, de forma que la proporción de los casos que comieron carne de hamburguesa cruda es del 0,46. De los 33 controles, 7 comieron carne de hamburguesa cruda; por lo tanto, la proporción correspondiente es del 0,21. Nos podríamos preguntar, entonces, si estos datos sugieren que la ingesta de carne de hamburguesa cruda fue el origen del brote.

	Hamburguesa Cruda		
	Comió	No comió	**Total**
Casos	17	20	37
Controles	7	26	33
Total	24	46	70

Proporción de casos: $17/37 = 0.46$

Proporción de controles: $7/33 = 0.21$

Debido a que éste es un estudio de casos y controles en lugar de un estudio de seguimiento, estas proporciones no calculan los riesgos para los casos y controles. Por lo tanto, ***no*** podemos calcular una razón de riesgos. Entonces, ¿qué medida de efecto debería utilizarse en los estudios de casos y controles? La respuesta el ***odds ratio*** (**OR, por sus siglas en inglés),** que se describe en la próxima actividad.

Resumen

- Se utilizó un estudio de casos y controles para investigar un brote producido por alimentos en un club de vacaciones del Caribe.
- En un estudio de casos y controles, no podemos calcular los riesgos para los casos y controles.

- Como consecuencia, no podemos utilizar una razón de riesgos (RR) como medida de efecto, sino que debemos utilizar la razón de chances, u *odds ratio* (OR, por sus siglas en inglés), en su lugar.

¿Por qué no podemos utilizar una razón de riesgos en los estudios de casos y controles?

En un estudio de casos y controles, no podemos calcular el riesgo, sino que calculamos las **probabilidades de exposición** para los casos y controles. La probabilidad de exposición para un caso es la probabilidad de que un individuo esté expuesto si ese individuo es un caso; esto no es equivalente a la probabilidad de que un individuo sea un caso si está expuesto, que es el riesgo para los expuestos.

En otras palabras, utilizando la notación de probabilidad condicional:

Pr(expuesto | caso) ≠ Pr(caso | **expuesto**), donde "|" significa "si".

De forma similar, la probabilidad de exposición para un control no es equivalente a 1 menos el riesgo para los expuestos. Es decir,

Pr(**expuesto** | control) ≠ 1 - Pr(caso | **expuesto**).

Lamentablemente, la razón de dos probabilidades, o chances, de exposición no es una razón de riesgos. Por lo tanto, en los estudios de casos y controles debemos utilizar una medida de efecto diferente, por ejemplo, la razón de chances, u *odds ratio*.

El *odds ratio* (continuación)

Para entender el *odds ratio*, debemos comenzar con el concepto de qué es ***odds, concepto también conocido como chance, ventaja, o momio***. El término ***odds*** se usa comúnmente en eventos deportivos. Podemos leer que los chances, u *odds,* son de 3 a 1 en contra de que un caballo en particular gane una carrera, o que los chances, u *odds,* son de 20 a 1 en contra de que España gane el próximo Mundial, o que los chances, u *odds,* son de 1 a 2 de que los New York Yankees vayan a llegar a la Serie Mundial este año. Cuando decimos que los chances, u *odds,* en contra de un caballo en particular son de 3 a 1, queremos decir que el caballo tiene 3 veces más probabilidades de perder que de ganar.

Los chances, u *odds,* de un evento son fáciles de calcular desde su probabilidad de aparición. Los *odds* pueden expresarse como **P**, la probabilidad de que un evento ocurra, dividido por **1-P**, la probabilidad de que el evento no ocurra.

$$\text{Odds} = \frac{\mathbf{P}}{\mathbf{1 - P}} = \frac{\text{P(El evento ocurra)}}{\text{P(El evento no ocurra)}}$$

En nuestro ejemplo de la carrera de caballos, si P denota las probabilidades de que el caballo pierda, entonces 1 – P denota las probabilidades contrarias: que el caballo gane. Por lo tanto, si las probabilidades de que el caballo pierda son de 0,75, entonces las probabilidades de que el caballo gane son de 0,25 y los chances, u *odds,* son de 3 ó de 3 a 1.

$$\text{Odds} = \frac{\mathbf{P}}{\mathbf{1 - P}} = \frac{\text{P(El caballo pierda)}}{\text{P(El caballo gane)}} = \frac{0,75}{0,25} = 3 \text{ ó } \frac{3}{1}$$

En el estudio de casos y controles del club de vacaciones de Haití, recuerde que el evento de interés ocurre si un individuo del estudio comió carne de hamburguesa cruda y, de ser así, decimos que el individuo está **expuesto**. Las probabilidades de exposición en los casos fueron de 0,46, por lo tanto, los chances, u ***odds,* de estar expuesto en los casos** es de 0,46 dividido por 1 - 0,46:

$$\hat{\text{Odds}}_{\text{Casos}} = \frac{0,46}{1 - 0,46} = 0,85$$

De forma similar, las probabilidades de exposición para los controles fue de 0,21, por lo tanto, los chances, u *odds,* para los controles es de 0,21 dividido por 1 - 0,21:

$$\hat{\text{Odds}}_{\text{Controles}} = \frac{0,21}{1 - 0,21} = 0,27$$

La razón de ***odds*** calculada para estos datos es la razón de *odds* en los casos dividida por *odds* en los controles, que da por resultado 3,2.

$$\text{Odds Ratio (OR)} = \frac{\hat{\text{Odds}}_{\text{Casos}}}{\hat{\text{Odds}}_{\text{Controles}}} = \frac{0,85}{0,27} = 3,2$$

¿Cómo interpretamos este valor del *odds ratio*? Una interpretación es que los ***odds* de exposición para los casos** son de aproximadamente 3,2 veces los ***odds* de exposición para los controles.** Ya que los que comieron carne de hamburguesa cruda son los individuos expuestos, los *odds* de que un caso haya comido carne de hamburguesa cruda resultan ser 3,2 veces superiores a los *odds* de que un individuo de control lo haya hecho.

Preguntas de estudio (P5.3)

Con el ejemplo del estudio de casos y controles de Haití:

1. ¿Cómo interpretaría un *odds ratio* de 2,5?
2. ¿Qué sucedería si el *odds ratio* fuese de 1,1?
3. ¿Y si el *odds ratio* fuera menor que 1, por ejemplo, 0,5?
4. ¿Cómo interpretaría un valor de 0,25?

Las razones de *odds*, como las razones de riesgos, pueden ser mayores, iguales o menores que 1. Una razón de *odds* mayor que 1 indica que los *odds* de exposición en los casos son **mayores** que los *odds* de exposición en los controles. Una razón de *odds* menor que 1 indica que los *odds* de exposición en los casos son **menores** que los *odds* de exposición en los controles. Una razón de *odds* igual a 1 indica que los *odds* de exposición en los casos y controles son iguales.

$$\hat{OR} = \hat{O}dds_{|Casos} / \hat{O}dds_{|Controles}$$

¿$\hat{OR}$ mayor que 1?

$$\hat{O}dds_{Casos} > \hat{O}dds_{Controles}$$

¿$\hat{OR}$ igual a 1?

$$\hat{O}dds_{Casos} = \hat{O}dds_{|Controles}$$ No hay efecto

¿$\hat{OR}$ menor que 1?

$$\hat{O}dds_{Casos} < \hat{O}dds_{Controles}$$

Resumen

- Los chances, u *odds,* de un evento pueden calcularse como P/(1-P), donde P es la probabilidad del evento.
- Los chances, u *odds ratio* (OR, por sus siglas en inglés), son la razón entre dos *odds*.
- En los estudios de casos y controles, el OR es dado por los *odds* de exposición para los casos dividido por los *odds* de exposición para los controles.
- Las razones de *odds*, como las razones de riesgos, pueden ser mayores, iguales o menores que 1, donde 1 es el valor nulo.



Auto-evaluación (P5.4)

En la década de los años veinte, se sospechó por primera vez una relación causal entre fumar cigarrillos y el cáncer pulmonar, sobre la base de observaciones clínicas. Para probar esta aparente asociación, se realizaron numerosos estudios entre 1930 y 1960. Un **estudio clásico de casos y controles** se llevó a cabo en 1947 para comparar los hábitos de fumar de los pacientes con cáncer pulmonar y los de otros pacientes.

1. En este estudio de casos y controles, es **???** calcular el riesgo de cáncer pulmonar entre los fumadores y, por lo tanto, la medida de asociación adecuada es la **???**

Opciones
Imposible **razón de *odds*** **posible** **razón de riesgos**

Consideremos los siguientes datos de un estudio clásico de casos y controles para evaluar la relación entre el tabaquismo y el cáncer pulmonar. Los casos fueron pacientes hospitalizados a los que recientemente se les diagnosticó cáncer pulmonar. Los controles fueron pacientes que padecían otros trastornos. Esta tabla de 2x2 compara los hábitos de fumar para los casos y controles de sexo masculino.

2. Las probabilidades de ser un fumador entre los casos son de . . **???**

3. Las probabilidades de ser un fumador entre los controles son de . . **???**

4. Los *odds* de fumar entre los casos son de . . . . **???**

5. Los *odds* de fumar entre los controles son de . . . . **???**

6. La razón de *odds* es de . . . . **???**

Opciones

0,11 1,04 10,50 1.296/1.357 1.350/1.357 1.350/2.646 192,86 21,25 7/68 9,08

	Fumador de cigarrillos	No fumador	Total
Casos	1.350	7	1.357
Controles	1.296	61	1.357
Total	2.646	68	2.714

En un estudio de casos y controles para descubrir el origen de un brote, la razón de *odds* de comer ensalada de col es definida como los *odds* en los casos divididos por los *odds* en los controles. Para los siguientes escenarios, ¿cuál sería la razón de *odds*?

7. Los casos tienen *odds* de comer ensalada de col tres veces más altas que los controles. **???**

8. Los casos tienen los mismos *odds* de comer ensalada de col que los controles **???**

9. Los controles tienen *odds* de comer ensalada de col tres veces mayores que los casos **???**

Opciones

0 0,25 0,333 1 3 4

5-2 Cálculo de la razón de *odds*/Diferentes diseños de estudios

Cálculo de la razón de *odds (odds ratio)*

Este diseño de tabla de 2x2 brinda una forma más conveniente para calcular la razón de *odds*. La fórmula es ***a*** por ***d*** sobre ***b*** por ***c***. Se denomina fórmula de la **razón de productos cruzados**, porque es la razón de un producto que está cruzado en la tabla dividida por el otro producto que está cruzado en la tabla.

Forma general de tabla 2x2

	Estado de exposición: Sí	Estado de exposición: No	**Total**
Casos	a	b	m_1
Controles	c	d	m_0
Total	n_1	n_0	m

Razón de productos cruzados $\quad \hat{OR} = \dfrac{a \times d}{b \times c}$

Como ejemplo de esta fórmula, considere los datos del brote en el club de vacaciones de Haití. La fórmula de productos cruzados nos brinda el mismo resultado, 3,2, que habíamos obtenido originalmente de la razón de *odds* de exposición en los casos y controles.

Estudio de casos y controles: brote en el club de vacaciones de Haití

	Hamburguesa Cruda: Sí	Hamburguesa Cruda: No	**Total**
Casos	a=17	b=20	m_1=37
Controles	c=7	d=26	m_0=33
Total	n_1=24	n_0=46	m=70

Razón de productos cruzados

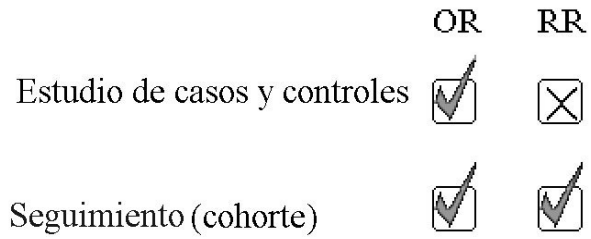
$$\hat{OR} = \frac{a \times d}{b \times c} = \frac{(17)(26)}{(20)(7)} = 3.2$$

$$= \frac{\hat{Odds}_{casos}}{\hat{Odds}_{controles}}$$

Preguntas de estudio (P5.5)

1. ¿Deberíamos calcular el OR para otras comidas que se ingirieron durante el brote antes de afirmar que el origen fue la carne de hamburguesa cruda?

Hay algunos tipos de estudios de casos y controles en los que sí se puede calcular el riesgo directamente (dependiendo de su temporalidad); sin embargo, en los estudios de casos y controles donde no se puede calcular el riesgo directamente, es obligatorio calcular el *odds ratio* como medida de efecto. En cambio, tanto la razón de riesgos como el *odds ratio* pueden calcularse en los estudios de seguimiento.



(Observe que el OR y la RR también pueden calcularse en los ensayos clínicos aleatorizados que tienen medidas de incidencia acumulada.)

Por ejemplo, consideremos el estudio de "dejar de fumar" para los pacientes que padecieron ataques cardíacos. El diseño de estudio aquí es un estudio de seguimiento. Anteriormente calculamos que el riesgo para los pacientes que continuaron fumando fue 2,1 veces mayor que el riesgo para aquellos que dejaron de fumar.

Estudios de seguimiento (cohorte)

	Fuma	Dejó	**Total**
Muerte	27	14	41
Sobrevivencia	48	67	115
Total	75	81	156

$$\hat{RR} = \frac{\text{Riesgo para los que continuaron fumando}}{\text{Riesgo para los que dejaron de fumar}} = \frac{27/75}{14/81} = \frac{0.36}{0.17} = 2.1$$

$$\hat{OR} = \frac{a \times d}{b \times c} = \frac{(27)(67)}{(14)(48)} = 2.7$$

El uso de la fórmula de productos cruzados con estos datos de seguimiento da 2,7. No debería sorprender el hecho de que estos números (la razón de riesgos y la razón de *odds*) no sean iguales, ya que la razón de riesgos y la razón de *odds* son dos medidas distintas. Pero los valores en este ejemplo no son muy diferentes. De hecho, estos dos cálculos tienen interpretaciones similares, ya que ambos sugieren que existe una relación moderada entre el estatus de dejar de fumar y el estatus de supervivencia.

Resumen

- Una fórmula conveniente del OR es la razón de productos cruzados: (ad)/(bc)
- El OR puede calcularse tanto en estudios de casos y controles como en estudios de seguimiento, mediante la fórmula de productos cruzados.

(Véanse las actividades en la página 5-4 de esta lección para una discusión sobre cómo puede aproximarse la razón de riesgos mediante la razón de *odds*.)

Auto-evaluación (P5.6)

Para estudiar la relación entre el uso de anticonceptivos orales y el cáncer ovárico, el CDC inició el estudio de cáncer y hormonas esteroides en 1980. Fue un estudio de casos y controles.

1. Con la fórmula de la razón de productos cruzados, el OR que compara la condición de exposición de los casos y los controles es el siguiente: (93) * (**???**) / (**???**) * (959), que es igual a **???**

2. Esto significa que las **???** de **???** entre los casos fue **???** las **???** de exposición entre los **???**

Opciones

0,23 **0,77** **1,3** **683** **86** **casos** **controles** **enfermedad** **expuestos** **exposición** **mayor que** **menor que** **no expuestos** ***odds*** **riesgo**

	Alguna vez utilizaron OC	Nunca utilizaron OC	Total
Casos	93	86	179
Controles	959	683	1.642
Total	1.052	769	1.821

El *odds ratio* en diferentes diseños de estudios

El *odds ratio* puede calcularse tanto en estudios de casos y controles como en estudios de seguimiento (cohorte). Debido a que un estudio de casos y controles requiere que calculemos probabilidades de exposición en lugar de riesgos, generalmente llamamos ***odds ratio* de exposición (EOR, por sus siglas en inglés)** al *odds ratio* calculado en los estudios de casos y controles. Por el contrario, debido a que un estudio de seguimiento nos permite calcular riesgos, generalmente llamamos ***odds ratio* de riesgo (ROR, por sus siglas en inglés)** al *odds ratio* calculado en los estudios de seguimiento.

El *odds ratio* también puede calcularse para estudios transversales. Debido a que un estudio transversal mide la prevalencia o condiciones existentes en un punto en el tiempo, generalmente llamamos ***odds ratio* de prevalencia (POR, por sus siglas en inglés)** al *odds ratio* calculado en un estudio transversal.

Odds Ratio

Estudio de casos y controles (probabilidades de exposición):

Odds ratio de exposición (EOR)

Estudio de seguimiento (riesgos):

Odds ratio de riesgo (ROR)

Estudios transversales (prevalencia):

Odds ratio de prevalencia (POR)

Como ejemplo del cálculo de un *odds ratio* de prevalencia para datos transversales, considere estos datos que se recabaron de una encuesta transversal diseñada para evaluar la relación entre la enfermedad coronaria y

varios factores de riesgo, uno de los cuales era el tipo de personalidad. Para estos datos transversales, podemos utilizar la fórmula general de la razón de productos cruzados para calcular un *odds ratio* de prevalencia. Los *odds* de tener una personalidad tipo A entre los que padecen enfermedad coronaria son 5 veces superiores a los *odds* de los que no padecen la enfermedad.

Datos transversales (POR):

		Personalidad Tipo A	B	Total
CHD	Sí	93	36	129
	No	46	89	135
	Total	139	125	264

Odds ratio de prevalencia (POR) = ad/bc

$$\hat{POR} = \frac{(93)(89)}{(46)(36)} = 5.0$$

$$\text{Odds de Tipo A}_{\text{CHD}} = 5\text{x Odds de Tipo A}_{\text{NO-CHD}}$$

En general, podemos utilizar la fórmula de la razón de productos cruzados para calcular un *odds ratio* de exposición, un *odds ratio* de riesgo o un *odds ratio* de prevalencia según el diseño del estudio utilizado.

Resumen

- El OR calculado en el estudio de casos y controles se llama *odds ratio* de exposición (EOR, por sus siglas en inglés).
- El OR calculado en el estudio de seguimiento se llama *odds ratio* de riesgo (ROR, por sus siglas en inglés).
- El OR calculado en el estudio transversal se llama *odds ratio* de prevalencia (POR, por sus siglas en inglés).
- Podemos utilizar la fórmula general de la razón de productos cruzados para calcular el EOR, el ROR o el POR, según el diseño del estudio utilizado.

¿ROR = EOR = POR?

No necesariamente. Aunque la fórmula de cálculo (es decir, ad/bc) es la misma, independientemente del diseño del estudio, podrían obtenerse valores diferentes del *odds ratio* calculado a partir de una tabla de 2x2 para diseños de estudios diferentes. Esto se debe a la posibilidad de sesgo de selección (descrito en la lección 8). Por ejemplo, un estudio de casos y controles que utiliza casos prevalentes podría dar un cálculo diferente del *odds ratio* que un estudio de seguimiento que sólo incluye casos incidentes.

Auto-evaluación (P5.7)

A continuación, se muestran los datos para un estudio transversal que tiene como propósito evaluar si fumar cigarrillos durante el embarazo es un factor de riesgo de que el bebé nazca con bajo peso.

1. Calcule el *odds ratio* que mide si las fumadoras tienen mayores probabilidades de dar a luz a bebés con bajo peso al nacer que las no fumadoras. OR=**???**

2. Este valor del *odds ratio* sugiere que las fumadoras tienen **???** de dar a luz a bebés con bajo peso al nacer que las no fumadoras.

3. Este *odds ratio* es un ejemplo de *odds ratio* de **???**

Opciones

0,48 **2,04** **2,18** **exposición** **menor probabilidad** **mayor probabilidad** **prevalencia** **riesgo**

	Fumadoras	No fumadoras	Total
Peso bajo al nacer	1.556	14.974	16.530
Peso alto al nacer	694	14.532	15.226
Total	2.250	29.506	31.756

5-3 Comparación del *odds ratio* con las aproximaciones de la razón de riesgos

Comparación de la razón de riesgos con el *odds ratio*

Hemos descrito dos medidas de efecto que se utilizan ampliamente: la razón de riesgos y el *odds ratio*. Frecuentemente, se prefieren las razones de riesgos, porque son más fáciles de interpretar. Pero, como hemos visto, en los estudios de casos y controles, no podemos calcular los riesgos y, en su lugar, debemos trabajar con un *odds ratio* de exposición (EOR, por sus siglas en inglés). Sin embargo, en los estudios de seguimiento tenemos la opción de calcular tanto una razón de riesgos como un *odds ratio* de riesgo (ROR, por sus siglas en inglés). ¿Cuál deberíamos preferir?

Puede demostrarse matemáticamente que si un valor de la razón de riesgos es igual o mayor que 1, entonces, el *odds ratio* de riesgo correspondiente es, al menos, tan grande como la razón de riesgos. Por ejemplo, con los datos de seguimiento del estudio sobre dejar de fumar realizado en pacientes que habían sufrido ataques cardíacos, vimos que la razón de riesgos fue 2,1, que es mayor que 1; el *odds ratio* correspondiente fue 2,7, que es mayor que 2,1.

Estudios de seguimiento: $\hat{RR}$ vs. $\hat{ROR}$

Si $\hat{RR} \geq 1$, entonces $\hat{ROR} \geq \hat{RR}$

Datos de seguimiento sobre dejar de fumar en pacientes que habían sufrido ataques cardíacos

	Fuma	Dejó	Total
Muerte	27	14	41
Sobrevivir	48	67	115
Total	75	81	156

$\hat{RR} = 2.1$

$\hat{ROR} = 2.7$

De igual manera, si la razón de riesgos es menor que 1, el *odds ratio* correspondiente es tan pequeño o más pequeño que la razón de riesgos. Por ejemplo, si intercambiamos las columnas del cuadro de dejar de fumar, entonces la razón de riesgos es 0,48, que es menor que 1, y el *odds ratio* correspondiente es 0,37, que es menor que 0,48.

Si $\hat{RR} \leq 1$, entonces $\hat{ROR} \leq \hat{RR}$

Datos de seguimiento sobre dejar de fumar en pacientes que habían sufrido ataques cardíacos

	Fuma	Dejó	Total
Muerte	14	27	41
Sobrevivir	67	48	115
Total	81	75	156

$\hat{RR} = 0.48$

$\hat{ROR} = 0.37$

Si $\hat{RR} \geq 1$, entonces $\hat{ROR} \geq \hat{RR}$

Si $\hat{RR} \leq 1$, entonces $\hat{ROR} \leq \hat{RR}$

También puede demostrarse que si la enfermedad es "poco frecuente", entonces el *odds ratio* de riesgo se aproximará mucho a la razón de riesgos. Para los estudios de seguimiento, este *supuesto de enfermedad poco frecuente* significa que el riesgo de que algún individuo del estudio desarrolle la enfermedad es lo suficientemente mínimo, de forma tal que los valores correspondientes del *odds ratio* y la razón de riesgos brindan básicamente la misma interpretación del efecto de exposición en la enfermedad.

Normalmente, se considera que una "enfermedad" es poco frecuente si se trata de una enfermedad que ocurre en forma tan esporádica en la población de interés que el riesgo para cualquier individuo del estudio es aproximadamente cero. Por ejemplo, si una de cada 100.000 personas desarrolla la enfermedad, el riesgo para esta población es cero si consideramos 4 espacios decimales. Bueno, esto es realmente poco frecuente.

Preguntas de estudio (P5.8)

1. ¿Indica un riesgo de 0,01 una enfermedad poco frecuente?
2. Suponga que para un determinado estudio de seguimiento, el riesgo verdadero no se considera poco frecuente. ¿Es posible que el ROR y la RR sean aproximadamente iguales?

Podemos escribir una fórmula que exprese el *odds ratio* de riesgo en términos de la razón de riesgos:

$$ROR = RR \times f$$

donde

$$f = \frac{(1 - R_0)}{(1 - R_1)}$$

R_0 es el riesgo para los no expuestos
R_1 es el riesgo para los expuestos
$RR=R_1/R_0$

Esta fórmula dice que el *odds ratio* de riesgo es igual a la razón de riesgos multiplicada por el factor ***f***, donde ***f*** se define como 1 menos el riesgo para el grupo no expuesto (R_0) dividido por 1 menos el riesgo para el grupo expuesto (R_1). De esta ecuación se deduce que si tanto **R_1** como **R_0** son aproximadamente igual a cero, entonces ***f*** es aproximadamente igual a 1, y el *odds ratio* de riesgo es aproximadamente igual a la razón de riesgos.

Preguntas de estudio (P5.9)

1. En el ejemplo de dejar de fumar, donde R_0 es 0,17 y R_1 es igual a 0,36, ¿cuánto es ***f***?
2. Para este valor de ***f***, ¿el ROR se aproxima a la RR?
3. ¿Qué le sucede a ***f*** si los riesgos se reducen a la mitad, es decir, R_0 = 0,17/2 = 0,085 y R_1 = 0,36/2 = 0.,180?

4. ¿Los valores del ROR y la RR se aproximan para este ***f***?
5. ¿Qué le sucede a ***f*** si volvemos a reducir los riesgos a la mitad, de forma que R_0=0,0425 y R_1=0,09?
6. ¿La aproximación es mejor?
7. Basándose en sus respuestas a las preguntas anteriores, ¿qué "poco" frecuentes tienen que ser los riesgos para que el *odds ratio* y la razón de riesgos sean aproximadamente iguales?

Resumen

- Si un valor de RR $\geq$ 1, entonces, el valor correspondiente de ROR es, al menos, tan grande como el valor de RR.
- Si un valor de RR $\leq$ 1, entonces, el valor correspondiente de ROR es tan pequeño como, o más pequeño que, el valor de RR.
- En los estudios de seguimiento (de cohorte), el "supuesto de enfermedad poco frecuente" dice que el riesgo para cualquier individuo del estudio es aproximadamente igual a cero.
- Según el supuesto de enfermedad poco frecuente, el *odds ratio* de riesgo (ROR) calculado en un estudio de seguimiento se aproxima a la razón de riesgos (RR) calculada a partir del mismo estudio.

Comparación de la RR y el OR en el estudio de Rotterdam

La osteoporosis es una enfermedad común de los ancianos, y provoca un mayor riesgo de fracturas de los

huesos. Para estudiar esta enfermedad, se realizó un seguimiento durante cuatro años de una cohorte compuesta por casi 1.800 mujeres posmenopáusicas que vivían en Rotterdam, Holanda. Los investigadores del estudio de Rotterdam querían saber qué factores genéticos determinan el riesgo de fracturas por osteoporosis. Se concentraron en la codificación del gen para uno de los colágenos involucrados en la formación ósea. La composición genética de cada persona está compuesta por dos alelos de este gen, y cada alelo puede tener una de dos formas alternativas, llamadas alelo A o alelo B. Los investigadores demostraron que las mujeres con dos alelos A tenían una mayor masa ósea que las mujeres con, al menos, un alelo B. Por lo tanto, formularon la hipótesis de que el riesgo de fracturas sería mayor en las mujeres con alelo B.

De las 1.194 mujeres con dos alelos A, 64, ó 5,36%, sufrieron una fractura durante el seguimiento. De las 584 mujeres con al menos un alelo B, 47, ó 8,05%, sufrieron una fractura.

Preguntas de estudio (P5.10)

1. Calcule la razón de riesgos para la presencia de fracturas en mujeres con al menos un alelo B, en comparación con mujeres con dos alelos A.

Debido a que el valor de la razón de riesgos es mayor que 1, esperamos que el *odds ratio* de riesgo sea al menos tan grande como la razón de riesgos.

Preguntas de estudio (P5.10) (continuación)

2. Calcule el *odds ratio* de riesgo para la presencia de fracturas en mujeres con al menos un alelo B, en comparación con mujeres con dos alelos A.

Observe que el riesgo de fracturas es relativamente poco frecuente en esta población; por lo tanto, el *odds ratio* de riesgo es aproximadamente igual a la razón de riesgos. Recuerde la fórmula ROR = RR * ***f***. Aquí, ***f*** se define como 1 menos el riesgo en las mujeres con dos alelos A dividido por 1 menos el riesgo en las mujeres con al menos un alelo B.

$$\mathrm{ROR} = \mathrm{RR} \times \frac{1 - \mathrm{R}(2 \text{ alelos B})}{1 - \mathrm{R}(1 \text{ alelo A})}$$

Preguntas de estudio (P5.10) (continuación)

3. Con la fórmula ROR = RR x ***f***, ¿puede demostrar que calculamos el *odds ratio* de riesgo correcto?

En este estudio, tanto la razón de riesgos como el *odds ratio* de riesgo llevan a la misma conclusión: las mujeres con al menos un alelo B tienen un 50% más de probabilidades de fracturas que las mujeres con dos alelos A. Los investigadores del estudio de Rotterdam llegaron a la conclusión de que la composición genética puede predisponer a las mujeres a sufrir fracturas osteoporósicas.

Auto-evaluación (P5.11): RR en contraposición a OR en los estudios de seguimiento

Se entregó un cuestionario a aquellas personas que asistieron a un evento social en el que 39 de los 87 participantes se enfermaron con una condición diagnosticada como salmonelosis. La siguiente tabla de 2x2 resume la relación entre el consumo de ensalada de papa y la enfermedad.

1. La razón de riesgos que compara los expuestos con los no expuestos es . . . **???**
2. El *odds ratio* es . . . . . **???**
3. ¿El *odds ratio* se aproxima mucho a la razón de riesgos? . . . . . **???**
4. ¿Considera que esta enfermedad es "poco frecuente"? . . . . . **???**

Opciones
0,25 **1,7** **3,7** **36,0** **9,8** **no** **sí**

	Expuestos	No expuestos	Total
Enfermos	36	3	39
Sanos	12	36	48
Total	48	39	87

Consideremos los datos de un estudio clásico de pelagra. La pelagra es una enfermedad causada por una deficiencia dietaria de niacina y caracterizada por dermatitis, diarrea y demencia. A continuación se muestran datos que comparan casos por sexo.

5. La razón de riesgos de la pelagra para las mujeres en contraposición a los hombres es (con un decimal) . . . . . **???**
6. El *odds ratio* es (con un decimal) . . . . . **???**
7. ¿El *odds ratio* se aproxima mucho a la razón de riesgos? . . . **???**
8. ¿Considera que esta enfermedad es "poco frecuente"? . . . **???**

Opciones
1,4 2,4 2,5 24,2 no sí

	Mujeres	Hombres	Total
Enfermos	46	18	64
Sanos	1.438	1.401	2.839
Total	1.484	1.419	2.903

Aproximación del *odds ratio* en los estudios de casos y controles

Comparación de la RR y el OR en los estudios de casos y controles

Ya hemos visto que, para los estudios de seguimiento, si la enfermedad es "poco frecuente", entonces, el *odds ratio* de riesgo será una aproximación cercana a la razón de riesgos calculada a partir de los mismos datos de seguimiento. Sin embargo, en los estudios de casos y controles, no se puede calcular una razón de riesgos, y se debe utilizar en su lugar un *odds ratio* de exposición. Por lo tanto, para los datos de casos y controles, si la enfermedad es "poco frecuente", ¿el odds ratio de exposición se aproxima a la razón de riesgos que hubiese resultado de un estudio de seguimiento comparable? La respuesta es afirmativa, según ciertas condiciones que deben satisfacerse, como describiremos a continuación.

Este cuadro de doble entrada categoriza el cáncer de pulmón y tabaquismo para una cohorte de médicos clínicos en una ciudad metropolitana grande, a los que se les hace un seguimiento durante 7 años. Cuarenta fumadores y veinte no fumadores desarrollaron cáncer de pulmón. La razón de riesgos es 2. Asimismo, para esta población, el *odds ratio* de riesgo es 2,02, básicamente igual que la razón de riesgos. Debido a que éstas son medidas de efecto para una población, no hemos colocado el acento circunflejo sobre la notación de la razón de riesgos y el *odds ratio* de riesgo.

Cohorte de médicos (7 años de seguimiento)

	¿Fumador? Sí	No	**Total**
CP	40	20	60
No CP	1960	1980	3940
Total	2000	2000	4000

medidas de efecto para una población

$$RR = \frac{40/2000}{20/2000} = 2$$

$$ROR = \frac{40 \times 1980}{1960 \times 20} = 2.02$$

Ahora consideremos los resultados que esperaríamos obtener si realizamos un estudio de casos y controles utilizando esta cohorte como nuestra población de origen. Supondremos que se ha realizado el seguimiento durante 7 años. También supondremos que existe un registro completo sobre casos de cáncer, de forma tal que pudimos encontrar los 60 casos incidentes que se desarrollaron durante el período de 7 años. Estos serían los casos en nuestro estudio de casos y controles. Ahora supondremos que seleccionamos en forma aleatoria 60 controles de la población de origen para que sean nuestro grupo de comparación. Ya que la mitad de toda la cohorte de 4.000 médicos clínicos estuvo expuesta y la otra mitad no estuvo expuesta, esperaríamos 30 expuestos y 30 no expuestos de los 60 controles.

Estudios de casos y controles

		E	no E	Total
(incidentes)	Casos	40	20	60
	Controles	30	30	60
	Total	70	50	120

Podemos utilizar la fórmula de la razón de productos cruzados para calcular el *odds ratio* de exposición esperado, que da como resultado 2. Este valor del *odds ratio* de exposición obtenido de los datos de casos y controles es el mismo que hubiésemos obtenido de la razón de riesgos y el *odds ratio* de riesgo si hubiésemos

realizado el estudio de seguimiento de esta cohorte de población. En otras palabras, el EOR esperado de este estudio de casos y controles se aproximaría mucho a la RR del estudio de seguimiento correspondiente, aún cuando el estudio de seguimiento nunca se haya realizado.

$$\hat{EOR} = 2.0 \approx \hat{RR} = 2$$

Podríamos preguntarnos si el EOR se aproximaría a la RR incluso si los 60 controles no se dividieran de forma pareja entre los grupos de expuestos y no expuestos como se esperaba. Esto puede ocurrir de casualidad por una selección aleatoria o si hacemos un trabajo deficiente al elegir los controles. Por ejemplo, suponga que hubo 40 expuestos y 20 no expuestos entre los controles. Entonces, el *odds ratio* de exposición estimado sería igual a 1, en lugar de 2; por lo tanto, en esta situación, el EOR sería bastante diferente de la RR obtenida de un estudio de seguimiento comparable.

$$\hat{EOR} = 1.0 \not\approx \hat{RR} = 2$$

Estudios de casos y controles

		E	no E	Total
(incidentes)	Casos	40	20	60
	Controles	30	30	60
	Total	70	50	120

$$\hat{EOR} = 1.0$$

Lo que hemos mostrado con el ejemplo refleja, en verdad, una advertencia importante cuando se aplica el supuesto de enfermedad poco frecuente a los datos de casos y controles. La elección de los controles en un estudio de casos y controles debe ser representativa de la población de origen de donde se desarrollan los casos. De no ser así, ya sea de casualidad o por una elección deficiente de los controles, el *odds ratio* de exposición no necesariamente se aproximará a la razón de riesgos, incluso si la enfermedad es poco frecuente. Hay otra advertencia importante sobre la aplicación del supuesto de enfermedad poco frecuente en un estudio de casos y controles. Los casos deben ser casos incidentes, es decir, los casos necesitan incluir todos los casos nuevos que se desarrollaron durante el período de tiempo considerado para determinar la condición de exposición. Si los casos consistieron sólo de casos prevalentes en el momento de la constatación de casos, entonces podría surgir un valor sesgado, porque la medida de efecto estaría calculando la prevalencia y no la incidencia.

El EOR calcula la RR en un estudio de casos y controles sin el supuesto de enfermedad poco frecuente

La fórmula típica del EOR en un estudio de ***casos y controles***:

$$EOR = \frac{P(E \mid D)/[1 - P(E \mid D)]}{P(E \mid \overline{D})/[1 - P(E \mid \overline{D})]}$$

En un estudio de casos y cohorte, las muestras de los controles se toman de la cohorte original (completa); por lo tanto, la fórmula del EOR en un estudio de ***casos y cohorte*** puede modificarse en el denominador de la siguiente forma:

$$EOR = \frac{P(E \mid D)/[1 - P(E \mid D)]}{P(E)/[1 - P(E)]}$$

Por principios algebraicos, surge entonces que la formula para la RR en un estudio de *casos y cohorte* es:

$$RR = \frac{P(D \mid E)}{P(D \mid \overline{E})} = \frac{P(E \mid D)P(D)/P(E)}{P(\overline{E} \mid D)P(D)/P(\overline{E})}$$

$$= \frac{P(E \mid D)/P(E)}{P(\overline{E} \mid D)/P(\overline{E})} = \frac{P(E \mid D)P(\overline{E})}{P(\overline{E} \mid D)P(E)}$$

$$= \frac{P(E \mid D)/[1 - P(E \mid D)]}{P(E)/[1 - P(E)]}$$

$$= EOR$$

Resumen

- En los estudios de casos y controles, el EOR se aproxima a una RR cuando se satisfacen las siguientes 3 condiciones:

1) se aplica el supuesto de enfermedad poco frecuente.
2) La elección de los controles en un estudio de casos y controles debe ser representativa de la población de origen de donde se tomaron los casos.
3) Los casos deben ser casos incidentes.



Auto-evaluación (P5.12): comprensión de la razón de riesgos

En un estudio de casos y controles, si se satisface el supuesto de enfermedad poco frecuente, entonces:

1. El **???** se aproxima a la **???** siempre que no haya **???** en la selección de **???**, y que los casos sean **???** en vez de casos **???**

Opciones
EOR ROR RR sesgo casos controles incidentes prevalentes aleatoriedad

En una comunidad de 1 millón de personas, se informaron 100 casos de una enfermedad, distribuidos por exposición según el cuadro que aparece a continuación.

2. Calcule la RR. . . **???**
3. Calcule el ROR. . . **???**
4. ¿Esta enfermedad es poco frecuente? . . **???**

	Expuestos	No expuestos	Total
Enfermos	90	10	100
Sanos	499.910	499.990	999.900
Total	500.000	500.000	1.000.000

Si no hubiese estado disponible la exposición del millón de personas en la población del estudio, el investigador podría haber realizado un estudio de casos y controles. Supondremos que se seleccionó una muestra aleatoria de 100 controles.

5. ¿Aproximadamente qué porcentaje de estos controles esperaría que estén expuestos? **???**

6. ¿Cuál es el EOR esperado en un estudio de casos y controles? . **???**

Opciones
0,11 **10** **50** **9,00** **90** **no** **sí**

Las matemáticas detrás de la aproximación de enfermedad poco frecuente

El argumento matemático, que explica el por qué el *odds ratio* de exposición (EOR) se aproxima a la razón de riesgos (RR) cuando la enfermedad de interés es poco frecuente, puede exponerse en líneas generales como se detalla en los párrafos siguientes.

Primero demostramos que el *odds ratio* de riesgo (ROR) calculado a partir de datos de seguimiento se aproxima a la razón de riesgos (RR) para estos mismos datos si la enfermedad es poco frecuente. Luego, utilizando probabilidades condicionales, demostramos que el *odds ratio* de riesgo y el *odds ratio* de exposición son iguales. De esto surge que el *odds ratio* de exposición también debe aproximarse a la razón de riesgos para las enfermedades poco frecuentes.

Paso 1: $\hat{\text{ROR}} \approx \hat{\text{RR}}$ (si la enfermedad es poco frecuente)

Paso 2: $\hat{\text{ROR}} = \hat{\text{EOR}}$

Paso 3: $\hat{\text{EOR}} \approx \hat{\text{RR}}$

Describamos primero la razón de riesgos y el *odds ratio* de riesgo en términos de probabilidades condicionales. La razón de riesgos puede expresarse como la razón de la probabilidad condicional de desarrollar la enfermedad, si están expuestos, dividido por la probabilidad condicional de desarrollar la enfermedad, si no están expuestos. Para describir el *odds ratio* de riesgo, comenzamos con los *odds* de desarrollar la enfermedad para las personas expuestas y los *odds* de desarrollar la enfermedad para las personas no expuestas.

El *odds ratio* de riesgo es dado, entonces, por la razón de estos dos *odds*, como se muestra aquí. Aplicando un poco de álgebra, podemos reescribir el *odds ratio* de riesgo de la siguiente manera:

Paso 1: $\hat{ROR} \approx \hat{RR}$ (si la enfermedad es poco frecuente)

$$RR = \frac{P(D|E)}{P(D|no\ E)}$$

$$ROR = \frac{\text{(odds de D dado E)}}{\text{(odds de D dado no E)}}$$

$$= \frac{\frac{P(D|E)}{1-P(D|E)}}{\frac{P(D|no\ E)}{1-P(D|no\ E)}} = \frac{P(D|E)}{P(D|no\ E)} * \frac{1-P(D|not\ E)}{1-P(D|E)}$$

Por lo tanto, podemos expresar el *odds ratio* de riesgo en términos de la razón de riesgos, reemplazando la primera parte del término del producto en la expresión de la derecha por la razón de riesgos, de forma tal que obtenemos la siguiente fórmula:

$$ROR = RR * \frac{1-P(D|no\ E)}{1-P(D|E)}$$

Ahora, si la enfermedad es poco frecuente, la probabilidad de la enfermedad es baja, independientemente de la exposición. Por consiguiente, tanto P(D|no E) como P(D|E) son aproximadamente igual a cero. Si sustituimos por cero estos términos en la fórmula del *odds ratio* de riesgo, obtenemos una expresión que dice que el *odds ratio* de riesgo se aproxima a la razón de riesgos cuando la enfermedad es poco frecuente.

$$ROR = RR * \frac{1-0}{1-0} \approx RR * 1$$

Enfermedad poco frecuente $\Rightarrow P(D|no\ E) \approx 0$ y $P(D|E) \approx 0$

Ahora utilizamos un famoso teorema sobre probabilidades condicionales, llamado **teorema de Bayes**, para mostrar que el *odds ratio* de exposición es igual al *odds ratio* de riesgo, de lo que surge lógicamente que el *odds ratio* de exposición se aproxima a la razón de riesgos para las enfermedades poco frecuentes. El teorema de Bayes expresa probabilidades condicionales de la forma P(D|E) y P(D|no E) en términos de probabilidades condicionales de la forma P(E|D) y P(E|no D). Necesitamos considerar P(E|D) y P(E|no D) porque, en los estudios de casos y controles, primero se da la condición de enfermedad y luego se determina la probabilidad condicional de la condición de exposición previa. Mediante el teorema de Bayes, podemos expresar P(D|E) y P(D|no E) de la siguiente forma:

Paso 2: $\hat{ROR} = \hat{EOR}$ Teorema de Bayes

$$P(D|E) = \frac{P(D)\,P(E|D)}{P(D)\,P(E|D) + [1 - P(D)]\,P(E|\text{no } D)}$$

$$P(D|\text{no } E) = \frac{P(D)\,[1-P(E|D)]}{P(D)\,[1-P(E|D)] + [1 - P(D)]\,[1-P(E|\text{no } D)]}$$

Si luego sustituimos estas expresiones por P(D|E) y P(D|no E) en la fórmula del *odds ratio* de riesgo y aplicamos una cantidad considerable de álgebra mediante los términos sustituidos, obtenemos la siguiente expresión. La expresión de la derecha es el *odds ratio* de exposición.

$$ROR = \frac{\frac{P(D|E)}{1- P(D|E)}}{\frac{P(D|\text{no } E)}{1- P(D|\text{no } E)}} = \frac{\frac{P(E|D)}{1- P(E|D)}}{\frac{P(E|\text{no } D)}{1- P(E|\text{no } D)}} = EOR$$

Por lo tanto, hemos demostrado que el *odds ratio* de riesgo y el *odds ratio* de exposición son iguales. Combinando el resultado de que el *odds ratio* de riesgo se aproxima a la razón de riesgos cuando la enfermedad es poco frecuente con el resultado de que el *odds ratio* de riesgo es igual al *odds ratio* de exposición, surge que el *odds ratio* de exposición se aproxima a la razón de riesgos cuando la enfermedad es poco frecuente.

√ Paso 1: $\hat{ROR} \approx \hat{RR}$ (si la enfermedad es poco frecuente)

√ Paso 2: $\hat{ROR} = \hat{EOR}$

√ Paso 3: $\hat{EOR} \approx \hat{RR}$ (si la enfermedad es poco frecuente)

El argumento matemático que acabamos de completar para la equivalencia del *odds ratio* de exposición y el *odds ratio* de riesgo requiere dos supuestos adicionales. Estos supuestos son necesarios para pasar de probabilidades teóricas a sus valores calculados a partir de datos de casos y controles. En primer lugar, la elección de los controles en un estudio de casos y controles debe ser representativa de la población de origen de donde se tomaron los casos. En segundo lugar, los casos deben ser incidentes, y no prevalentes.

Resumen: las matemáticas detrás de la aproximación de enfermedad poco frecuente

- Utilice principios algebraicos que involucren probabilidades condicionales y el teorema de Bayes.
- Teorema de Bayes: probabilidades condicionales de la forma P(D|E) y P(D|no E) en términos de P(E|D) y P(E|no D).
- También se requieren dos supuestos: controles representativos y casos incidentes.
- Paso 1: el ROR calculado a partir de datos de seguimiento se aproxima a la RR para estos mismos datos si la enfermedad es poco frecuente.
- Paso 2: utilizando el teorema de Bayes, el ROR y el EOR son iguales.

- Paso 3: combinando el paso 1 y el paso 2, el EOR se aproxima a la RR para las enfermedades poco frecuentes.

5-4 La razón de tasas y sus características

La razón de tasas

Una **razón de tasas** es una razón entre dos tasas promedio. También suele llamársele **razón de densidades de incidencia** o **razón de riesgos instantáneos.** Recuerde la fórmula general para una tasa promedio: **I** indica la cantidad de casos nuevos del resultado en la salud, y **PT** indica el valor persona-tiempo acumulado durante el seguimiento.

A continuación se muestra un modelo general de organización de datos para calcular una razón de tasas. $\mathbf{I_1}$ e $\mathbf{I_0}$ hacen referencia a la cantidad de casos nuevos en los grupos expuestos y no expuestos, y $\mathbf{PT_1}$ y $\mathbf{PT_0}$ hacen referencia al valor persona-tiempo acumulado correspondiente a estos dos grupos. También se brinda la fórmula para la **razón de tasas** o **razón de densidades de incidencia (RDI, por sus siglas en español o IDR, por sus siglas en inglés)**. Hemos utilizado la notación IDR para denotar la razón de tasas, a fin de evitar confusión con nuestro uso previo de RR para denotar la razón de riesgos.

$$\text{Tasa promedio: } \frac{I}{PT}$$

Modelo general para calcular una razón de tasas (i.e., IDR)

	Expuesto	No expuesto	Total
Casos Nuevos	I_1	I_0	I
Persona-tiempo	PT_1	PT_0	PT

Razón de tasas:

$$IDR = \frac{\frac{I_1}{PT_1}}{\frac{I_0}{PT_0}}$$

De la misma forma que con las medidas de razón de riesgos y de *odds ratio*, la razón de tasas puede ser >1, <1, o =1. Si la razón de tasas es igual a 1, significa que no hay relación entre la exposición y la enfermedad utilizando esta medida de efecto.

Para ejemplificar el cálculo de una razón de tasas, consideramos los datos sobre la relación entre el nivel de colesterol en sangre y la mortalidad de un estudio de casi 40.000 personas del área de Chicago realizado en 1992. Los datos mostrados comparan a hombres blancos con niveles de colesterol moderadamente altos y hombres blancos con niveles de colesterol normales. Los individuos, incluyendo las personas de otras categorías de raza y sexo, fueron incluidos en el estudio entre 1967 y 1973, se les realizaron estudios para detectar factores de riesgo de enfermedad cardiovascular (CVD, por sus siglas en inglés) y luego se les realizó un seguimiento durante un promedio de 14 a 15 años. Sobre la base de 36.581 personas-años de seguimiento, hubo un total de 26 muertes relacionadas con la CHD entre hombres blancos de 25 a 39 años con colesterol moderadamente alto al incorporarse al estudio. Esto da una tasa de 71,1 muertes por 100.000 personas-años. Entre el grupo de comparación hubo 14 muertes relacionadas con la CHD sobre la base de 68.239 personas-años de seguimiento, lo

que da una tasa de 20,5 muertes por 100.000 personas-años. Por lo tanto, los hombres blancos de 25-39 años con colesterol moderadamente alto tienen una tasa de mortalidad 3,5 veces más alta que aquellos con colesterol normal, lo que indica que incluso las personas con colesterol moderadamente alto tienen un mayor riesgo de mortalidad por CHD.



1992, 40,000 personas, Chicago (hombres blancos, edad 25 - 39)

	Nivel de colesterol	
	Moderadamente alto	Normal
Muertes	26	14
Persona-años	36,581	68,239
$\frac{I}{PT}$	$\frac{71.1}{100{,}000}_{\text{persona-años}}$	$\frac{20.5}{100{,}000}_{\text{persona-años}}$

$$\hat{IDR} = \frac{\frac{I_1}{PT_1}}{\frac{I_0}{PT_0}} = \frac{71.1}{20.5} = \mathbf{3.5}$$

Resumen: razón de tasas

- La razón entre dos tasas promedio se llama razón de tasas (es decir, razón de densidades de incidencia, razón de riesgos instantáneos).
- La fórmula para la razón de tasas (IDR, por sus inglés) es:

$$IDR = \frac{\frac{I_1}{PT_1}}{\frac{I_0}{PT_0}}$$

donde I_1 e I_0 son la cantidad de casos nuevos y PT_1 y PT_0 corresponden al valor persona-tiempo acumulado para los grupos 1 y 0, respectivamente.

- Al igual que con la razón de riesgos o el *odds ratio*, la razón de tasas puede ser >1, <1 o =1.

Aproximación del *odds ratio* a la razón de tasas

Ya hemos visto que, en los estudios de casos y controles, el *odds ratio* de exposición calcula la razón de riesgos para el estudio de seguimiento correspondiente, siempre que el resultado en la salud sea poco frecuente. En esta actividad demostramos que, en ciertas condiciones, el *odds ratio* que surge de un estudio de casos y controles calcula la razón de tasas a partir de un estudio de cohorte comparable que utiliza información de persona-tiempo. No es necesario suponer una enfermedad poco frecuente para que esta aproximación sea válida. La condición clave requerida es que la población de origen de la cual se derivan los casos y controles esté en **estado estacionario**, un término que definiremos a la brevedad.

Casos y Controles Seguimiento (cohorte)

$\hat{EOR} \approx \hat{RR}$

si la enfermedad es poco frecuente

Casos y Controles Seguimiento (cohorte)

$\hat{EOR} \approx \hat{IDR}$

enfermedad no poco frecuente
población de origen tiene que estar en

Estado estacionario

Supondremos que la población de origen en el momento t_0 contiene N_1 personas expuestas que no padecen la enfermedad y N_0 personas no expuestas que no padecen la enfermedad. Supondremos, además, que después de Δt años de seguimiento, se desarrollaron I_1 casos nuevos de la enfermedad de aquellos expuestos en el momento T_0 e I_0 casos nuevos de aquellos no expuestos en el momento T_0. Supondremos, asimismo, que la población no sufre cambios demográficos importantes, de forma que el tamaño de la población de origen es básicamente constante durante los Δt años de seguimiento. Entonces, se considera que la población de origen aquí es estable o que se encuentra en estado estacionario.

Seguimiento de la población de origen para Δt años

	Expuesto	No expuesto
Casos nuevos en $(t_0, t_0+\Delta t)$	I_1	I_0
No casos	$N_1 - I_1$	$N_0 - I_0$
$\text{Total}_{\text{sin enfermedad en momento } t_0}$	N_1	N_0

Población de origen es constante $\Rightarrow$ estado estacionario

En dichas condiciones de estado estacionario, se puede aproximar la cantidad de personas-años de observación para cada grupo de exposición utilizando la fórmula $PT = N^* \times \Delta t$, donde N^* denota el tamaño de esta población de origen estable para cada grupo de exposición. Por lo tanto, la cantidad de personas-años para las personas expuestas es $N_1 \times \Delta t$ y la cantidad de personas-años para las personas no expuestas es $N_0 \times \Delta t$.

Seguimiento de la población de origen para Δt años

Condición estado estacionario

	Expuesto	No expuesto
Casos nuevos en $(t_0, t_0+\Delta t)$	I_1	I_0
Total sin enfermedad en momento t_0	N_1	N_0

$$PT = N * \Delta t$$

$$\text{Rázon de tasas} = \hat{IDR} = \frac{\frac{I_1}{PT_1}}{\frac{I_0}{PT_0}} = \frac{\frac{I_1}{N_1 \Delta t}}{\frac{I_0}{N_0 \Delta t}}$$

La razón de tasas para esta cohorte, que hemos denotado IDR, pos sus siglas en inglés, surge entonces de sustituir las expresiones de PT_1 y PT_0 en la fórmula de IDR para obtener la expresión que se muestra aquí.

Podemos cancelar Δt en esta expresión para simplificarla de la siguiente manera:

$$\hat{IDR} = \frac{\frac{I_1}{I_0}}{\frac{N_1}{N_0}}$$

Ahora, supondremos que realizamos un estudio de casos y controles utilizando esta población de origen. Seleccionamos una muestra aleatoria de casos incidentes que se desarrollaron entre los momentos t_0 y $t_0 + \Delta t$ y una muestra aleatoria de controles de la población de origen. Luego determinamos la condición de exposición previa para los casos y controles. Esto nos da la siguiente tabla de doble entrada para los datos de casos y controles.

Forma general de tabla 2x2

	Estado de exposición Sí	No	Odds de Exposición
Casos	a	b	$a/b = I_1 / I_0$
Controles	c	d	$c/d = N_1 / N_0$ población de origen

$$\hat{IDR} = \frac{I_1 / I_0}{N_1 / N_0} = \frac{a/b}{c/d} = \frac{ad}{bc} = \hat{EOR}$$

De este cuadro, los *odds* de exposición para los casos son dados por a/b y los *odds* de exposición para los controles son dados por c/d. Suponiendo que el muestreo es "ciego" a la condición de exposición, esperaríamos que la proporción de casos en el estudio que estuvieron expuestos fuese igual, en promedio, a la proporción de casos en la cohorte completa que estuvieron expuestos. Aplicando un poco de álgebra, surge que a/b, los *odds* de exposición entre los casos, es igual a I_1/I_0, los *odds* de exposición correspondientes en la población de origen. De forma similar, puede demostrarse que los *odds* de exposición para los controles c/d deberían, en promedio, ser iguales a los *odds* de exposición entre las personas sin la enfermedad en la población de origen. Si ahora

sustituimos a/b por I_1/I_0 y c/d por N_1/N_0 en la fórmula de la razón de tasas, obtenemos ad/bc, el *odds ratio* del estudio de casos y controles.

Por lo tanto, hemos demostrado que en condiciones de estado estacionario y utilizando los casos incidentes, el *odds ratio* de exposición se aproximará a la razón de tasas, sin requerir que la enfermedad sea poco frecuente.

Resumen: cuando el *odds ratio* de exposición se aproxima a la razón de tasas

- En condiciones de estado estacionario, el *odds ratio* de un estudio de casos y controles se aproximará a la razón de tasas de un estudio de cohorte comparable que utiliza información de persona-tiempo.
- Esta aproximación no requiere el supuesto de enfermedad poco frecuente.
- *Estado estacionario* significa que no hay un cambio importante en la composición demográfica de la población de origen.

Auto-evaluación (P5.13)

A continuación se muestran los datos para un estudio de seguimiento que tiene como propósito comparar las tasas de mortalidad entre diabéticos y no diabéticos.

1. La tasa de mortalidad para los diabéticos es . **???**
2. La tasa de mortalidad para los no diabéticos es . . **???**
3. La razón de tasas es . . . **???**

Opciones

13,9 **13,9 por 1.000 personas-años** **2,8** **2,8 por 1.000 personas-años**
38,7 **38,7 por 1.000 personas-años**

	Diabéticos	No diabéticos	Total
Muertos	72	511	583
Vivos	146	3.312	3.458
Personas-años	1.862,4	36.532,2	38.394,6

4. La razón de tasas que compara las tasas de mortalidad de los diabéticos y no diabéticos es de 2,8. ¿Cuál de las siguientes es la interpretación correcta de esta medida?

   A. Los que sufren diabetes tienen 2,8 veces más probabilidades de morirse que los que no la padecen.
   B. Las personas tienen 2,8 veces más probabilidades de morirse por diabetes que por cualquier otra enfermedad.
   C. Los diabéticos mueren a una tasa 2,8 veces más alta que la de los no diabéticos.

Nomenclatura

Organización en forma de tabla para los estudios de cohorte, casos y controles y prevalencia:

	Expuestos	No expuestos	Total
Enfermedad presente/casos	a	b	m_1
Sin enfermedad/controles	c	d	m_0

Total	n_1	n_0	n

Organización en forma de tabla para los datos de la cohorte con persona-tiempo:

	Expuestos	No expuestos	Total
Enfermedad (casos nuevos)	I_1	I_0	I
Sin enfermedad	-	-	-
Persona-tiempo total sin enfermedad	PT_1	PT_0	PT

Δt Cambio en el tiempo o duración de tiempo
EOR *Odds ratio* de exposición; *odds* de exposición en los enfermos dividido por los *odds* de exposición en los sanos
I Incidencia promedio o número total de casos nuevos
I_0 Número de casos nuevos en los no expuestos
I_1 Número de casos nuevos en los expuestos
RDI Razón de densidades de incidencia ("razón de tasas"): RDI=tasa en los expuestos/tasa en los no expuestos, o IDR, por sus siglas en inglés
N Tamaño de la población bajo estudio
N_0 Tamaño de la población **no expuesta** bajo estudio en el tiempo t = 0
N_1 Tamaño de la población **expuesta** bajo estudio en el tiempo t = 0
OR *Odds ratio (OR)*: ad/bc
P Probabilidad de un evento
P(D|E) Probabilidad de la enfermedad dada la exposición
P(D|no E) Probabilidad de la enfermedad dada la ausencia de la exposición
P(E|D) Probabilidad de la exposición dada la enfermedad
P(E|no D) Probabilidad de la exposición dada la ausencia de la enfermedad
POR *Odds ratio* de prevalencia; un *OR* calculado con datos de prevalencia
PT Persona-tiempo contribuida libre de la enfermedad
PT_0 Persona-tiempo contribuida libre de la enfermedad en los **no expuestos**
PT_1 Persona-tiempo contribuida libre de la enfermedad en los **expuestos**
R_0 Riesgo en los **no expuestos**
R_1 Riesgo en los **expuestos**
ROR *Odds ratio* de riesgo; un *OR* calculado a partir de los datos de riesgo de la cohorte
RR Razón de riesgos: el riesgo en los expuestos dividido por el riesgo en los no expuestos
T o **t** Tiempo

Fórmulas

$RDI= (I_1/PT_1) / (I_0/PT_0)$

***Odds* = P / (1-P)**

***Odds ratio* = ad/bc**

$ROR = RR * f$ donde $f=(1-R_0)/(1-R_1)$

$RR=R_1 / R_0$

Referencias

Doll R, and Hill AB. *Smoking and carcinoma of lungs: preliminary report*. Br Med J 1950;2:739-48.

Dyer AR, Stamler J, Shekelle RB. *Serum cholesterol and mortality from coronary heart disease in young, middle-aged, and older men and women in three Chicago epidemiologic studies*. Ann of Epidemiol 1992;2(1-2): 51-7.

Greenberg RS, Daniels SR, Flanders WD, Eley JW, Boring JR. *Medical Epidemiology* (3° edición). Lange Medical Books, Nueva York, 2001.

Hammond EC, Horn D. *The relationship between human smoking habits and death rates*. JAMA 1958;155:1316-28.

Johansson S, Bergstrand R, Pennert K, Ulvenstam G, Vedin A, Wedel H, Wilhelmsson C, Wilhemsen L, Aberg A. *Cessation of smoking after myocardial infarction in women. Effects on mortality and reinfarctions*. Am J Epidemiol 1985;121(6): 823-31.

Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic Research: Principles and Quantitative Methods*. John Wiley and Sons Publishers, Nueva York, 1982.

Spika JS, Dabis F, Hargett-Bean N, Salcedo J, Veillard S, Blake PA. *Shigellosis at a Caribbean Resort. Hamburger and North American origin as risk factors*. Am J Epidemiol 1987;126 (6): 1173-80.

Steenland K (ed.). *Case studies in occupational epidemiology*. Oxford University Press, Nueva York, NY: 1993.

Uitterlinden AG, Burger H, Huang Q, Yue F, McGuigan FE, Grant SF, Hofman A, van Leeuwen JP, Pols HA, Ralson SH. *Relation of alleles of the collagen type I alpha1 gene to bone density and the risk of osteoporotic fractures in postmenopausal women*. N Engl J Med 1998;338(15):1016-21.

Wynder EL, Graham EA. *Tobacco smoking as a possible etiologic factor in bronchogenic carcinoma: a study of six hundred and eighty-four proved cases*. JAMA 1950;143:329-36.

Ejercicios

EPI-1. Medidas de efecto: mascar tabaco versus leucoplaquia oral

Se realiza un estudio para investigar la asociación entre mascar tabaco y la leucoplaquia oral (una lesión precancerosa) entre jugadores de baseball profesional actualmente activos en el sureste de Estados Unidos. Se obtiene una lista de todos los jugadores activos (n=500). Todos los individuos potenciales del estudio acuerdan participar. Cada individuo tiene una entrevista sobre el hábito actual de mascar tabaco y un dentista examina su boca. De los 500 individuos, 125 individuos mascan tabaco y 375 no lo hacen. De los que mascan tabaco, 24 muestran pruebas de leucoplaquia oral. De los que no mascan tabaco, 15 muestran pruebas de leucoplaquia oral. Se realizó un seguimiento de la totalidad de los 500 jugadores durante un período de 5 años. De los que habían mostrado pruebas de leucoplaquia oral, 18 murieron como consecuencia de algún tipo de cáncer oral.

a. Grafique un cuadro de 2x2 donde demuestre la relación entre mascar tabaco y la leucoplaquia oral. Al graficar este cuadro, coloque la variable exposición en las columnas y la variable de resultado de salud en las filas.
b. Grafique un segundo cuadro de 2x2 donde demuestre la relación entre mascar tabaco y la leucoplaquia oral, pero esta vez, coloque la variable exposición en las filas y la variable de resultado de salud en las columnas.
c. Mediante el cuadro graficado en el ítem a, calcule la razón de prevalencia y el *odds ratio* de prevalencia de la leucoplaquia oral para los que mascan tabaco en comparación con los que no mascan tabaco. ¿Estos dos valores se acercan? ¿Por qué éstas son medidas de prevalencia y no de incidencia?
d. Mediante el cuadro graficado en el ítem b, calcule la razón de prevalencia y el *odds ratio* de prevalencia de la leucoplaquia oral para los que mascan tabaco en comparación con los que no mascan tabaco. ¿Estos valores son iguales a sus valores correspondientes calculados utilizando el cuadro graficado en el ítem a? Desarrolle su respuesta.
e. Ignorando el tema de la inferencia estadística y el control de otras variables, ¿qué dicen estos resultados sobre la relación entre mascar tabaco y la presencia de leucoplaquia oral?
f. Calcule la tasa de letalidad (en verdad, un riesgo) en este estudio. ¿Por qué ésta es una medida de riesgo?

g. Basándose en la información brindada, ¿por qué no puede evaluar si los que mascan tabaco tienen un riesgo de letalidad más alto que los que no mascan tabaco?

EPI-2. Razones de tasas: muertes por cáncer de colon

El cuadro siguiente muestra la cantidad de muertes por cáncer de colon y la cantidad de personas-años de riesgo por la frecuencia de aspirina para hombres y mujeres.

Cuadro 1. Tasas de mortalidad como consecuencia de cáncer de colon, según la frecuencia de uso de aspirina en la cohorte ***antes*** de que se excluyeran los pacientes con la enfermedad al momento de la inclusión en el estudio.

		Uso de aspirina (veces por mes)			
		0	<1	1-15	≥16
Hombres	Número de muertes	378	184	127	85
	Personas-años en riesgo	646.346	486.620	389.083	201.636
	Tasa de mortalidad por 100.000	_____	_____	_____	_____
	Razón de tasas	1,00	_____	_____	_____
Mujeres	Número de muertes	284	157	100	73
	Personas-años en riesgo	705.064	671.927	505.854	265.424
	Tasa de mortalidad por 100.000	_____	_____	_____	_____
	Razón de tasas	1,00	_____	_____	_____

a. Calcule las tasas de mortalidad y las razones de tasas para cada una de las categorías de uso de aspirina en el cuadro anterior.
b. Dados los resultados en este cuadro, ¿cuál es su conclusión sobre la asociación entre el uso de aspirina y el cáncer de colon mortal?

EPI-3. Razones de tasas: AINES

Los AINES (es decir, antiinflamatorios no esteroides) se recetan o se adquieren sin receta en el caso de enfermedades agudas y crónicas, percibidas y diagnosticadas. Por esta razón, los investigadores en el estudio descrito en la pregunta 2 también analizaron los datos excluyendo a aquellos individuos con enfermedades seleccionadas al comienzo del seguimiento. El cuadro 2 muestra la cantidad de muertes por cáncer de colon por la frecuencia de uso de aspirina después de la exclusión de aquellos individuos con enfermedades seleccionadas.

Cuadro 2. Tasas de mortalidad por cáncer de colon, según la frecuencia de uso de aspirina en la cohorte ***después*** de que se excluyeran los pacientes con la enfermedad al momento de la inclusión en el estudio.

		Uso de aspirina (veces por mes)			
		0	<1	1-15	≥16
Hombres	Cantidad de muertes	171	101	63	28
	Personas-años en riesgo	487.932	385.321	302.106	116.947
	Tasa de mortalidad por 100.000	_____	_____	_____	_____
	Razón de tasas	1,00	_____	_____	_____
Mujeres	Cantidad de muertes	126	98	54	32
	Personas-años en riesgo	521.467	531.469	396.956	175.409

Continúa

	Tasa de mortalidad por 100.000	_____	_____	_____	_____
	Razón de tasas	1,00	_____	_____	_____

a. Calcule las tasas y las razones de tasas para cada una de las categorías de uso de aspirina en el cuadro anterior.
b. Dados los resultados del cuadro 1 (de la pregunta 2) y el cuadro 2, ¿cuál es su conclusión sobre la asociación entre el uso de aspirina y el cáncer de colon mortal?

EPI-4. Medidas de efecto de incidencia: abandono del tabaquismo

Los siguientes datos surgen de un estudio de enfoques de autoayuda para dejar de fumar. Los fumadores que querían dejar de fumar fueron clasificados en forma aleatoria en uno de cuatro grupos (C = control, M = manual para dejar de fumar, MS = manual + folleto de apoyo social, MST = manual + folleto de apoyo social + asesoría por teléfono). La condición de tabaquismo se midió por medio de encuestas enviadas por correo, 8, 16 y 24 meses después de la formación aleatoria de los grupos. Estos son los resultados del mes 16:

	Grupo de formación aleatoria				
Condición	**C**	**M**	**MS**	**MST**	**Total**
Dejaron de fumar	84	71	67	109	331
Fuman	381	396	404	365	1.546
Total	465	467	471	474	1.877

a. La "tasa de abandono del tabaquismo" se calcula como la proporción en abstinencia (los que dejaron de fumar) al momento del seguimiento. ¿Cuál fue la tasa general de abandono del tabaquismo a los 16 meses para todos los individuos? Basándose en las tasas de abandono del tabaquismo, ¿cuál de los grupos de intervención fue el menos exitoso hasta el seguimiento de 16 meses? Justifique su respuesta.
b. ¿Es una medida de tipo incidencia acumulada o de tipo densidad de incidencia esta "tasa de abandono del tabaquismo"? Justifique su respuesta.
c. Compare la tasa de abandono del tabaquismo para el grupo MST con la del grupo de control, calculando tanto una CIR como un OR. Muestre su trabajo. Brinde una interpretación de la CIR.

EPI-5. Tasa de densidad de incidencia: la radioterapia entre los niños.

En un estudio sobre los efectos adversos de la radioterapia en los niños de Israel, de los registros originales del tratamiento, se identificaron 10.834 niños irradiados. Éstos se combinaron con 10.834 individuos de comparación no irradiados que fueron seleccionados de la población general. Se realizó un seguimiento de los individuos durante una media de 26 años. La cantidad de personas-años de observación fue: individuos irradiados, 279.901 personas-años; individuos de comparación, 280.561 personas-años. Durante el período de seguimiento hubo 49 muertes por cáncer entre los individuos irradiados, y 44 entre los individuos de comparación de la población no irradiados.

a. ¿Cuáles son las tasas de mortalidad por cáncer (por 105 personas-años) en cada uno de los dos grupos?
b. Calcule e interprete la razón de tasas (IDR, por sus siglas en inglés) de la mortalidad por cáncer, comparando los individuos irradiados con los no irradiados.

EPI-6. ***Odds ratio*: consumo de alcohol vs. infarto de miocárdio.**

Se realizó un estudio de casos y controles para evaluar la relación entre el consumo de alcohol y el infarto de miocardio (MI, por sus siglas en inglés). Los casos fueron hombres de entre 40 y 65 años que habían sufrido su primer MI durante los seis meses anteriores a la inclusión en el estudio. Un grupo de hombres de las mismas edades que nunca habían experimentado un MI fueron seleccionados como controles. A continuación se resumen los datos de este estudio:

	Expuestos	**No expuestos**	**Totales**
Enfermos	158	201	359
Sanos	252	170	422
Totales	410	371	781

a. ¿Cuáles son los *odds* de exposición entre los controles?
b. Calcule e interprete el *odds ratio* de exposición para estos datos.
c. ¿Cree que el OR calculado en el ítem b anterior representa un buen cálculo de la razón de riesgos correspondiente para la relación entre el alcohol y el OR? ¿Por qué sí o por qué no?

Respuestas a las preguntas de estudio y auto-evaluaciones

P5.1

1. El riesgo en cinco años para los fumadores que continúan fumando es 4½ veces más alto que el riesgo para los fumadores que dejan de fumar.
2. La razón de riesgos es muy cercana a 1,0, lo que indica que no hay una diferencia relevante entre los riesgos de los dos grupos.
3. Piense en una situación inversa.
4. Ya debería poder hacer esto sin problemas.

P5.2

1. 2
2. 1
3. 0,5
4. 0,0104
5. 0,0236
6. 0,44 - En general, la razón de riesgos que compara dos grupos se define como el riesgo para un grupo dividido por el riesgo para el otro grupo. Es importante especificar claramente qué grupo está en el numerador y qué grupo está en el denominador. Si, por ejemplo, los dos grupos se llaman *grupo 1* y *grupo 0,* y el riesgo para el grupo 1 está en el numerador, entonces, decimos que la razón de riesgos compara el grupo 1 con el grupo 0.

P5.3

1. Los *odds* de que un caso coma una hamburguesa cruda son aproximadamente dos veces y media mayores que los *odds* de que un individuo de control coma una hamburguesa cruda.
2. Debido a que el odds ratio es tan cercano a uno, esto sería considerado un caso nulo, lo que implica que los *odds* de que un caso coma una hamburguesa cruda son aproximadamente iguales a los *odds* de que un individuo de control coma una hamburguesa cruda.
3. Un *odds ratio* menor que 1 significa que los *odds* de que un individuo que es un caso coma una hamburguesa cruda son menores que los *odds* de que un individuo de control coma una hamburguesa cruda.
4. Ya debería poder hacer esto sin problemas.

P5.4

1. Imposible, *odds ratio* – El riesgo de enfermedad es definido como la proporción de la población inicialmente sin la enfermedad que subsecuentemente desarrolló la enfermedad bajo estudio durante un período determinado. En general, en un estudio de casos y controles no se puede determinar el riesgo.
2. 1.350/1.357
3. 1.296/1.357
4. 192,86
5. 21,25
6. 9,08 - En general, el *odds ratio* se define como: *odds,* o posibilidades, de exposición de los casos de enfermedad dividido por *odds,* o posibilidad, de exposición de los controles. Los *odds* para cada grupo pueden calcularse mediante la fórmula P/(1-P), donde P es la probabilidad de exposición.
7. 3
8. 1
9. 0,333

P5.5

1. Definitivamente. Es posible, por ejemplo, que la mayonesa contuviese, en verdad, la bacteria que causó el brote y quizás la mayoría de los casos que comieron una hamburguesa cruda utilizaron mayonesa.

P5.6

1. 683, 86, 0,77
2. *odds*, exposición, menor que, *odds*, controles – Si el *odds ratio* calculado es menor que 1, entonces, los *odds* de exposición para los casos son menores que los *odds* de exposición para los controles. Si el *odds ratio* calculado es mayor que 1, entonces los *odds* de exposición para los casos son mayores que los *odds* de exposición para los controles.

P5.7

1. 2,18
2. mayor probabilidad
3. prevalencia

P5.8

1. Eso depende de la enfermedad en estudio y del período de tiempo de seguimiento a lo largo del cual se calcula el riesgo. Sin embargo, para la mayoría de las enfermedades crónicas y para períodos de tiempo cortos, un riesgo de 0,01 no es poco frecuente.
2. Sí, ya que aunque el riesgo puede no ser poco frecuente, puede ser lo suficientemente pequeño para que la ROR y la RR sean aproximadamente iguales.

P5.9

1. $f = (1 - 0{,}17) / (1 - 0{,}36) = 1{,}30$
2. No, ya que para estos datos, la RR es igual a 2,1, mientras que la ROR es igual a 2,7.
3. $f = (1 - 0{,}085) / (1 - 0{,}180) = 1{,}12$
4. Sí, ya que la RR es nuevamente 2,1, (0,180/0,085), pero la ROR es 2,4.
5. f=1,05
6. Sí, ya que la ROR es ahora 2,2.
7. En el contexto del ejemplo de abandono del tabaquismo, los riesgos menores que 0,10 para ambos grupos indican una enfermedad “poco frecuente”.

P5.10

1. La razón de riesgos en este estudio es 0,0805 dividido por 0,0536, que da como resultado 1,50.
2. El *odds ratio* de riesgo es 47/537 dividido por 64/1130, lo cual es igual a 1,54.
3. f= (1-0,0536) / (1-0,0805) = 1,03. La ROR = 1,03*RR = 1,03*1,50=1,54.

P5.11

1. 9,8
2. 36,0
3. No
4. No - La razón de riesgos que compara dos grupos se define como el riesgo para un grupo dividido por el riesgo para el otro grupo. El *odds ratio* puede calcularse mediante la fórmula de productos cruzados ad/bc. En general, una enfermedad se considera "poco frecuente" cuando el OR se aproxima mucho a la RR.
5. 2,44
6. 2,49
7. Sí
8. Sí

P5.12

1. EOR, RR, sesgo, controles, incidentes, prevalentes
2. 9
3. 9
4. Sí - Una enfermedad se considera "poco frecuente" cuando el ROR se aproxima mucho a la RR.
5. 50
6. 9,00

P5.13

1. 38,7 por 1.000 personas-años – La tasa de mortalidad para los diabéticos es igual a 72/1.862,4 personas-años = 38,7 por 1.000 personas-años.
2. 13,9 por 1.000 personas-años – La tasa de mortalidad para los no diabéticos es igual a 511/36.653,2 personas-años = 13,9 por 1.000 personas-años.
3. 2,8 – La razón de tasas es 38,7 por 1.000 personas-años/13,9 por 1.000 personas-años = 2,8.
4. C

LECCIÓN 6

MEDIDAS DE IMPACTO POTENCIAL

6-1

*En la lección anterior, nos concentramos exclusivamente en **las medidas de efecto en forma de razones**. En esta lección, consideramos **las medidas de efecto en forma de diferencias** y otras medidas relacionadas que permiten al investigador considerar el impacto potencial en la salud pública de los resultados obtenidos en un estudio epidemiológico.*

Diferencia de riesgos: ejemplo

*La **diferencia de riesgos** es la diferencia entre dos estimaciones de riesgo, mientras que la **razón de riesgos** es la razón de dos estimaciones de riesgo. Ejemplificaremos una diferencia de riesgos mediante un estudio de cohorte de pacientes que sufrieron un ataque cardíaco y que continuaron fumando o dejaron de fumar después de dicho ataque.*

Considere nuevamente los resultados de un estudio de seguimiento de cinco años para determinar si los fumadores que han tenido un ataque cardíaco reducirán su riesgo de morir si dejan de fumar. La razón de riesgos es 2,1; esto significa que el riesgo para los fumadores que continuaron fumando era 2,1 veces el riesgo para los fumadores que dejaron de fumar.

Diferencia de Riesgos

Pacientes ataque cardíaco	Fuma	Dejó	**Total**
Muerte	27	14	41
Sobrevivencia	48	67	115
Total	75	81	156

Riesgo de morir en 5 años
continúan fumando: 27/75 = 0.36
dejaron de fumar: 14/81 = 0.17

Razón de Riesgos = 0.36/ 0.17 = **2.1**

Diferencia de Riesgos = 0.36 - 0.17 = .19

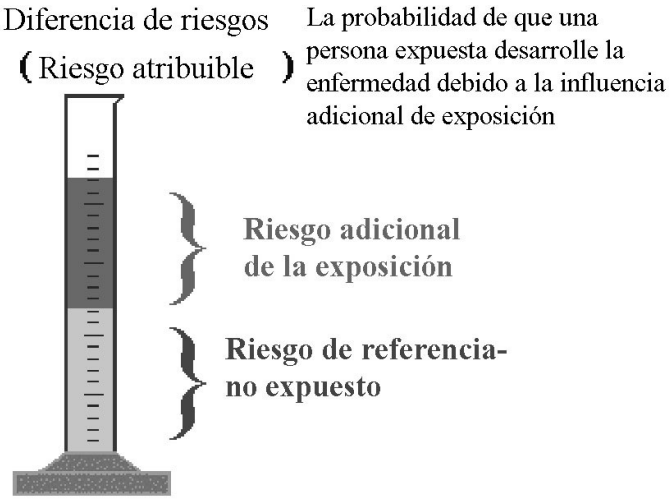
Ahora nos concentramos en la diferencia entre las dos estimaciones de riesgo calculadas, en lugar de la razón de riesgos. ¿Qué tipo de interpretación podemos hacer de este cálculo de diferencia? La **diferencia de riesgos** (**DR**) de 0,19 da el **exceso de riesgo** asociado con continuar fumando después de un ataque cardíaco. El

valor del riesgo de morir en el grupo de los que dejaron de fumar, 0,17, es el nivel de referencia o "esperado" con el cual se compara el riesgo de 0,36 en el grupo de los que continúan fumando.

Preguntas de estudio (P6.1)

1. ¿Cuántas muertes habrían ocurrido entre los 75 pacientes que continuaron fumando después de su ataque cardíaco si, por el contrario, estos 75 pacientes hubiesen dejado de fumar?
2. ¿Cuántas muertes en exceso hubo entre los 75 pacientes que continuaron fumando después de su ataque cardíaco?
3. ¿Cuál es la proporción de las muertes en exceso entre los que continúan fumando?

El valor nulo (en este caso, cero) describe "ningún exceso de riesgo". No habría ningún exceso de riesgo si los dos riesgos calculados fuesen iguales. Debido a que la diferencia de riesgos describe el **exceso** de riesgo, también se llama **riesgo atribuible**, y calcula el riesgo adicional "atribuible" a la exposición.



La diferencia de riesgos, por lo tanto, puede interpretarse como la probabilidad de que una persona expuesta desarrolle la enfermedad debido a la influencia adicional de exposición sobre el riesgo de referencia. En este ejemplo, el riesgo atribuible en cinco años de 0,19 calcula la probabilidad de que los que continúan fumando mueran debido a que han continuado haciéndolo.

Preguntas de estudio (P6.1) (continuación)

4. Si el estudio incluyera a 1.000 pacientes que sufrieron un ataque cardíaco y que continuaron fumando después de dicho ataque, ¿cuántas muertes podrían evitarse (es decir, podrían ser atribuibles a la exposición) para obtener una diferencia de riesgos de 0,19 si todos los pacientes dejaran de fumar?
5. ¿Cómo podría evaluar si el exceso de riesgo de 0,19 es clínicamente (no estadísticamente) excesivo?
6. ¿Se le ocurre un valor de referencia para compararlo con el exceso de riesgo? De ser así, ¿cómo interpretaría esta comparación relativa?

Resumen

- La diferencia de riesgos es la diferencia entre dos valores de riesgo.
- El valor nulo de la diferencia de riesgos es 0, mientras que el valor nulo de la razón de riesgos es 1.
- La diferencia de riesgos refleja un **exceso de riesgo** atribuible a la exposición.
- El exceso de riesgo describe la proporción de casos que podrían evitarse entre los individuos expuestos si éstos tuviesen el mismo riesgo que los individuos no expuestos.
- La diferencia de riesgos también se llama **riesgo atribuible**.

Motivación matemática del exceso de riesgo

El concepto de **exceso de riesgo** puede explicarse matemáticamente de la siguiente manera: considere la siguiente tabla de 2x2 que describe los datos de un estudio de cohorte que le permiten a usted calcular el riesgo individual mediante la incidencia acumulada:

	Expuestos	No expuestos	Total
Casos	a	b	m_1
No casos	c	d	m_0
Total	n_1	n_0	n

De este cuadro, los riesgos calculados para los expuestos, los no expuestos y los grupos totales son dados por las siguientes fórmulas de incidencia acumulada:

$$\hat{I}A_E = \frac{a}{n_1}, \hat{I}A_{no\,E} = \frac{b}{n_0}, \hat{I}A_{Total} = \frac{m_1}{n} = \frac{(a+b)}{n}$$

De las definiciones anteriores, podemos calcular el número en exceso de casos atribuibles a la exposición, que denotamos como **N(AE)**, de la siguiente manera:

N(AE) = número de casos nuevos entre los expuestos – número de casos nuevos esperados si hay ausencia de exposición

$= n_1 \text{ x } IA_E - n_1 \text{ x } IA_{no\,E}$

$= n_1 \text{ x } [IA_E - IA_{no\,E}]$

$= n_1 \text{ x}$ diferencia de riesgos

Dividiendo por n_1 ambos lados de la ecuación final, surge que:

Diferencia de riesgos = $\frac{N(AE)}{n_1}$ = exceso de riesgo entre los sujetos expuestos

Una forma alternativa para determinar N(AE) considerando todos los casos en lugar de los casos expuestos es la siguiente:

N(AE) = número de casos totales – número de casos totales esperados si hay ausencia de exposición

$= n \text{ x } IA_{Total} - n \text{ x } IA_{no\,E}$

$= n\left[\frac{[n_1 \cdot IA_E \cdot IA_{no\,E}]}{n}\right] - n \cdot IA_{no\,E}$

$= [n_1 \text{ x } IA_E + n_0 \text{ x } IA_{no\,E}] - (n_1 + n_0) \text{ x } IA_{no\,E}$

$= [n_1 \text{ x } IA_E + n_0 \text{ x } IA_{no\,E}] - [n_1 \text{ x } IA_{no\,E} + n_0 \text{ x } IA_{no\,E}]$

$= n_1 \text{ x } [IA_E - IA_{no\,E}]$

$= n_1 \text{ x}$ diferencia de riesgos

Medidas de efecto en forma de diferencias

Las medidas de efecto en forma de diferencias pueden calcularse en ensayos clínicos aleatorizados, en estudios de cohorte y transversales, pero *no* en estudios de casos y controles. En los estudios de cohorte que calculan el riesgo individual mediante las medidas de incidencia acumulada, la medida de diferencia de interés se llama **diferencia de riesgos**. Se calcula como la diferencia entre $\hat{I}A_1$, la incidencia acumulada en el grupo expuesto, e $\hat{I}A_0$, la incidencia acumulada en el grupo no expuesto.

En los estudios de cohorte que calculan la tasa promedio mediante la información de persona-tiempo, la medida de diferencia es la **diferencia de tasas**. Puede calcularse como la diferencia entre dos tasas calculadas o densidades de incidencia, $\hat{D}I_1$ y $\hat{D}I_0$.

En los estudios transversales, la medida de diferencia se llama **diferencia de prevalencias** y se calcula como la diferencia entre dos cálculos de prevalencia estimada.

Estudios de Cohorte

Riesgo individual $(\hat{IA})$:

Diferencia de riesgo $(\hat{DIA}) = \hat{IA}_1 - \hat{IA}_0$

Tasa promedio $(\hat{DI})$:

Diferencia de tasas $(\hat{DDI}) = \hat{DI}_1 - \hat{DI}_0$

Estudios transversales

Prevalencia $(\hat{P})$:

Diferencia de prevalencia $(\hat{DP}) = \underset{\text{expuesto}}{\hat{P}_1} - \underset{\text{no expuesto}}{\hat{P}_0}$

Las medidas de efecto en forma de diferencias no pueden calcularse en los estudios de casos y controles porque en dichos estudios no pueden calcularse adecuadamente ni el riesgo, ni la tasa, ni la prevalencia.

Ejemplificaremos el cálculo de la diferencia de tasas. Nuevamente consideremos los datos sobre la relación entre el nivel de colesterol en sangre y la mortalidad de un estudio de casi 40.000 personas de Chicago, Illinois realizado en 1992. Entre los hombres blancos de 25 a 39 años con colesterol moderadamente alto, hubo 71,1 muertes por 100.000 personas-años. Entre el grupo de comparación, hubo 20,5 muertes por 100.000 personas-años.

1992, 40,000 personas, Chicago (hombres blancos, edad 25 - 39)

	Nivel de colesterol	
	Moderadamente alto	Normal
Muertes	26	14
Personas-Años	36,581	68,239

$\hat{DI}_1 = \frac{71.1}{100{,}000 \text{ PA}}$ $\hat{DI}_0 = \frac{20.5}{100{,}000 \text{ PA}}$

La razón de tasas que compara estos dos grupos es 3,5. La diferencia de tasas, o DDI, es 50,6 muertes por 100.000 personas-años. ¿Qué tipo de interpretación podemos hacer de esta diferencia de tasas?

$$\hat{RDI} = \frac{\frac{71.1}{100{,}000\ PA}}{\frac{20.5}{100{,}000\ PA}} = 3.5 \qquad \hat{DDI} = \frac{71.1}{100{,}000\ PA} - \frac{20.5}{100{,}000\ PA} = \frac{50.6}{100{,}000\ PA}$$

La diferencia de tasas indica un exceso de tasa de 50,6 muertes por 100.000 personas-años asociado a tener colesterol moderadamente alto, en comparación con el nivel normal de colesterol. Aquí estamos utilizando la tasa estimada de muertes relacionadas con la enfermedad coronaria (EC) en el grupo no expuesto como el nivel de *referencia* o *esperado* con el cual se compara la tasa en el grupo expuesto. La diferencia de tasas también se llama **tasa atribuible**, ya que brinda la tasa adicional atribuible a la exposición.

Preguntas de estudio (P6.2)

1. ¿Cuántas muertes relacionadas con la EC por 100.000 personas-años (es decir, PA) podrían evitarse (es decir, podrían ser atribuibles a la exposición) entre las personas con colesterol moderadamente alto si estas personas pudiesen disminuir su nivel de colesterol a los valores normales?
2. ¿Cuál es el exceso de tasa de muertes relacionadas con la EC por 100.000 PA entre las personas con colesterol moderadamente alto?
3. ¿Cómo podría evaluar si el exceso de tasa de 50,6 es clínicamente (no estadísticamente) excesivo?
4. ¿Se le ocurre un valor de referencia para compararlo con el exceso de tasa? De ser así, ¿cómo interpretaría esta comparación relativa?

Resumen

- Las medidas de diferencia pueden calcularse en estudios de cohorte y transversales, pero no en estudios de casos y controles.
- Si se calcula el **riesgo**, la medida de diferencia es la **diferencia de riesgos**.
- Si se calcula la **tasa**, la medida de diferencia es la **diferencia de tasas**.
- Si se calcula la **prevalencia**, la medida de diferencia es la **diferencia de prevalencias**.
- Las medidas de efecto en forma de diferencias le permiten calcular el (exceso de) riesgo atribuible a la exposición sobre el riesgo de referencia brindado por los individuos no expuestos.

El número necesario de pacientes a tratar

La diferencia de riesgos describe el exceso de riesgo de la enfermedad que es atribuible a la exposición. La diferencia de riesgos también se utiliza para calcular el "número necesario de pacientes a tratar" o NNT. El NNT representa el número de pacientes que deben tratarse para evitar que ocurra un resultado.

Como ejemplo, utilizamos los datos del Consejo Británico de Investigación Médica en un estudio de pacientes con hipertensión leve. Se comparó el tratamiento de la hipertensión con un antagonista beta-adrenérgico o un diurético con el uso de un placebo. La incidencia acumulada en 10 años con respecto al accidente cerebrovascular (ACV) en pacientes que recibieron el placebo fue 2,6%; en los pacientes que fueron tratados, la incidencia acumulada en 10 años fue 1,4%.

Tratamiento:

Placebo vs. antagonista beta-adrenérgico diurético
Incidencia en 10 años del ACV:
$\hat{I}A_1 = 2{,}6\%$ $\hat{I}A_0 = 1{,}4\%$

Preguntas de estudio (P6.3)

1. ¿Cuál es la diferencia de riesgos?

Ahora queremos saber cuántos pacientes con hipertensión necesitamos tratar para evitar un ACV. La fórmula utilizada para calcular el NNT es 1 dividido por la diferencia de riesgos.

$$\text{NNT} = \frac{1}{\hat{\text{IA}}_1 - \hat{\text{IA}}_0}$$

Preguntas de estudio (P6.3) (continuación)

2. ¿Cuántos pacientes con hipertensión necesitamos tratar para evitar un ACV?
3. ¿Durante cuánto tiempo necesitamos tratar a estos pacientes?

Los resultados de este estudio mostraron que si 83 individuos con hipertensión se tratasen con antagonistas beta-adrenérgicos o diuréticos durante 10 años podría evitarse la incidencia de un ACV.

Resumen

- La diferencia de riesgos puede utilizarse para calcular el "número de pacientes necesario a tratar" (NNT).
- El NNT representa el número de pacientes que deben tratarse para evitar que ocurra un resultado.
- La fórmula utilizada para calcular el NNT es 1 dividido por la diferencia de riesgos.

El número necesario de pacientes a tratar: definición y base lógica

En el ejemplo de esta actividad, se determinó la incidencia acumulada en 10 años de ACV para pacientes con hipertensión leve. Para los que se trataron con antagonista beta-adrenérgico o diurético, la IA calculada fue 1,4%. Para los que recibieron un placebo (es decir, quienes no fueron tratados), la IA calculada fue 2,6%. La diferencia de riesgos se calcula entonces como 2,6% - 1,4% = 1,2%. En otras palabras, el riesgo de ACV atribuible al uso de un placebo (o no tratamiento) fue 1,2%.

El **número necesario de pacientes a tratar (NNT)** se define como el número esperado de pacientes que deben tratarse con una terapia experimental para poder evitar un resultado adverso adicional (o, según el contexto, esperar un resultado beneficioso adicional). Podemos calcular el NNT invirtiendo el valor de la diferencia de riesgos, es decir:

NNT = l/(Diferencia de riesgos).

Para este ejemplo, por lo tanto, NNT = 1/0,012, que da 83,3. En otras palabras, 83 pacientes deben tratarse

con antagonista beta-adrenérgico o diurético para evitar un ACV.

Según el ejemplo anterior, la base lógica de la fórmula de NNT es la siguiente: si 1.000 pacientes fueran tratados con antagonista beta-adrenérgico o diurético, esperaríamos 14 ACV (1,4%) durante 10 años de seguimiento. Sin embargo, si estos 1.000 pacientes recibieran un placebo en lugar de tratamiento, esperaríamos 26 ACV (2,6%) durante 10 años.

Por lo tanto, para 1.000 pacientes seguidos durante 10 años, esperaríamos evitar 26 - 14 = 12 ACV (2,6% - 1,4% = 1,2%) si los 1.000 pacientes recibieran el tratamiento, en lugar de que ninguno de ellos recibiera tratamiento (es decir, recibiera un placebo). Ya que 1.000/12 = 83,3/1, esto significa que podríamos esperar prevenir un ACV durante 10 años por cada 83 pacientes que son tratados; es decir, 12 es a 1.000 como 1 es a 83,3.

Auto-evaluación (P6.4)

¿Cuáles de los siguientes términos son sinónimos de diferencia de riesgos?

1. Riesgo absoluto **???**
2. Riesgo atribuible **???**
3. Exceso de riesgo **???**
4. Razón de riesgos **???**

Opciones
No Sí

Durante el brote de encefalitis del Nilo Occidental en Nueva York en 1999, la incidencia varió según la ubicación. Las tasas informadas fueron:

Queens	16,4 por millón	Bronx	7,5 por millón
Brooklyn	1,3 por millón	Manhattan	0,7 por millón
Staten Island	0,0 por millón	Total ciudad de Nueva York	6,1 por millón

Para calcular la diferencia de tasas para los residentes de Queens, ¿qué ubicación(es) podría(n) utilizarse para la tasa de referencia o esperada?

5. Queens **???**
6. Bronx **???**
7. Brooklyn **???**
8. Manhattan **???**
9. Staten Island **???**
10. Total ciudad de Nueva York **???**
11. Bronx + Brooklyn + Manhattan + Staten Island **???**

Opciones
No Sí

12. Calcule la diferencia de tasas entre Queens y Manhattan. **???**

Opciones
15,7 15,7 por millón 23,4 23,4 por millón

Los investigadores entrevistaron a todas las personas que habían asistido a la boda de Smith y Jones dos días antes, comparando la proporción que desarrolló gastroenteritis entre los que comieron y los que no comieron ciertos alimentos. Ahora quieren determinar el impacto de comer ensalada de papa en la gastroenteritis.

13. La medida adecuada del impacto potencial es **???**

Los investigadores realizaron una encuesta transversal, identificaron a los encuestados a los que se les había diagnosticado diabetes y calcularon un índice de obesidad mediante las alturas y los pesos informados. Ahora quieren determinar el impacto de la obesidad en la diabetes.

14. La medida adecuada del impacto potencial es **???**

Los investigadores incluyeron en un estudio de seguimiento a estudiantes de primer año recién matriculados en la universidad. Los investigadores les administraron cuestionarios y les tomaron una muestra de sangre anual para identificar factores de riesgo de adquisición y seroconversión al virus Epstein-Barr (VEB, el agente etiológico de la mononucleosis). Utilizando el número de personas-años de observación, ahora los investigadores quieren determinar el impacto de residir en un dormitorio mixto sobre la seroconversión al VEB.

15. La medida adecuada de impacto potencial es **???**

Opciones

no calculable diferencia de *odds* diferencia de prevalencias diferencia de tasas diferencia de riesgos

Medidas de efecto en forma de diferencias versus tasas

Considere este cuadro de información hipotética que describe las relaciones separadas de cuatro exposiciones distintas a la misma enfermedad.

UBICACIÓN	TABACO	CAFÉ	ALCOHOL
$IA_1 = 0.010$	$IA_1 = 0.005$	$IA_1 = 0.050$	$IA_1 = 0.050$
$IA_0 = 0.010$	$IA_0 = 0.001$	$IA_0 = 0.046$	$IA_0 = 0.010$
RR = 1.000	RR = 5.000	RR = 1.087	RR = 5.000
DR = 0.000	DR = 0.004	DR= 0.004	DR= 0.040

Primero concéntrese en la ubicación, rural versus urbana, para la cual la razón de riesgos es uno y la

diferencia de riesgos es 0. No hay evidencia de un efecto de la ubicación en la enfermedad, ya sea que se considere la razón de riesgos o la diferencia de riesgos. De hecho, si la razón de riesgos es exactamente 1, entonces la diferencia de riesgos debe ser exactamente 0, y viceversa.

Ubicación: rural vs. urbana
RR=1,000 DR = 0,000
Sin efecto

Ahora, veamos el efecto de mascar tabaco en la enfermedad. La razón de riesgos para mascar tabaco es 5; esto indica una relación muy fuerte entre mascar tabaco y la enfermedad. Pero, la diferencia de riesgos de 0,004 parece bastante cercana a cero, lo que sugiere que mascar tabaco no produce ningún efecto.

Mascar tabaco
RR = 5,000 DR = 0,004
Efecto fuerte Efecto leve

Por lo tanto, es posible llegar a una conclusión distinta si utilizamos la razón de riesgos o la diferencia de riesgos. ¿Sólo una de estas dos medidas de efecto brinda la conclusión correcta o ambas son correctas en cierto punto? En verdad, ambas medidas brindan información relevante sobre dos aspectos distintos de la relación entre exposición y enfermedad.

Comparemos ahora el efecto de mascar tabaco con el efecto de tomar café.

Tomar café
RR = 1,087 DR = 0,004

Las razones de riesgos para estas dos exposiciones son muy distintas, aunque las diferencias de riesgos son exactamente las mismas y cercanas a cero. Parece que el efecto de tomar café es muy leve, si es que lo hay. Por lo tanto, ¿existe o no un efecto de mascar tabaco?

Si nos preguntamos si consideraríamos que mascar tabaco es un factor de riesgo fuerte para la enfermedad, nuestra respuesta sería afirmativa, ya que una persona que masca tabaco tiene 5 veces más riesgo de sufrir la enfermedad que una persona que no lo hace. Es decir, mascar tabaco parece estar asociado a la etiología de la enfermedad, ya que es un factor de riesgo muy fuerte.

Sin embargo, si nos preguntamos si mascar tabaco es una carga para la salud pública al tener un gran número de casos de pacientes que requieren tratamiento, nuestra respuesta sería negativa. Para ver las consecuencias en la salud pública, recuerde que la diferencia de riesgos de 0,004 para mascar tabaco muestra el exceso de riesgo que resultaría si se eliminara totalmente el mascar tabaco en la población del estudio. Por lo tanto, por poner un ejemplo, de 1.000 personas que mascan tabaco, un exceso de 1.000 multiplicado por 0,004, o sólo 4 personas que mascan tabaco, desarrollarían la enfermedad debido a este hábito. No hay muchos pacientes que haya que tratar en relación con las 1.000 personas que mascan tabaco y que están en riesgo de padecer la enfermedad.

Preguntas de estudio (P6.5)

1. Compare el efecto de mascar tabaco con el efecto de consumir alcohol en la enfermedad. ¿Ambos hábitos tienen el mismo efecto en términos de la etiología de la enfermedad?
2. ¿Mascar tabaco y consumir alcohol tienen las mismas consecuencias en la salud pública con respecto al tratamiento de la enfermedad?
3. Explique su respuesta a la pregunta anterior en términos de la idea de exceso de riesgo.

Resumen

- Si la razón de riesgos es exactamente igual a 1, entonces la diferencia de riesgos es exactamente igual a 0, y viceversa, y no hay efecto de exposición en el resultado de salud.
- Si la razón de riesgos difiere mucho de 1, todavía es posible que la diferencia de riesgos sea cercana a cero.
- Si la diferencia de riesgos es cercana, pero no exactamente igual a 0, es posible que la diferencia de riesgos sea lo suficientemente grande como para indicar un problema en la salud pública para tratar la enfermedad.
- Las medidas en forma de razones se utilizan principalmente para aprender sobre la etiología de una enfermedad u otro resultado de salud.
- Las medidas en forma de diferencias se utilizan para determinar la importancia en la salud pública de una enfermedad u otro resultado de salud.

Auto-evaluación (P6.6)



Durante el brote de encefalitis del virus del Nilo Occidental (VNO) en la ciudad de Nueva York en 1999, las tasas informadas fueron:

Queens	16,4 por millón de habitantes
Resto de la ciudad de Nueva York	2,4 por millón
Total de la ciudad de Nueva York	6,1 por millón

Coloque **verdadero** o **falso** en cada una de las siguientes afirmaciones:

1. Si Queens hubiese tenido la misma tasa de VNO que el resto de la ciudad de Nueva York, se habrían presentado 10,3 casos menos por millón, es decir, la diferencia de tasas sería de 10,3 por millón. **???**

2. El exceso de tasa en Queens fue de 14,0 casos por millón (comparado con el resto de la ciudad de Nueva York). . . . . **???**

3. La tasa atribuible (es decir, la diferencia de tasas) en Queens fue de 16,4 casos por millón. . **???**

4. La medida de efecto más común para comparar a Queens con el resto de la ciudad de Nueva York es 6,8. . . . . **???**

Determine si cada una de las siguientes afirmaciones coincide más con la **diferencia de riesgos**, **razón de riesgos**, **ambas** o **ninguna**.

5. Una medida más de carga a la salud pública. . . . . **???**

6. Una medida más de etiología. . . . . **???**

7. El valor nulo es 0,0. . . . . **???**

8. Puede ser un número negativo. . . . . **???**
9. Puede ser un número entre 0,0 y 1,0. . . . . **???**
10. Puede calcularse a partir de la mayoría de los estudios de seguimiento. . . . **???**
11. Puede calcularse a partir de la mayoría de los estudios de casos y controles. . . **???**
12. No tiene unidades. . . . . **???**
13. Un valor muy cercano a 0,0 indica un efecto fuerte. . . . . **???**
14. Es sinónimo de riesgo atribuible. . . . . **???**

Considere los datos en el cuadro que figura a continuación, al igual que los siguientes cálculos de riesgo sobre el tabaquismo y la incidencia del cáncer de pulmón y la enfermedad coronaria (EC).

	Cáncer de pulmón	**EC**
Razón de tasas	12,9	2,1
Diferencia de tasas	79,0/100k/año	190,4/100k/año

15. ¿Qué enfermedad está más fuertemente asociada al tabaquismo? . **???**
16. ¿La eliminación del tabaquismo reduciría la mayoría de los casos de qué enfermedad? **???**

Incidencia del cáncer de pulmón

	Fumadores	**No fumadores**	**Total**
Nuevos casos de cáncer de pulmón	60.000	10.000	70.000
Cantidad estimada de personas-años	70.000.000	150.000.000	220.000.000
Densidad de incidencia	$\hat{D}I_1$=85,7 por 100.000 personas-años	$\hat{D}I_0$=6,7 por 100.000 personas-años	$\hat{D}I$=31,8 por 100.000 personas-años

Incidencia de la enfermedad coronaria (EC)

	Fumadores	**No fumadores**	**Total**
Casos nuevos de EC	250.000	250.000	500.000
Cantidad estimada de personas-años	70.000.000	150.000.000	220.000.000
Densidad de incidencia	$\hat{D}I_1$=357,1 por 100.000 personas-años	$\hat{D}I_0$=166,7 por 100.000 personas-años	$\hat{D}I$=227,3 por 100.000 personas-años

6-2 Medidas de impacto potencial (continuación)

Impacto potencial: concepto

La medida de impacto potencial brinda una perspectiva de la salud pública sobre una relación exposición-enfermedad que se está estudiando. De manera más específica, la medida de impacto potencial trata de responder a la pregunta ¿en cuánto se reduciría la carga de la enfermedad en una población específica si se cambiara la distribución de una variable de exposición? Por carga de la enfermedad nos referimos a las personas con una enfermedad de interés que requerirían atención médica en un punto específico del tiempo.

La típica medida de impacto potencial es una proporción, generalmente expresada como porcentaje, del número de casos que no se hubiesen convertido en tales si todas las personas bajo estudio tuviesen la misma condición de exposición. Por ejemplo, cuando se determina el impacto potencial de fumar sobre el desarrollo de cáncer de pulmón, la medida de impacto potencial de fumar da la proporción de nuevos casos de cáncer de pulmón que no se hubiesen desarrollado si nadie fumara en la población.

O bien, uno podría determinar el impacto potencial de una vacuna hipotética en la prevención de una enfermedad, por ejemplo, el VIH, en personas de alto riesgo. La medida de impacto potencial de la vacuna brinda la proporción de todos los casos potenciales de VIH evitados por la vacuna; si no hubiese existido la vacuna, todos estos casos se habrían presentado.

Estos ejemplos ilustran dos tipos de medidas de impacto potencial. La medida de impacto de fumar en el cáncer de pulmón considera una exposición que está asociada a un mayor riesgo de la enfermedad y se llama **fracción etiológica**. El ejemplo de la medida de impacto de una vacuna hipotética para evitar el VIH considera una exposición que está asociada a un menor riesgo de la enfermedad y se denomina **fracción prevenida**.

Resumen

- La medida de impacto potencial brinda una perspectiva de la salud pública sobre el efecto de una relación exposición-enfermedad.
- En general, las medidas de impacto potencial indican qué proporción de casos desarrollaron la enfermedad como resultado de la supuesta influencia de la exposición.
- La fracción etiológica es una medida de impacto potencial que considera una exposición que es una causa potencial de la enfermedad.
- La fracción prevenida es una medida de impacto potencial que considera una exposición que previene la enfermedad.

Medidas generalizadas del impacto potencial

Esta lección sobre las medidas de impacto potencial se centra en la cuantificación de la reducción potencial de la incidencia de la enfermedad a partir de un cambio específico en la prevalencia de una única variable de exposición binaria. La **fracción etiológica** (**FE**) considera el potencial de beneficios futuros (es decir, el impacto potencial) que surge de eliminar totalmente la presencia de una exposición nociva, por ejemplo, fumar. La **fracción prevenida** (**FP**) considera los beneficios pasados que surgen de introducir una exposición protectora (eliminando completamente su ausencia), por ejemplo, una vacuna o programa de ejercicios. Estas dos medidas, sin embargo, están limitadas de muchas maneras, lo que ha conducido a medidas más generalizadas de impacto potencial. La discusión detallada de dichas medidas generalizadas está fuera del alcance de esta presentación. No obstante, a continuación describimos brevemente varias generalizaciones importantes.

1. **Exposiciones multiniveles**. La fórmula de la FE puede extenderse de la siguiente manera, si los cálculos del riesgo están disponibles para k+ 1 categorías de exposición:

$$FE = 1 - \left[\frac{1}{\sum_{i=0}^{k} (p_i \cdot RR_i)} \right]$$

donde p_i denota la proporción de la población en el grupo de exposición i, donde i aumenta de 0 a k, y RR_i es la razón de riesgos que comparó la categoría de exposición i-ava con el grupo referente 0. Si los datos involucran tasas en lugar de riesgos, sustituya RDI_i por RR_i en la fórmula anterior.

2. **Ajuste por otros factores**. Tres enfoques:
   a) medidas separadas para subgrupos:
   b) análisis estratificado:
      i) uso de promedio ponderado de FE para cada subgrupo, o
      ii) uso de valor ajustado de RR;
   c) regresión logística.

3. **Impacto de la modificación parcial de la exposición (en lugar de su completa eliminación)**. Ejemplos:
   a) medir la reducción de la mortalidad por cáncer de pulmón si se reduce, pero no se elimina, el hábito de fumar;
   b) medir la reducción de la enfermedad coronaria por mayores niveles de actividad física.

A continuación se brinda una fórmula para la fracción de impacto (**FI**) generalizada:

$$FI = \frac{\sum_{i=1}^{k} \left[(p_i^* - p_i^{**}) \cdot RR_i \right]}{\sum_{i=0}^{k} \left[p_i^* \cdot RR_i \right]}$$ donde

FI es la reducción del riesgo de la enfermedad como resultado de un cambio en la distribución de una variable de exposición multinivel;
$\mathbf{p_i^*}$ es la proporción de la población candidata en la categoría de exposición i-ava antes de la intervención o cambio planeados; y
$\mathbf{p_i^{**}}$ es la proporción de la población candidata en la categoría de exposición i-ava después del cambio.

Fracción etiológica

La **fracción etiológica** responde a la pregunta ¿qué proporción de casos nuevos que ocurren durante un cierto período de seguimiento son atribuibles a la exposición de interés? Otros nombres para esta medida son **fracción etiológica *de la población***, **fracción atribuible *de la población***, **riesgo atribuible de la población** y **porcentaje de riesgo atribuible de la población**.

En términos matemáticos, la fracción etiológica está dada por la fórmula **I*** dividido por **I**, donde **I*** denota el número de casos nuevos atribuibles a la exposición e **I** denota el número de casos nuevos que realmente ocurren. El numerador, **I***, puede encontrarse como la diferencia entre el número real de casos nuevos y el número de casos nuevos que se hubiesen presentado en ausencia de la exposición.

$$FE = \frac{I^*}{I}$$

Para ejemplificar el cálculo de la fracción etiológica, considere nuevamente los resultados de un estudio de seguimiento durante cinco años para determinar si los fumadores que han tenido un ataque cardíaco reducirán su riesgo de morir si dejan de fumar. Aquí, la razón de riesgos es 2,1 y la diferencia de riesgos es 0,19.

Pacientes ataque cardíaco	Fuma	Dejó	Total
Mortalidad	27	14	41
Supervivencia	48	67	115
Total	75	81	156

Riesgo			
continúan fumando	0.36	Razón de riesgos	2.1
dejaron de fumar	0.17	Diferencia de riesgo	.19

$$\hat{FE} = \frac{\hat{IA} - \hat{IA}_0}{\hat{IA}} = \frac{0.263 - 0.17}{0.263} \approx 0.35$$

Aquí se muestra una fórmula para el cálculo de la fracción etiológica, donde $\hat{IA}$ denota el valor de la incidencia acumulada o riesgo para todos los individuos, expuestos y no expuestos en forma combinada, en el estudio. $\hat{IA}_0$ denota el valor de la incidencia acumulada para los individuos no expuestos. Observe que el numerador en esta fórmula ***no*** es la diferencia de riesgos, que involucraría $\hat{IA}_1$, el valor del riesgo para las personas expuestas, en lugar de $\hat{IA}$, el valor del riesgo general.

Entonces, para calcular la fracción etiológica mediante nuestros datos, primero debemos calcular $\hat{IA}$, que es igual a 0,263 ó aproximadamente 26%. Ya sabemos que $\hat{IA}_0$ es 0,173 ó aproximadamente 17%. Si sustituimos estos valores en la fórmula, encontramos que la fracción etiológica es 0,35 ó 35%.

¿Cómo interpretamos este resultado? La fracción etiológica de 0,35 nos dice que el 35% de todos los casos que realmente ocurrieron se debieron a la continuación del hábito de fumar. En otras palabras, si pudiésemos haber hecho que todos los pacientes dejaran de fumar después de haber sufrido un ataque cardíaco, habría habido una reducción de 35% en el número total de muertes. Ésta es la razón por la cual la fracción etiológica se llama a menudo porcentaje de riesgo atribuible de la población. Brinda el porcentaje de todos los casos de la población que son atribuibles a la exposición, en el sentido de aportar exceso de riesgo.

Preguntas de estudio (P6.7)

Si nos basamos en el ejemplo de los fumadores de la página anterior:

1. ¿Cuántos casos se habrían esperado si todos los individuos no hubiesen estado expuestos?
2. ¿Cuál es el número en exceso de casos totales esperados en ausencia de exposición?
3. ¿Cuánto es I*/I para estos datos?

Resumen

- La fracción etiológica está dada por la fórmula I*/I, donde I denota el número de casos nuevos que realmente ocurren e I* denota el número de casos nuevos atribuibles a la exposición.
- El numerador, I*, puede cuantificarse como la diferencia entre el número real de casos nuevos y el número de casos nuevos que se hubiesen presentado en ausencia de exposición.
- Una fórmula para calcular la fracción etiológica es FE= (IA – IA_0) / IA, donde IA es la incidencia acumulada.
- A menudo FE se llama porcentaje de riesgo atribuible de la población, porque brinda el porcentaje de todos los casos en la población que son atribuibles a la exposición.

Fórmula alternativa de la fracción etiológica

En un estudio de cohorte que calcula el riesgo, la fracción etiológica puede calcularse a partir de los valores de la incidencia acumulada para toda la cohorte y para las personas no expuestas. Una fórmula equivalente puede escribirse en términos de la razón de riesgos y la proporción (p) de personas expuestas en la cohorte.

Fórmula alternativa/ fracción etiológica

$$\hat{FE} = \frac{\hat{IA} - \hat{IA}_0}{\hat{IA}} = \frac{p\,(\hat{RR} - 1)}{p(\hat{RR} - 1) + 1}$$

p = proporción expuestas

Para ejemplificar esta fórmula alternativa, considere nuevamente los resultados de un estudio de seguimiento durante cinco años para determinar si los fumadores que han tenido un ataque cardíaco reducirán su riesgo de morir dejando de fumar. Mediante la primera fórmula, calculamos previamente la fracción etiológica, que es 0,35 ó 35%. Por lo tanto, el 35% de todos los casos que realmente ocurrieron se deben a aquellos que continuaron fumando.

Estudio de seguimiento durante 5 años

	Fuma	Dejó	Total
Murió	27	14	41
Sobrevivió	48	67	115
Total	75	81	156

$\hat{RR} = 2.1$

$\hat{DR} = .19$

Para utilizar la segunda fórmula, primero calculamos la proporción de la cohorte expuesta, que es 0,481.

$p = 75/156 = .481$

Ahora sustituimos este valor y la razón de riesgos de 2,1 en la segunda fórmula. El resultado es 0,35, que es exactamente igual al obtenido anteriormente, ya que ambas fórmulas son equivalentes.

$$\hat{FE} = \frac{.481(2.1 - 1)}{.481(2.1 - 1) + 1} = 0.35$$

La segunda fórmula nos brinda información adicional sobre el significado de la fracción etiológica. Esta fórmula nos dice que el tamaño de la fracción etiológica depende del tamaño de la razón de riesgos y la proporción expuesta. En particular, el impacto potencial para un determinante fuerte de la enfermedad (es decir, cuando la razón de riesgos es alta) puede ser pequeño si están expuestas relativamente pocas personas de la población.

Suponga en nuestro ejemplo que sólo el 10%, en lugar del 48%, de la cohorte estuviese expuesta, de forma que *p* es igual a 0,10. Entonces, la fracción etiológica se reduciría a 0,10 ó 10%, lo que indica un impacto de la exposición mucho más pequeño que el 35%. Además, si toda la cohorte no estuviese expuesta, la fracción etiológica sería cero.

$$\hat{FE} = \frac{0.10(2.1 - 1)}{0.10(2.1 - 1) + 1} = 0.10 < 0.35$$

Ahora suponga que el 90%, en lugar del 48%, de la cohorte estuviese expuesta, de forma que p es igual a 0,90. Entonces, la fracción etiológica aumenta a 0,50 ó 50%. Si toda la cohorte estuviese expuesta, la fracción etiológica aumentaría a su valor máximo posible de 0,52 ó 52% para una razón de riesgos de 2,1.

$$\hat{FE} = \frac{0.90(2.1 - 1)}{0.90(2.1 - 1) + 1} = 0.50$$

En general, para un valor fijo de la razón de riesgos, la fracción etiológica puede oscilar entre cero, si toda la cohorte no estuviese expuesta, y un valor máximo de RR menos uno sobre RR, si toda la cohorte estuviese expuesta.

Valor fijo de $\hat{RR}$:

$$0 \leq \hat{FE} \leq \frac{\hat{RR} - 1}{\hat{RR}}$$

$p = 0$ (toda cohorte no expuesta) $\qquad$ $p = 1$ (toda cohorte expuesta)

Preguntas de estudio (P6.8)

1. Utilice la fórmula (RR – 1)/RR para calcular el valor máximo posible de FE cuando la RR es 2,1.
2. ¿Cuál es el valor máximo posible de FE cuando RR es 10?
3. A medida que RR aumenta hacia el infinito, ¿a qué se acerca el valor máximo posible de FE?
4. Si la RR es muy grande, por ejemplo, 100, ¿FE puede ser todavía relativamente pequeña? Desarrolle su respuesta.

Resumen

- Una fórmula alternativa de la fracción etiológica es FE = p(RR-1) / [p(RR-1)+1], donde RR es la razón de riesgos y p es la proporción expuesta en toda la cohorte.
- El tamaño de la fracción etiológica depende del tamaño de la razón de riesgos y la proporción expuesta.
- Para un valor fijo de la razón de riesgos, la fracción etiológica puede oscilar entre cero y un valor máximo de (RR – 1)/RR.
- El impacto potencial para un determinante fuerte de la enfermedad (es decir, una razón de riesgos alta) puede ser pequeño si relativamente pocas personas de la población están expuestas.

Fracción etiológica para estudios de cohorte con persona-tiempo

En los estudios de cohorte que calculan una tasa utilizando información de persona-tiempo, la **fracción etiológica** puede calcularse utilizando valores de **densidad de incidencia**, en lugar de la incidencia acumulada.

$$\hat{\text{FE}} = \frac{\hat{\text{DI}} - \hat{\text{DI}}_0}{\hat{\text{DI}}}$$

En la fórmula anterior, $\hat{\text{DI}}$ denota la densidad de incidencia o tasa calculada para todos los individuos, expuestos y no expuestos en forma combinada. $\hat{\text{DI}}_0$ denota la densidad de incidencia calculada para los individuos no expuestos.

Una versión equivalente de esta fórmula puede escribirse en términos del valor de la razón de densidades de incidencia, $\hat{\text{RDI}}$, y la proporción p*, de persona-tiempo total para las personas expuestas.

$$\hat{\text{FE}} = \frac{(\text{p}^*)(\hat{\text{RDI}} - 1)}{(\text{p}^*)(\hat{\text{RDI}} - 1) + 1}$$

En esta fórmula, $\hat{\text{RDI}}$ es la razón de las tasas para los grupos expuestos y no expuestos; y p* se calcula como L_1 dividido por ($L_1 + L_0$), donde L_1 y L_0 es la información de persona-tiempo para los grupos expuestos y no expuestos, respectivamente.

$$\hat{\text{RDI}} = \frac{\frac{\text{I}_1}{\text{PT}_1}}{\frac{\text{I}_0}{\text{PT}_0}}, \quad \text{p}^* = \frac{\text{L}_1}{\text{L}_1 + \text{L}_0}$$

Para ejemplificar el cálculo de cada fórmula, consideramos de nuevo los datos sobre el nivel de colesterol en sangre y la mortalidad. Para calcular la fracción etiológica utilizando la fórmula con densidades de incidencia, primero debemos calcular el valor de la densidad de incidencia general, que es 38,2 por 100.000 personas-años

([26+14]/[36.581+68.239]). Sustituyendo $\hat{DI}$ y $\hat{DI}_0$ en la fórmula por 38,2 y 20,5 respectivamente, obtenemos una fracción etiológica de 0,463 ó 46,3%.

	Nivel de colesterol	
	Moderadamente alto	Normal
Muertes	26	14
Personas-Años	36,581	68,239

$$\hat{DI}_1 = \frac{26}{36,581} = \frac{71.1}{100,000} \qquad \hat{RDI} = 71.1 / 20.5 = 3.47$$

$$\hat{DI}_0 = \frac{14}{68,239} = \frac{20.5}{100,000} \qquad \hat{DDI} = \frac{71.1 - 20.5}{100,000} = \frac{50.6}{100,000}$$

$$\hat{FE} = \frac{\hat{DI} - \hat{DI}_0}{\hat{DI}} = \frac{38,2 - 20,5}{38,2} = 46,3\%$$

Utilizando la segunda fórmula, descubrimos que p* es 0,349 y, ya que la razón de tasas es 3,47 (36.581/[36.581+68.239]), la fracción etiológica basada en esta fórmula también es del 46,3%:

$$\hat{FE} = \frac{(p^*)(R\hat{D}I - 1)}{(p^*)(R\hat{D}I - 1) + 1} = \frac{0,349(3,47 - 1)}{0,349(3,47 - 1) + 1} = 46,3\%$$

¿Cómo interpretamos este resultado?

Preguntas de estudio (P6.9)

1. ¿Cuál de las siguientes afirmaciones **no** es correcta sobre la FE (46,3%) calculada en el ejemplo anterior?
   A. Casi la mitad de todas las muertes se deben a personas con colesterol moderadamente alto.
   B. La proporción de muertes en exceso a raíz de la exposición (tener colesterol moderadamente alto) del total de muertes es 0,463.
   C. La reducción porcentual del total de muertes si las personas con colesterol moderadamente alto pudiesen disminuirlo a niveles normales es 46,3.
   D. 46% de todas las muertes entre las personas con colesterol moderadamente alto se deben a su nivel de colesterol.

2. En el cuadro siguiente, ¿cuál es el número de muertes en exceso del total de muertes si todas las personas en la cohorte tuviesen niveles normales de colesterol?
   A. 18,5
   B. 40,0
   C. 21,5
   D. 14,0

	Nivel de colesterol	
	Moderadamente alto	Normal
Muertes	26	14
Personas-Años	36,581	68,239

$\hat{DI}_1 = 26 / 36{,}581 = 71.1/100{,}000$ $\hat{RDI} = 71.1 / 20.5 = 3.47$

$\hat{DI}_0 = 14 / 68{,}239 = 20.5/100{,}000$ $\hat{DDI} = \frac{71.1 - 20.5}{100{,}000} = \frac{50.6}{100{,}000}$

Resumen

- En los estudios de cohorte que utilizan información de persona-tiempo, la fracción etiológica puede definirse en términos de densidades de incidencia o como una combinación de la razón de densidades de incidencia y la proporción de persona-tiempo total para las personas expuestas.
- La fórmula que involucra densidades de incidencia es:

$$\hat{FE} = \frac{\hat{DI} - \hat{DI}_0}{\hat{DI}}$$ donde $\hat{DI}$=densidad de incidencia total y $\hat{DI}_0$=densidad de incidencia entre los no expuestos

- Una fórmula equivalente es: $\hat{FE} = \frac{(p^*)(\hat{RDI} - 1)}{(p^*)(\hat{RDI} - 1) + 1}$ donde RDI es la razón de densidades de incidencia

$p^* = \frac{L_1}{L_1 + L_0}$ donde L_1 y L_0 es la información de persona-tiempo para los grupos expuestos y no expuestos, respectivamente.

Fracción etiológica entre los expuestos

Hay dos formulaciones conceptuales de la fracción etiológica. Una, que ya hemos descrito anteriormente, se centra en el impacto potencial de la exposición sobre el número total de casos, que se muestra a continuación como **I**. La segunda se enfoca en el impacto potencial de la exposición sobre el número de casos expuestos, que denotamos como $\mathbf{I}_1$. Esta medida se llama **fracción etiológica *entre los expuestos,* fracción atribuible *entre los expuestos*** o **porcentaje de riesgo atribuible *entre los expuestos***. En términos matemáticos, la fracción etiológica entre los expuestos está dada por la fórmula **I*** dividido por $\mathbf{I_1}$, donde **I*** denota el número de casos expuestos atribuibles a la exposición e $\mathbf{I_1}$ denota el número de casos expuestos que realmente ocurren.

1. Impacto potencial sobre la cantidad total de casos (I)

$$FE = \frac{I^*}{I} = \frac{\hat{CI} - \hat{CI}_0}{\hat{CI}} = \frac{p(\hat{RR} - 1)}{[p(\hat{RR} - 1) + 1]}$$

2. Impacto potencial sobre la cantidad de casos expuestos (I_1)

$$FE_e = \frac{I^*}{I_1}$$

I^* → cantidad de casos expuestos atribuibles a la exposición.

I_1 → cantidad de casos expuestos.

El denominador (en la fórmula de **FEe**) es el número de casos expuestos. Esto difiere del denominador en **FE**. El motivo es que el grupo referente para **FEe** es el número de casos expuestos que ocurren en la cohorte, en lugar de el número total de casos en **FE**. El numerador en ambas fórmulas es el mismo, concretamente **I***. En particular, **I*** tanto en **FE** como en **FEe** puede cuantificarse como la diferencia entre el número real de casos y el número de casos que hubiesen ocurrido en ausencia de exposición.

Para ejemplificar el cálculo de la fracción etiológica entre los expuestos, considere nuevamente los resultados de un estudio de seguimiento durante cinco años para determinar si los fumadores que han tenido un ataque cardíaco reducirán su riesgo de morir dejando de fumar. La fracción etiológica calculada anteriormente o, en forma equivalente, el porcentaje de riesgo atribuible de la población calculado para estos datos fue del 35%.

Pacientes ataque cardíaco	Fuma	Dejó	**Total**
Mortalidad	27	14	41
Supervivencia	48	67	115
Total	75	81	156

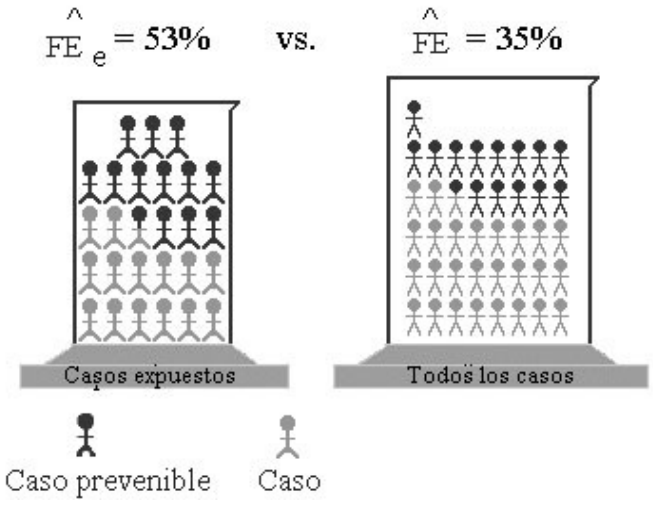
Riesgos				
	continúan fumando:	0.36	Razón de riesgos	2.1
	dejan de fumar:	0.17	Diferencia de riesgos	.19

$$\hat{FE} = \frac{\hat{IA} - \hat{IA}_0}{\hat{IA}} = 35\%$$

$$\hat{FE}_e = \frac{\hat{IA}_1 - \hat{IA}_0}{\hat{IA}_1} = \frac{.19}{.36} = \mathbf{53\%}$$

La fracción etiológica entre los expuestos (**FEe**) puede calcularse para estos mismos datos utilizando la fórmula anterior. El término $\hat{IA}_1$ denota el valor de la incidencia acumulada o el riesgo para todos los individuos expuestos en el estudio; además, $\hat{IA}_0$ denota el valor de la incidencia acumulada para los individuos no expuestos. El numerador en esta fórmula es el valor de la diferencia de riesgos ($\hat{DR}$). Ya que el valor de la diferencia de riesgos es 0,19 y el riesgo para las personas expuestas es 0,36, podemos sustituir estos valores en la fórmula de FEe para obtener 0,53 ó 53%. ¿Cómo interpretamos este resultado?

La fracción etiológica de 0,53 nos dice que el 53% de todas las muertes entre los fumadores que continuaron fumando se deben a la continuación del hábito de fumar. En otras palabras, si pudiésemos haber logrado que todos los fumadores que continuaron fumando y murieron dejaran de fumar después de haber sufrido su ataque cardíaco, habría habido una reducción del 53% en las muertes entre estas personas.



Preguntas de estudio (P6.10)

Pacientes ataque cardíaco	Fuma	Dejó	Total
Mortalidad	27	14	41
Supervivencia	48	67	115
Total	75	81	156

Riesgos { continúan fumando: 0.36 ; dejaron de fumar: 0.17

Razón de riesgos: 2.1

Diferencia de riesgos: .19

1. ¿Cuál es el número en exceso de casos expuestos (es decir, muertes entre los fumadores que continuaron fumando) esperados en ausencia de exposición?
2. Complete los espacios vacíos en este ejemplo, _____ de las _____ muertes entre los fumadores que continuaron fumando podrían haberse evitado.
3. Utilice la fórmula I^*/I_1 para calcular FEe para estos datos.
4. Una fórmula alternativa para la fracción etiológica entre los expuestos es FEe=(RR-1)/RR, donde RR es la razón de riesgos. Utilice esta fórmula para calcular FEe para los datos del estudio sobre ataque cardíaco.
5. El porcentaje de riesgo atribuible de la población (FE) calculado para estos datos es del 35%, mientras que el porcentaje de riesgo atribuible entre los expuestos (FEe) es del 53%. ¿Cómo explica estas diferencias?
6. Para los estudios de cohorte que utilizan información de persona-tiempo, establezca una fórmula para la fracción etiológica entre los expuestos que involucre densidades de incidencia en los grupos expuestos y no expuestos.
7. Como en la pregunta anterior, establezca una fórmula alternativa para FEe que involucre la razón de densidades de incidencia.
8. Para los estudios de casos y controles en los que no puede estimarse el riesgo o la tasa, ¿puede sugerir las

fórmulas para FE y FEe?

Resumen

- La fracción etiológica entre los expuestos (FEe) se centra en el impacto potencial de la exposición en el número de casos expuestos, en lugar del número total de casos.
- FEe se define como I^*/I_1, donde I^* es el número en exceso de casos expuestos debido a la exposición e I_1 es el número real de casos expuestos.
- Para los estudios de cohorte que calculan el riesgo:

$$\hat{\text{FEe}} = \frac{(\hat{\text{IA}}_1 - \hat{\text{IA}}_0)}{\hat{\text{IA}}_1} = \frac{\hat{\text{RR}} - 1}{\hat{\text{RR}}}$$

- Para los estudios de cohorte que calculan la tasa:

$$\hat{\text{FEe}} = \frac{(\hat{\text{DI}}_1 - \hat{\text{DI}}_0)}{\hat{\text{DI}}_1} = \frac{\text{R}\hat{\text{DI}} - 1}{\text{R}\hat{\text{DI}}}$$

Fracción etiológica: ejemplo

Se sabe que el hipotiroidismo, un estatus de enfermedad en el que hay una reducción de la producción de hormona tiroides, aumenta el riesgo de enfermedad cardiovascular. En las mujeres ancianas, la forma subclínica del hipotiroidismo es muy prevalente. Por lo tanto, los investigadores del estudio de Rotterdam examinaron el impacto potencial del hipotiroidismo subclínico en la incidencia del infarto de miocardio.

En este estudio de casi 1.000 mujeres de 55 años y más, la prevalencia del hipotiroidismo subclínico fue del 10,8%. Considere el cuadro de 2x2 representado aquí. La incidencia acumulada del infarto de miocardio es del 2,9% (3/103) en las mujeres con hipotiroidismo subclínico, del 1,2% (10/854) en las mujeres sin hipotiroidismo y del 1,4% (13/957) en general.

	Hipotiroidismo subclínico	Sin hipotiroidismo subclínico	Total
IM	3	10	13
No IM	100	844	944
Total	103	854	957
	$\hat{\text{IA}}_1$=2,9%	$\hat{\text{IA}}_0$=1,2%	$\hat{\text{IA}}$=1,4%

Preguntas de estudio (P6.11)

1. Utilizando estos datos, ¿puede calcular la fracción etiológica?

La fracción etiológica es: FE = (1,4 -1,2)/ 1,4 = 14%. Esto indica que de todos los infartos de miocardio que se presentan en las mujeres ancianas, el 14% se deben a la presencia de hipotiroidismo subclínico. En otras palabras, si el hipotiroidismo subclínico pudiese evitarse, habría un 14% menos de infartos de miocardio en esta población.

Preguntas de estudio (P6.11) (continuación)

2. ¿Puede calcular la fracción etiológica mediante la fórmula alternativa:

FE = [p(RR-1)]/[p(RR-1) + 1]?

3. ¿Puede calcular la fracción etiológica en los expuestos?

La fracción etiológica en los expuestos (FEe) es (2,9-1,2), que es igual a la diferencia de riesgos, dividido por 2,9, lo que da como resultado 60%. Por lo tanto, entre las mujeres afectadas, un 60% de los infartos de miocardio pueden atribuirse a la presencia de hipotiroidismo subclínico.

Resumen

- Los investigadores del estudio de Rotterdam examinaron el impacto potencial del hipotiroidismo subclínico en la incidencia del infarto de miocardio.
- De todos los infartos de miocardio que se presentan en las mujeres ancianas, 14% se deben a la presencia de hipotiroidismo subclínico.
- Entre las mujeres afectadas, 60% de los infartos de miocardio pueden atribuirse a la presencia de hipotiroidismo subclínico.

Auto-evaluación (P6.12)

Considere los datos en el siguiente cuadro sobre el tabaquismo y la incidencia del cáncer de pulmón y la enfermedad coronaria (EC).

Incidencia del cáncer de pulmón

	Fumadores	**No fumadores**	**Total**
Nuevos casos de cáncer de pulmón	60.000	10.000	70.000
Cantidad estimada de personas-años	70.000.000	150.000.000	220.000.000
Densidad de incidencia	$\hat{DI}_1$=85,7 por 100.000 personas-años	$\hat{DI}_0$=6,7 por 100.000 personas-años	$\hat{DI}$=31,8 por 100.000 personas-años

Incidencia de la enfermedad coronaria (EC)

	Fumadores	**No fumadores**	**Total**
Casos nuevos de EC	250.000	250.000	500.000
Cantidad estimada de personas-años	70.000.000	150.000.000	220.000.000
Densidad de incidencia	$\hat{DI}_1$=357,1 por 100.000 personas-años	$\hat{DI}_0$=166,7 por 100.000 personas-años	$\hat{DI}$=227,3 por 100.000 personas-años

1. La prevalencia del tabaquismo en esta población es . . . . **???**.
2. La fracción etiológica para el cáncer de pulmón es . . . **???**.
3. La fracción etiológica para la enfermedad coronaria es . . . . **???**.
4. La fracción etiológica entre los expuestos para el cáncer de pulmón es . . **???**.
5. La fracción etiológica entre los expuestos para la EC es . . . . **???**.
6. La proporción de cáncer de pulmón entre los fumadores atribuible a su tabaquismo es **???**.

Opciones

0,0% **26,7%** **31,8%** **53,3%** **79,0%** **92,2%**

En cada una de las siguientes frases indique si se trata de **diferencia de riesgos/tasas**, **fracción etiológica** o **fracción etiológica entre los expuestos**.

7. Porcentaje de riesgo atribuible entre los expuestos. . . **???**
8. Riesgo atribuible de la población. . . . . **???**
9. Exceso de riesgo. . . . . **???**
10. Influenciada por la prevalencia de la exposición en la población. . **???**
11. Tiene las mismas unidades que una medida de la ocurrencia. . **???**
12. Puede ser un número negativo. . . . . **???**

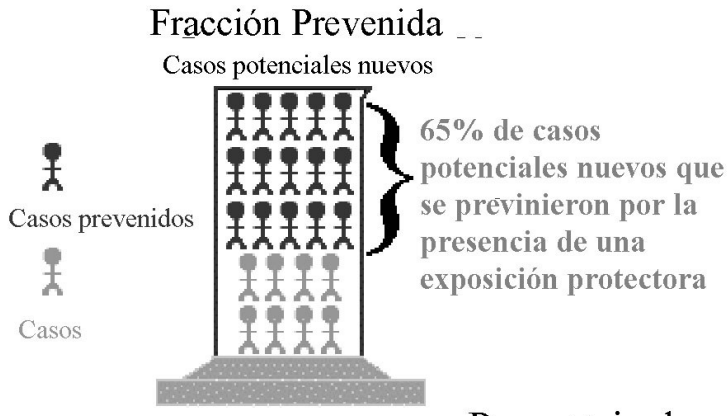
13. **???** nunca puede ser mayor que **???**

Opciones (para la pregunta 13)
Fracción etiológica **Fracción etiológica entre los expuestos**

6-3 Medidas de impacto potencial (continuación)

Fracción prevenida

La **fracción prevenida** brinda la proporción de casos potenciales nuevos que se previnieron por la presencia de una exposición **protectora**. Otros nombres para la fracción prevenida son **fracción prevenida *de la población***, **riesgo poblacional prevenido** y **porcentaje de riesgo prevenido de la población**.



Riesgo prevenido de la población = FP = Porcentaje de riesgo prevenido de la población

La fracción prevenida puede expresarse matemáticamente de la siguiente forma:

$$\hat{\text{FP}} = \frac{\text{I}^{**}}{\text{I}^{**} + \text{I}}$$

El numerador **I**** denota el número de casos nuevos que se esperarían en ausencia de exposición *menos* el número de casos nuevos que se presentaron realmente. En otras palabras, el numerador muestra el número de casos prevenidos por la presencia de una exposición protectora.

La **I** en el denominador indica el número total de casos nuevos que realmente se presentaron durante el seguimiento de la cohorte. La suma de **I**** más **I** indica el número de casos nuevos que se esperarían en *ausencia* de exposición. Es decir, el denominador indica el número potencial de casos nuevos.

Para ejemplificar el cálculo de una fracción prevenida, consideramos un estudio de una epidemia de sarampión en Texarkana, una ciudad bisecada por el límite estatal entre Texas y Arkansas. Esta ciudad nunca tuvo una campaña de vacunación contra el sarampión, aunque una gran proporción de los niños había sido vacunada anteriormente. A continuación se muestran los resultados del estudio:

Estudio Texarkana	Vacunado	No vacunado	Total
Enfermo	27	512	539
No enfermo	6323	4323	10646
Total	6350	4835	11185

Incidencia acumulada por 1,000

Vacunado: 4.2

No vacunado: 105.9

General: 48.2

$\hat{\mathbf{RR}} = 0.04$

$\hat{\mathbf{DR}} = -101.7$ por 1,000

Una fórmula para calcular la fracción prevenida es:

$$\hat{\text{FP}} = \frac{\hat{\text{IA}}_0 - \hat{\text{IA}}}{\hat{\text{IA}}_0}$$

Como en la fórmula de la fracción etiológica, $\hat{\text{IA}}_0$ denota la incidencia acumulada en los no expuestos e $\hat{\text{IA}}$ denota la incidencia acumulada general. Esta fórmula difiere de la fórmula de la fracción etiológica *tanto* en el numerador *como* en el denominador; es necesario cuantificar la exposición como protectora, en lugar de nociva.

En nuestro ejemplo, $\hat{\text{IA}}_0$ es 105,9 por 1.000 personas e $\hat{\text{IA}}$ es 48,2 por 1.000 personas. Si sustituimos estos valores en la fórmula, descubrimos que la fracción prevenida aquí es 0,55 ó 55%.

$$\hat{\text{FP}} = \frac{\hat{\text{IA}}_0 - \hat{\text{IA}}}{\hat{\text{IA}}_0} = \frac{105{,}9 - 48{,}2}{105{,}9} = 0{,}55 = 55\%$$

Preguntas de estudio (P6.13)

Si nos basamos en el ejemplo de Texarkana:

1. ¿Cuántos niños enfermos se esperarían en total si ninguno de los individuos hubiese sido inmunizado?
2. ¿Cuántos niños se enfermaron, en verdad, en toda la cohorte?
3. ¿Cuántos casos fueron prevenidos mediante el programa de inmunización?
4. Utilizando los valores I**=645,5 e I=539, calcule la fracción prevenida.

¿Cómo interpretamos este resultado? La fracción prevenida 0,55 nos dice que de un total de 1.184,5 casos de enfermedad que se esperarían si ningún individuo estuviese inmunizado, se previnieron 645,4 casos de enfermedad mediante la inmunización. Es decir, hemos observado 645,4 casos menos de enfermedad que los esperados debido a la vacunación. Por lo tanto, el 55% del total de casos esperados se previnieron mediante la vacunación.

Resumen

- La fracción prevenida (FP) brinda la proporción de nuevos casos potenciales que se previnieron por la exposición.
- Otros nombres para la FP son riesgo prevenido de la población y porcentaje de riesgo prevenido de la población.
- FP=I** / (I** + I) donde I** denota los casos prevenidos totales e I denota los casos totales que se presentaron.
- Mediante los datos de incidencia acumulada,

  $\hat{FP} = \frac{\hat{IA}_0 - \hat{IA}}{\hat{IA}_0} = p(1 - \hat{RR})$ donde $\hat{IA}$, $\hat{RR}$ y p son iguales que en la fórmula de FE.
- Mediante los datos de persona tiempo,

  $\hat{FP} = \frac{(\hat{DI}_0 - \hat{DI})}{\hat{DI}_0} = (p^*)(1 - R\hat{DI})$ donde $\hat{DI}_0$, $\hat{DI}$, $R\hat{DI}$,

  $p^* = \frac{L_1}{L_1 + L_0}$ y L_1 y L_0 son iguales a los términos de la fórmula de FE correspondiente.

La relación entre FP y FE

Hemos visto las fórmulas para la **fracción etiológica** (**FE**):

$$FE = \frac{\hat{I}A - \hat{I}A_0}{\hat{I}A} = \frac{\hat{p}(\hat{R}R - 1)}{\hat{p}(\hat{R}R - 1) + 1}$$

Aquí se muestran las fórmulas correspondientes para la **fracción prevenida** (**FP**):

$$FP = \frac{\hat{I}A_0 - \hat{I}A}{\hat{I}A_0} = \hat{p}(1 - \hat{R}R)$$

Con un poco de álgebra, podemos escribir FP en términos de FE de la siguiente manera:

$$FP = \frac{FE}{FE - 1}$$

También podemos escribir FPe en términos de FEe de la siguiente manera:

$$FPe = \frac{FEe}{FEe - 1}$$

Ejemplificamos la relación entre FP y FE mediante los resultados del estudio de seguimiento de 5 años en pacientes que sufrieron un ataque cardíaco, descrito en varias actividades en esta lección. FE = 0,35 y FEe = 0,53 cuando se mide el impacto potencial de continuar fumando. Calculamos FE en lugar de FP porque los fumadores que continuaron fumando tenían un riesgo estimado más alto (0,36) que los fumadores que dejaron de fumar (0,17). Ahora suponga que invertimos nuestra pregunta sobre el impacto potencial y preguntamos cuál es la fracción prevenida (**FP**) de dejar de fumar. Para responder esto, debemos intercambiar las incidencias acumuladas, de forma que $\mathbf{IA_0}$ indica la incidencia acumulada calculada en el grupo de los fumadores que continuaron fumando (0,36) e $\mathbf{IA_1}$ indica la incidencia acumulada calculada en el grupo de los ex fumadores (0,173). La incidencia acumulada general **IA** continúa siendo 0,263.

La FP en este caso está dada por (0,36-0,263)/0,36 = 0,269 = 26,9%. Por lo tanto, el 26,9% de las muertes esperadas pueden prevenirse dejando de fumar. Esto se contrapone con la fracción etiológica de 0,35, que dice que el 35% de todas las muertes son atribuibles a seguir fumando. Observe que la FE para dejar de fumar es el valor negativo dado por (0,263-0,36)/0,263=-0,368. Substituyendo FE por -0,368 en la fórmula FP=FE/(FE-1), obtenemos FP=-0,368/(-0,368-1) = 0,269, que es la fracción prevenida de dejar de fumar. Por lo tanto, la fracción etiológica del 35% para continuar fumando no es igual a la fracción prevenida del 26,9% de dejar de fumar. Tampoco obtendremos el valor de 0,269 si sustituimos el valor de FE de 0,35 en la fórmula que relaciona FP con FE. Esta fórmula sólo funciona si cambiamos las incidencias acumuladas de IA_1 a IA_0 y viceversa.

Fracción prevenida entre los expuestos

Al igual que con la fracción etiológica, hay dos formulaciones conceptuales de la fracción prevenida. Una, que hemos descrito anteriormente, se centra en el impacto potencial de una exposición protectora sobre **el número total de casos** que se esperarían en ausencia de exposición.

$$\hat{FP} = \frac{I^{**}}{I^{**} + I} = \frac{\hat{IA}_0 - \hat{IA}}{\hat{IA}_0} = p(1 - RR)$$

Una segunda formulación de la fracción prevenida se centra en el impacto potencial de una exposición protectora sobre el número esperado de **casos expuestos**. Esta medida se llama **fracción prevenida entre los expuestos**, o en forma alternativa, **porcentaje de riesgo prevenido entre los expuestos**.

$$\hat{FP}_e = \frac{I^{**}}{I^{**} + I_1}$$

Aunque el denominador en la fracción prevenida entre los expuestos (FPe) es diferente del denominador en la fracción prevenida (FP), los numeradores son iguales.

$$\hat{FP} = \frac{I^{**}}{I^{**} + I} = \frac{\text{Casos prevenidos}}{\text{Casos prevenidos} + \text{Cantidad total de casos}}$$

$$\hat{FP}_e = \frac{I^{**}}{I^{**} + I_1} = \frac{\text{Casos prevenidos}}{\text{Casos prevenidos} + \text{Casos expuestos}}$$

Para ejemplificar el cálculo de la fracción prevenida entre los expuestos, considere nuevamente los resultados de la epidemia de sarampión en Texarkana en 1970 y 1971. La fracción prevenida calculada anteriormente, o en forma equivalente, el porcentaje de riesgo prevenido de la población para estos datos es del 55%.

Fracción prevenida

Estudio Texarkana	Vacunado	No vacunado	Total
Enfermo	27	512	539
No enfermo	6323	4323	10646
Total	6350	4835	11185

Incidencia acumulada por 1,000

Vacunado: 4.2
No vacunado: 105.9
General: 48.2

$\hat{RR} = 0.04$

$\hat{DR} = \frac{-101.7}{1000}$

$\hat{FP} = .55$ or 55%

$$\hat{FP}_e = \frac{\hat{IA}_0 - \hat{IA}_1}{\hat{IA}_0} = \frac{101.7}{105.9} = 0.96 = 96\%$$

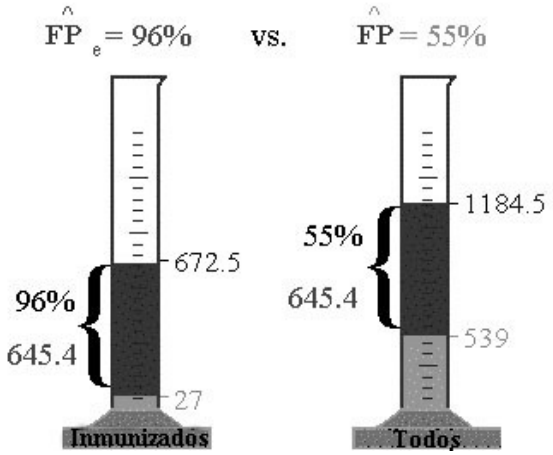
La fracción prevenida entre los expuestos puede calcularse para estos mismos datos mediante la fórmula anterior. $\hat{I}A_0$ e $\hat{I}A_1$indican las incidencias acumuladas calculadas para los individuos no expuestos y expuestos, respectivamente. El numerador en esta fórmula es el negativo de la diferencia de riesgos calculada. La diferencia de riesgos calculada es menos 101,7 por 1.000 y la incidencia acumulada para las personas no expuestas ($\hat{I}A_0$) es 105,9 por 1.000. Por lo tanto, la fracción prevenida entre los expuestos es 101,7 dividido por 105,9, que da 0,96 ó 96%.

Preguntas de estudio (P6.14)

Utilizando los datos de Texarkana:

1. ¿Cuántos niños enfermos estarían entre los inmunizados si la vacunación no fue efectiva?
2. ¿Cuántos niños se enfermaron realmente entre los que fueron inmunizados?
3. ¿Cuántos casos expuestos fueron prevenidos mediante el programa de vacunación?
4. Utilizando los valores I**=645,5 e I_1 = 27, calcule la fracción prevenida entre los expuestos.

¿Cómo interpretamos este resultado? La fracción prevenida de 0,96 entre los expuestos nos dice que el 96% del total de casos esperados entre los inmunizados fueron prevenidos por la vacunación. De un total de 672,5 casos de enfermedad esperados entre los inmunizados, 96% ó 645,4 casos de enfermedad fueron prevenidos por la vacunación. Es decir, hubo 645,4 casos menos de enfermedad que los esperados entre los niños debido a la vacunación.



Preguntas de estudio (P6.15)

1. Una fórmula alternativa para la fracción prevenida entre los expuestos es FPe = (1 – RR), donde RR es la razón de riesgos. Utilice esta fórmula para calcular FPe para los datos de vacunación.
2. El porcentaje de riesgo prevenido de la población (FP) calculado para estos datos es del 55%, mientras que el porcentaje de riesgo prevenido entre los expuestos (FPe) es del 96%. ¿Cómo explica la diferencia entre estos dos valores?
3. Para los estudios de cohorte que utilizan información de persona-tiempo, establezca una fórmula para la fracción prevenida entre los expuestos que involucre densidades de incidencia en los grupos expuestos y

no expuestos.

4. Como en la pregunta anterior, establezca una fórmula alternativa para FPe que involucre la razón de densidades de incidencia.
5. Para los estudios de casos y controles en los que no puede estimarse el riesgo o la tasa, ¿puede sugerir las fórmulas para FP y FPe?

Resumen

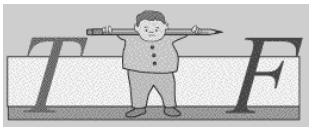
- La fracción prevenida entre los expuestos (FPe) se centra en el impacto potencial de la exposición *en los casos expuestos*, en lugar del número total de casos.
- FPe se define como I** / (I** + I_1), donde I** es el número de casos expuestos prevenidos por la exposición e I_1 es el número real de casos expuestos.
- Para los estudios de cohorte que calculan el riesgo:

$$\hat{\text{FPe}} = \frac{\hat{\text{IA}}_0 - \hat{\text{IA}}_1}{\hat{\text{IA}}_0} = (1 - \hat{\text{RR}})$$

- Para los estudios de cohorte que calculan la tasa:

$$\hat{\text{FPe}} = \frac{(\hat{\text{DI}}_0 - \hat{\text{DI}}_1)}{\hat{\text{DI}}_0} = (1 - \hat{\text{RDI}})$$

Auto-evaluación (P6.16)



Coloque **verdadero** o **falso** en cada una de las siguientes afirmaciones:

1. Cuando la diferencia de riesgos es < 0, se debe usar la fracción prevenida en lugar de la fracción etiológica. . . . . **???**
2. La fracción prevenida puede calcularse simplemente como 1 –FE cuando la diferencia de riesgos es < 0. . . . . **???**
3. La fracción prevenida entre los expuestos puede calcularse simplemente como 1 -RR. . **???**
4. La fracción prevenida se basa en casos teóricos o inexistentes, es decir, casos que no ocurrieron pero que se hubieran presentado en ausencia de la intervención o exposición. . . **???**
5. El denominador tanto de FP como de FPe es el riesgo (o la tasa) en el grupo no expuesto. **???**
6. Entre FP y FPe, sólo la FP está influenciada por la prevalencia de la exposición. . . **???**

Mediante los datos del siguiente cuadro, identifique cada una de las siguientes afirmaciones como:

fracción prevenida, **fracción prevenida entre los expuestos** o **ninguna de las dos**.

7. Podría calcularse como ([44/160] -[48/888]) / [44/160] = 80,3%. . . . . **???**
8. Podría calcularse como ([44/160] -[92/1048]) / [44/160] = 68,1%. . . . . **???**
9. Si se hubiese vacunado a menos niños, esta medida **aumentaría**. . . . **???**

10. Si se hubiese vacunado a menos niños, esta medida **disminuiría** . . . **???**

11. La proporción de casos potenciales en la comunidad prevenidos por la vacunación. . **???**

	Vacunados	No vacunados	Total
Sarampión	48	44	92
Sin sarampión	840	116	956
Total	888	160	1.048

IA en no vacunados = 27,5%; IA en vacunados = 5,4%; IA general = 8,8%.

Nomenclatura:

Organización en forma de cuadro para los estudios de cohorte, casos y controles y prevalencia:

	Expuestos	No expuestos	Total
Enfermedad/casos	a	b	n_1
Sin enfermedad/controles	c	d	n_0
Total	m_1	m_0	N

Organización en forma de cuadro correspondiente a datos de cohorte con persona-tiempo:

	Expuestos	No expuestos	Total
Enfermedad (casos nuevos)	I_1	I_0	I
Sin enfermedad	-	-	-
Persona-tiempo total sin enfermedad	PT_1 o L_1	PT_0 o L_0	PT

$\hat{I}A$	Incidencia acumulada en la población (n_1/n)
$\hat{I}A_0$	Incidencia acumulada en los no expuestos (b/m_0)
$\hat{I}A_1$	Incidencia acumulada en los expuestos (a/m_1)
DIA	Diferencia de incidencias acumuladas o diferencia de riesgos, $IA_1 - IA_0$
FE	Fracción etiológica
FEe	Fracción etiológica entre los expuestos
I	Número de casos nuevos que se presentan
I*	Número de casos nuevos atribuibles a la exposición
I**	Número de casos que se esperarían en ausencia de exposición
I_0	Número de casos nuevos en los no expuestos
I_1	Número de casos nuevos en los expuestos
$\hat{D}I$	Densidad de incidencia (o "tasa") en la población (I/PT)
$\hat{D}I_0$	Densidad de incidencia (o "tasa") en los no expuestos (I_0/PT_0)
$\hat{D}I_1$	Densidad de incidencia (o "tasa") en los expuestos (I_1/PT_1)
DDI	Diferencia de densidades de incidencia o diferencia de tasas, $DI_1 - DI_0$
RDI	Razón de densidades de incidencia o razón de tasas: DI_1 / DI_0
L_0	Persona-tiempo sin enfermedad en los no expuestos
L_1	Persona-tiempo sin enfermedad en los expuestos
N	Tamaño de la población en estudio
n_0	Tamaño de la población en estudio en los no expuestos en el momento cero
n_1	Tamaño de la población en estudio en los expuestos en el momento cero
NNT	Número necesario a tratar, $1/(IA_1-IA_0)$
OR	*Odds ratio*: ad/bc
P_1	Prevalencia en los expuestos
P_0	Prevalencia en los no expuestos
DP	Diferencia de prevalencias
FP	Fracción prevenida
FPe	Fracción prevenida en los expuestos
PT	Persona-tiempo sin enfermedad
PT_0	Persona-tiempo sin enfermedad en los no expuestos
PT_1	Persona-tiempo sin enfermedad en los expuestos
R_0	Riesgo en los no expuestos
R_1	Riesgo en los expuestos

DR	Diferencia de riesgos: el riesgo en los expuestos menos el riesgo en los no expuestos
RR	Razón de riesgos: el riesgo en los expuestos dividido por el riesgo en los no expuestos

Fórmulas para medidas de efecto en forma de diferencias por datos de riesgo, datos de tasa y datos de prevalencia.

Datos de riesgo	**Datos de tasa**	**Datos de prevalencia**
$\hat{DIA} = \hat{IA}_1 - \hat{IA}_0$	$\hat{DDI} = \hat{DI}_1 - \hat{DI}_0$	$\hat{DP} = \hat{P}_1 - \hat{P}_0$

Fórmulas para comparar FP con FE y FPe con FEe y fórmula de NNT:

$FP = \frac{FE}{FE\text{-}1}$	$FPe = \frac{FEe}{FEe - 1}$	$NNT = \frac{1}{\hat{IA}_1 - \hat{IA}_0}$

Fórmulas para fracción etiológica, fracción etiológica en los expuestos, fracción prevenida y fracción prevenida en los expuestos para datos de riesgo, datos de tasa, así como estudios de casos y controles.

	Datos de riesgo	**Datos de tasa**	**Casos y controles***
FE	$\hat{FE} = \frac{I^*}{I}$	$\hat{FE} = \frac{I^*}{I}$	
	$\hat{FE} = \frac{\hat{IA} - \hat{IA}_0}{\hat{IA}}$	$\hat{FE} = \frac{\hat{DI} - \hat{DI}_0}{\hat{DI}}$	
	$\hat{FE} = \frac{(p)(\hat{RR} - 1)}{(p)(\hat{RR} - 1) + 1}$	$\hat{FE} = \frac{(p^*)(R\hat{DI} - 1)}{(p^*)(R\hat{DI} - 1) + 1}$	$\hat{FE} = \frac{(p')(\hat{OR} - 1)}{(p')(\hat{OR} - 1) + 1}$
FEe	$\hat{FE}_e = \frac{I^*}{I_1}$	$\hat{FE}_e = \frac{I^*}{I_1}$	
	$\hat{FE}_e = \frac{\hat{IA}_1 - \hat{IA}_0}{\hat{IA}_1}$	$\hat{FE}_e = \frac{\hat{DI}_1 - \hat{DI}_0}{\hat{DI}_1}$	
	$\hat{FE}_e = \frac{\hat{DR}}{\hat{IA}_1}$	$\hat{FE}_e = \frac{\hat{DDI}}{\hat{DI}_1}$	
	$\hat{FE}_e = \frac{\hat{RR} - 1}{\hat{RR}}$	$\hat{FE}_e = \frac{R\hat{DI} - 1}{R\hat{DI}}$	$\hat{FE}_e = \frac{\hat{OR} - 1}{\hat{OR}}$
FP	$\hat{FP} = \frac{I^{**}}{I^{**} + I}$	$\hat{FP} = \frac{I^{**}}{I^{**} + I}$	
	$\hat{FP} = \frac{\hat{IA}_0 - \hat{IA}}{\hat{IA}_0}$	$\hat{FP} = \frac{\hat{DI}_0 - \hat{DI}}{\hat{DI}_0}$	
	$\hat{FP} = p(1 - \hat{RR})$	$\hat{FP} = (p^*)(1 - R\hat{DI})$	$\hat{FP} = p'(1 - \hat{OR})$
FPe	$\hat{FP}_e = \frac{I^{**}}{I^{**} + I_1}$	$\hat{FP}_e = \frac{I^{**}}{I^{**} + I_1}$	
	$\hat{FP}_e = \frac{\hat{IA}_0 - \hat{IA}_1}{\hat{IA}_0}$	$\hat{FP}_e = \frac{\hat{DI}_0 - \hat{DI}_1}{\hat{DI}_0}$	
	$\hat{FP}_e = 1 - \hat{RR}$	$\hat{FP}_e = 1 - R\hat{DI}$	$\hat{FP}_e = 1 - \hat{OR}$

donde $p = \frac{m_1}{n}$ $p^* = \frac{L_1}{L_1 + L_0}$ $p' = \frac{c}{n_0}$

*En los estudios de casos y controles, la FE, FEe, FP, FPe basadas en el *odds ratio* serán buenos cálculos cuando el OR es un buen cálculo de la RR (por ejemplo, el supuesto de enfermedad poco frecuente).

Referencias

Benichou J, Chow WH, McLaughlin JK, Mandel JS, and Fraumeni JF Jr. *Population attributable risk of renal cell cancer in Minnesota*. Am J Epidemiol 1988; 148(5):424-30.

Cook RJ, Sackett DL. *The number needed to treat: a clinically useful measure of treatment effect*. BMJ 1995; 310(6977):452-4.

Greenberg RS, Daniels SR, Flanders WD, Eley JW, Boring JR. *Medical Epidemiology* (3rd Ed). Lange Medical Books, New York, 2001.

Hak AE, Pols HA, Visser TJ, Drexhage HA, Hofman A, Witteman JC. *Subclinical hypothyroidism is an independent risk factor for atherosclerosis and myocardial infarction in elderly women: the Rotterdam Study*. Ann Intern Med 2000; 132(4):270-8.

Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic Research: Principles and Quantitative Methods*. John Wiley and Sons Publishers, New York, 1982.

Landrigan PJ, *Epidemic measles in a divided city*. JAMA 1972; 221(6):567-70.

Medical Research Council trial of treatment of mild hypertension: principal results. MRC Working Party. BMJ 1985:291; 97-104.

Morgenstern H and Bursic ES. *A method for using epidemiologic data to estimate the potential impact of an intervention on the health status of a target population*. J Community Health 1982; 7(4):292-309.

Spirtas R, Heineman EF, Bernstein L, Beebe GW, Keehn RJ, Stark A, Harlow BL, Benichou J. *Malignant mesothelioma: attributable risk of asbestos exposure*. Occup Environ Med 1994; 51(12):804-11.

Wacholder S, Benichou J, Heineman EF, Hartge P, Hoover RN. *Attributable risk: advantages of a broad definition of exposure*. Am J Epidemiol 1994; 140(4):303-9..

Walter SD. *Calculation of attributable risks from epidemiological data*. Int J Epidemiol 1978; 7(2):175-82.

Walter SD. *Prevention for multifactorial diseases*. Am J Epidemiol 1980; 112(3):409-16.

Walter SD. *Attributable risk in practice*. Am J Epidemiol 1998; 148(5):411-3.

Walter SD. *Number needed to treat (NNT): estimation of a measure of clinical benefit*. Stat Med 2001:20(24); 3947-62.

Wilson PD, Loffredo CA, Correa-Villasenor A, Ferencz C. *Attributable fraction for cardiac malformations*. Am J Epidemiol 1998; 148(5):414-23.

Ejercicios

EPI-1. Medidas del impacto potencial: dejar de fumar II

Los siguientes datos surgen de un estudio de enfoques de auto ayuda para dejar de fumar. (Estos son los mismos datos descritos en la pregunta 4 de los ejercicios de la Lección 5). Los fumadores que querían dejar de fumar fueron clasificados en forma aleatoria en uno de cuatro grupos (C = control, M = manual para dejar de fumar, MS = manual + folleto de apoyo social, MST = manual + folleto de apoyo social + asesoría por teléfono). La condición de tabaquismo se midió por medio de encuestas por correo enviadas 8, 16 y 24 meses después de la formación de los grupos en forma aleatoria. Estos son los resultados del mes 16:

	Grupo formado en forma aleatoria				
Condición	**C**	**M**	**MS**	**MST**	**Total**
Dejaron de fumar	84	71	67	109	331
Fuman	381	396	404	365	1.546
Total	465	467	471	474	1.877

Calcule una medida adecuada de impacto, comparando los grupos MST y MS. Interprete el resultado.

EPI-2. Medidas de impacto potencial: la radioterapia entre los niños II

(Esta es la continuación de la pregunta 5 de los ejercicios de tarea de la lección 5). En un estudio sobre los efectos adversos de la radioterapia en los niños de Israel, se identificaron 10.834 niños irradiados de los registros originales del tratamiento y se combinaron con 10.834 individuos de comparación no irradiados que fueron seleccionados de la población general. Se realizó un seguimiento de los individuos durante una media de 26 años. La cantidad de personas-años de observación fue: individuos irradiados, 279.901 personas-años; individuos de comparación, 280.561 personas-años. Durante el período de seguimiento hubo 49 muertes por cáncer en los individuos irradiados, y 44 en los individuos de comparación de la población no irradiados.

a. Suponiendo la causalidad, ¿cuántas muertes por cáncer por 100.000 individuos irradiados por año se debieron al efecto de la radioterapia?
b. Suponiendo nuevamente la causalidad, ¿qué proporción de muertes por cáncer en individuos irradiados pueden atribuirse al efecto de la radioterapia?

EPI-3. Fracción etiológica y fracción etiológica en los expuestos: EC y cáncer de pulmón

El siguiente cuadro muestra la mortalidad por enfermedad cardiovascular (ECV) y la mortalidad por cáncer de pulmón (CP) entre médicos fumadores y no fumadores obtenidas de un estudio prospectivo que incluyó a médicos enumerados en el Registro Médico Británico que vivían en Inglaterra y Gales desde octubre de 1951. La información sobre la causa de muerte fue obtenida del certificado de defunción y los registros de mortalidad durante los siguientes 10 años.

a. Complete el cuadro siguiente calculando la razón de tasas, la fracción etiológica (FE) y la fracción etiológica en los expuestos (FEe) para cada categoría de tabaquismo. Redondee todos los cálculos con dos decimales.

ECV	No fumadores	Todos	Fumadores con mayor consumo de cigarrillos (≥25 cig/día)	Población general
Tasa de mortalidad por 1.000 PA	7,32	9,51	9,53	9,40
Razón de tasas		_____	_____	
Fracción etiológica (FE)		_____	_____	
Fracción etiológica en los expuestos (FEe)		_____	_____	
Cáncer de pulmón				
Tasa de mortalidad por 1.000 PA	0,07	1,30	2,27	0,93
Razón de tasas		_____	_____	
Fracción etiológica (FE)		_____	_____	
Fracción etiológica en los expuestos (FEe)		_____	_____	

b. Si sólo considera a "todos los fumadores" (comparado con los no fumadores), ¿para cuál de las dos enfermedades el tabaquismo tiene mayor importancia etiológica? ¿Por qué?
c. Si sólo considera a "todos los fumadores" (comparado con los no fumadores), ¿para cuál de las dos enfermedades el tabaquismo tiene mayor importancia con relación a la salud pública? ¿Por qué?
d. ¿Cómo contestaría las dos partes anteriores (es decir, b y c) si considera a los "fumadores con mayor consumo de cigarrillos" en comparación con los "no fumadores"?

EPI-4. Fracción etiológica: exposición al amianto vs. mesotelioma

El cuadro a continuación representa los resultados de un estudio que investigó si existe una asociación entre la exposición al amianto y el cáncer poco frecuente llamado mesotelioma:

		Exposición al amianto	
		Sí	**No**
Mesotelioma	**Sí**	20	80
	No	3	93

a. ¿Por qué tiene sentido pensar que el diseño de estudio utilizado fue un estudio de casos y controles, en lugar de un estudio de cohorte?
b. Suponiendo que la medida de efecto de interés es el *odds ratio*, calcule medidas de la fracción etiológica de la población y la fracción etiológica entre los expuestos mediante las siguientes fórmulas:

FE = p'(OR - 1) / [p'(OR - 1) +1] y FEe = (OR - 1) / OR, donde OR es el *odds ratio* y p' es la proporción de todos los controles que están expuestos.

c. ¿Cómo interpreta estas medidas?

EPI-5. Impacto potencial: cólera

En 1963, el Laboratorio de Investigación del Cólera en Bangladesh evaluó el estatus nutricional de una muestra de probabilidad de niños de 12 a 23 meses. Estos niños fueron seguidos durante dos años y se identificaron todas las muertes. En el cuadro siguiente se muestran los resultados:

Número de muertes durante dos años de seguimiento de niños de 12 a 23 meses al momento de la incorporación al estudio: Bangladesh, 1963-5.

	Estatus nutricional		
	Normal	**Desnutrición moderada**	**Desnutrición severa**
Muertes	20	44	47
Sobrevivientes	526	1002	380
Total	546	1046	427

a. Calcule el riesgo de mortalidad para el grupo con desnutrición moderada causada por estar moderadamente desnutridos (es decir, la diferencia de riesgos para la desnutrición moderada comparada con la nutrición normal).
b. ¿Cuál es la fracción etiológica para los niños moderadamente desnutridos (en comparación con los niños con nutrición normal)?
c. Calcule el riesgo de mortalidad para el grupo con desnutrición severa causada por estar severamente desnutridos (es decir, la diferencia de riesgos para la desnutrición severa comparada con la nutrición normal).
d. ¿Cuál es la fracción etiológica para los niños severamente desnutridos (en comparación con los niños con nutrición normal)?
e. En una sola oración y utilizando las cifras que ha calculado para los niños con desnutrición severa como ejemplo, explique cuál es el significado de la diferencia de riesgos.
f. En una sola oración y utilizando las cifras que ha calculado para los niños con desnutrición severa como ejemplo, explique cuál es el significado de la fracción etiológica.

EPI-6. Impacto potencial: defectos del tubo neural

En 1988, Mulinare et al informaron sus resultados respecto a la asociación entre los defectos del tubo neural (DTN) y el uso periconcepcional de multivitaminas (**JAMA 260**:3141-3145). Seleccionaron varios grupos de bebés. Un grupo estaba compuesto por "todos los bebés nacidos vivos o muertos con el diagnóstico de anencefalia o espina bífida desde 1968 a 1980 que fueron registrados en el Programa de defectos congénitos del área metropolitana de Atlanta (MACDP, por sus siglas en inglés). El segundo grupo estaba compuesto por bebés "nacidos vivos" sin defectos congénitos elegidos en forma aleatoria entre todos los bebés nacidos vivos en el área de vigilancia del MACDP. "Obtuvimos datos sobre el empleo de multivitaminas y definimos su uso como 'el consumo de multivitaminas o multivitaminas prenatales durante cada mes del período de seis meses completo […] desde tres meses antes de la concepción hasta el tercer mes de embarazo' ". El siguiente cuadro muestra los resultados parciales de ese informe (con algunas simplificaciones para los fines de este ejercicio).

Distribución del uso de vitamina periconcepcional entre las madres que dieron a luz a bebés con y sin tubo neural

	Uso de multivitaminas		
	Uso periconcepcional	**Sin uso de vitaminas**	**Total**
Casos	24	159	183
Controles	411	1.052	1.463

a. Con los datos anteriores, ¿puede calcular la prevalencia de los defectos del tubo neural en el área del MACDP? De ser así, hágalo, si no, indique por qué no puede.
b. Utilizando los datos anteriores, ¿puede calcular el riesgo de defectos del tubo neural en el área del MACDP? De ser así, hágalo, si no, indique por qué no puede.
c. Con los datos anteriores, ¿puede calcular la fracción etiológica para los DTN provocados por la falta de uso periconcepcional de multivitaminas por parte de la madre? De ser así, hágalo, si no, indique por qué no puede.

Respuestas a las preguntas de estudio y auto-evaluaciones

P6.1

1. 75 * 0,17 = 12,75 muertes habrían ocurrido si los 75 pacientes hubiesen dejado de fumar.
2. 27 – 12,75 = 14,25 muertes en exceso entre los que continuaron fumando.
3. p(muertes en exceso entre los que continuaron fumando) = 14,25 / 75 = 0,19 = diferencia de riesgos.
4. 1.000 x 0,19 = 190 muertes en exceso podrían evitarse.
5. La mayor diferencia de riesgos posible es $\pm$ 1. Sin embargo, esto no significa que 0,19 es pequeña en relación con el valor de referencia clínicamente significativo, que sería lo deseable.
6. Una opción para el valor de referencia es el riesgo para los expuestos, es decir, 0,36. La razón 0,19/0,36 = 0,53 indica que el 53% del riesgo para el grupo de los fumadores que continuaron fumando se reduciría si este grupo hubiese dejado de fumar.

P6.2

1. 100.000 x 14 / 68.239 = 20,5 es el número esperado de muertes relacionadas con la EC por 100.000 personas-años si las personas con colesterol moderadamente alto lo bajaran a valores normales. Por lo tanto, se podrían evitar 71,1 – 20,5 = 50,6 muertes relacionadas con la EC por 100.000 personas-años.
2. 18,5 / 36.581 = 50,6 (71,1-20,5) muertes en exceso relacionadas con la EC por 100.000 personas-años. Este valor de 50,6 por 100.000 es la diferencia de tasas o tasa atribuible.
3. La mayor diferencia de tasas posible es infinita. Sin embargo, esto no significa que 50,6 es pequeña en relación con un valor de referencia clínicamente significativo.
4. Una opción para un valor de referencia es 71,1; la tasa para los expuestos. La razón 50,6 / 71,1 = 0,72 indica que la tasa en el grupo con colesterol moderadamente alto se reduciría un 72% si este grupo pudiese hacer disminuir su colesterol a niveles normales.

P6.3

1. La diferencia de riesgos es 2,6% – 1,4% = 1,2%.
2. 1/0,012 = 83
3. Ya que calculamos el número necesario a tratar utilizando las incidencias acumuladas durante 10 años, la respuesta a la pregunta anterior indica que 83 pacientes deben tratarse durante 10 años para evitar un ACV.

P6.4

1. No. El riesgo absoluto describe el riesgo en un grupo específico, en lugar de la diferencia de riesgos de ambos grupos.
2. Sí
3. Sí
4. No. El riesgo relativo es la ***razón*** (en lugar de la ***diferencia***) de riesgos, o RR, entre dos grupos.
5. No. Para las preguntas 5 a 11: cualquier ubicación que no incluya a Queens podría utilizarse para una tasa de referencia o tasa esperada. Por lo tanto, Queens y la ciudad de Nueva York no serían buenas opciones, ya que ambas incluyen a Queens.
6. Sí
7. Sí
8. Sí
9. Sí
10. No
11. Sí

12. 15,7 por millón. Ya que las tasas individuales están en unidades *por millón,* la diferencia de las tasas tendrá la misma unidad de medición.
13. Diferencia de riesgos. En un brote como éste, los investigadores están comparando dos riesgos. La medida adecuada del impacto aquí es la diferencia de riesgos.
14. Diferencia de prevalencias. En un estudio de prevalencia, la medida apropiada de frecuencia de la enfermedad es la prevalencia. Una medida del impacto correspondiente es la diferencia de prevalencias.
15. Diferencia de tasas. En un estudio de seguimiento podemos utilizar personas-años de observación para calcular una tasa. La medida adecuada del impacto potencial aquí es la diferencia de tasas.

P6.5

1. Sí. Porque tanto mascar tabaco como consumir alcohol tienen el mismo valor (5) de la razón de riesgos.
2. No. Mascar tabaco tiene poco efecto en la salud pública, mientras que el consumo de alcohol tiene un efecto mucho más fuerte en la salud pública.
3. De 1.000 personas con mayor consumo de alcohol (es decir, alto consumo de alcohol), 40 personas desarrollarían la enfermedad debido a ese hábito. Por lo contrario, sólo 4 de 1.000 personas que mascan tabaco desarrollarían la enfermedad por dicho consumo.

P6.6

1. Falso. La diferencia de tasas entre Queens y el resto de la ciudad de Nueva York es 16,4 por millón – 2,4 por millón = 14,0 por millón. Por lo tanto, la tasa en exceso (es decir, la diferencia de tasas) en Queens es 14,0 casos por millón de la población.
2. Verdadero. Véase el punto anterior para esta respuesta.
3. Falso. 16,4 casos por millón no es la tasa atribuible (es decir, la diferencia de tasas), sino la tasa absoluta de encefalitis del Nilo Occidental en Queens. La diferencia de tasas es 16,4 – 2,4 = 14,0 por millón de la población.
4. Verdadero. La medida de efecto más común para comparar a Queens con el resto de la ciudad de Nueva York es 6,8 y ésta es la razón de tasas calculada como 16,4/2,4.
5. Diferencia de riesgos, una medida de la carga en la salud pública.
6. Razón de riesgos, una medida de la etiología de la enfermedad.
7. Diferencia de riesgos. El valor nulo de la diferencia de riesgos es 0; el valor nulo de la razón de riesgos es 1.
8. Diferencia de riesgos. La diferencia de riesgos puede ser negativa si el riesgo de referencia es mayor que el riesgo en el grupo de interés. La razón de riesgos nunca puede ser negativa porque es una razón de dos números positivos.
9. Ambas. Sí es *posible* tener una diferencia de riesgos y una razón de riesgos en el rango de 0,0 a 1,0. Observe que el CD dice que la respuesta correcta es Ninguna, con el fundamento de que ni la razón de riesgos ni la diferencia de riesgos están restringidas a valores entre 0,0 y 1,0.
10. Ambas, ya que el riesgo puede calcularse a partir de la mayoría de los estudios de seguimiento, entonces puede calcularse tanto la razón de riesgos como la diferencia de riesgos.
11. Ninguna, ya que el riesgo no puede calcularse a partir de los estudios de casos y controles, no puede calcularse ni la diferencia de riesgos ni la razón de riesgos.
12. Razón de riesgos. La razón de riesgos no tiene unidades, ya que es matemáticamente una división de cantidades (riesgos) que tienen las mismas unidades.
13. Razón de riesgos. Una razón de riesgos cercana a cero indicaría un efecto protector fuerte. Una diferencia de riesgos cercana a cero indicaría una ausencia de efecto.
14. Diferencia de riesgos.
15. Cáncer de pulmón. La razón de tasas es mucho más alta para el cáncer de pulmón que para la EC.
16. EC; la eliminación del tabaquismo eliminaría 190,4 casos por 100.000 de EC por año.

P6.7

1. 156 x 0,173 = 27, donde 0,173 es el riesgo para los individuos no expuestos.
2. 41 – 27 = 14 = I*
3. I*/I = 14/41 = 0,35 = FE

P6.8

1. (2,1 – 1) / 2,1 = 0,52
2. (RR – 1) / RR = (10 – 1) / 10 = 0,90
3. El valor máximo posible de la FE se acerca a 1 a medida que RR se acerca a infinito.
4. Sí, aún cuando la RR es muy grande, la FE puede ser pequeña, incluso cercana a cero, si la proporción expuesta en la población es muy pequeña.

P6.9

1. La afirmación **in**correcta es la opción D. Es errónea porque considera sólo las muertes entre las personas expuestas, es decir, aquellos con colesterol moderadamente alto. Una versión corregida de esta afirmación, que sería básicamente equivalente a la opción A, es que el 46% de todas las muertes se deben a las personas que poseen colesterol moderadamente alto.
2. La respuesta correcta es la opción A. Aquí se muestra el cálculo de este valor.

Exceso de Muertes	=	(Número total de muertes)	-	(Número esperado de muertes si todas las personas tuviesen nivel normal de colesterol)
	=	40	-	(104820 * 14 / 68239)
	=	40	-	21.5
	=	18.5		

P6.10

1. 75 x 0,19 = 14,25 = I*, donde 75 es la cantidad de individuos expuestos y 0,19 es la diferencia de riesgos.
2. En este ejemplo, 14 de las 27 muertes entre los fumadores que continuaron fumando podrían haberse evitado.
3. I^*/I_1 = 14,25 / 27 = 0,53 = FEe.
4. FEe = (2,1 –1) / 2,1 = 0,52. Este valor es igual al valor 0,53 obtenido anteriormente, si sacamos el error de redondeo.
5. La FE considera el impacto potencial de la exposición en "todos los casos" de la cohorte, mientras que la FEe se focaliza en el impacto potencial de la exposición sólo en los "casos expuestos" en la cohorte. Ambas medidas son relevantes, pero tienen un enfoque distinto.
6. FEe = $(DI_1 – DI_0) / DI_1$, donde DI_1 y DI_0 son las densidades de incidencia (es decir, tasas) para las personas expuestas y no expuestas en la cohorte.
7. FEe = (RDI – 1) / RDI
8. FE = p'(OR - 1)/[p'(OR – 1) + 1] y FEe = (OR – 1) / OR, donde OR es el odds ratio y p' es la proporción de todos los controles que están expuestos.

P6.11

1. La fracción etiológica es (1,4-1,2) / 1,4 = 14%.
2. p=0,108
   RR=(3/103)/(10/854)=2,5
   $$\mathrm{FE} = \frac{0{,}108(2{,}5-1)}{0{,}108(2{,}5-2)+1} = 14\%$$
3. FEe=(2,9-1,2)/2,9=60%

P6.12

1. 31,8% - la prevalencia del tabaquismo = 70 millones/220 millones = 31,8%.
2. 79,0% - La FE para el cáncer de pulmón es (31,8-6,7)/21,8=79,0%.
3. 26,7%
4. 92,2%
5. 53,3%
6. 92,2%
7. Fracción etiológica entre los expuestos
8. Fracción etiológica
9. Diferencia de riesgos/tasas
10. Fracción etiológica
11. Diferencia de riesgos/tasas
12. Diferencia de riesgos/tasas
13. Fracción etiológica, fracción etiológica entre los expuestos

P6.13

1. Multiplique la cantidad total de individuos del estudio, 11.185, por la incidencia acumulada para los no expuestos (105,9 por 1.000 personas) para obtener 1.184,5 casos potenciales.
2. 539, que puede verse en el cuadro o calcularse multiplicando 11.185 por la incidencia acumulada general de 48,2 por 1.000 personas. Este número es **I** en la fórmula inicial para FP.
3. 1.184,5 – 539 = 645,5 = I** casos totales fueron prevenidos.
4. FP = I** / (I** + I) = 645,5 / 1.184,5 = 0,55 ó 55%.

P6.14

1. Multiplique la cantidad total de niños inmunizados, 6.350, por la incidencia acumulada para los no expuestos (105,9 por 1.000 personas) para obtener 672,5 casos esperados de niños enfermos.
2. 27, que puede verse en el cuadro o calcularse multiplicando 6.350 por la incidencia acumulada para los niños inmunizados de 4,2 por 1.000 personas. Este número es I_1 en la fórmula inicial para FPe.
3. 672,5 – 27 = 645,5 = I** casos totales fueron prevenidos.
4. FPe = I** / (I** + I_1) = 645,5 / 672,5 = 0,96, o en porcentaje, 96%.

P6.15

1. FEe = (1 – 0,4) = 0,96. Este valor es igual al que obtuvimos mediante la fórmula que incluye IA_0 e IA.
2. La FP considera el impacto potencial de la exposición en “todos los casos” de la cohorte, mientras que la FPe se focaliza en el impacto potencial de la exposición sólo en los “casos expuestos” en la cohorte. Ambas medidas son relevantes, pero tienen un enfoque distinto.

3. FPe = (DI_0 – DI_1) / DI_1, donde DI_0 y DI_1 son las densidades de incidencia (es decir, tasas) para las personas no expuestas y expuestas, respectivamente, en la cohorte.
4. FPe = (1 – RDI)
5. FP = p'(1 – OR) y FPe = (1 – OR), donde OR es el odds ratio y p' es la proporción de todos los controles que están expuestos.

P6.16

1. Verdadero. Cuando la diferencia de riesgos es < 0, la exposición parece ser protectora y se prefiere la fracción prevenida.
2. Falso; 1-FE no es igual a la fracción prevenida. La diferencia en las dos medidas ocurre tanto en el numerador como en el denominador.
3. Verdadero. Una fórmula alternativa para la FP entre los expuestos en una cohorte o ensayo clínico es (1-RR).
4. Verdadero. La fracción prevenida es la proporción de casos que no ocurrieron y que habrían ocurrido si la exposición no hubiera estado presente.
5. Verdadero. El denominador tanto de FP como de FPe es IA_0 para los datos de riesgo o DI_0 para los datos de tasas.
6. Verdadero. Una fórmula alternativa para FP es FP=p(1-RR) donde p es la prevalencia de la exposición.
7. Fracción prevenida entre los expuestos – $\hat{\text{F}}\text{Pe} = \dfrac{\hat{\text{I}}\text{A}_0 - \hat{\text{I}}\text{A}_1}{\hat{\text{I}}\text{A}_0}$
8. Fracción prevenida - $\hat{\text{F}}\text{P} = \dfrac{\hat{\text{I}}\text{A}_0 - \hat{\text{I}}\text{A}}{\hat{\text{I}}\text{A}_0}$
9. Ninguna de las dos. Una fórmula alternativa para FP es FP=p(1-RR) donde p es la prevalencia de la exposición. Si se hubiese vacunado a menos niños, la FP disminuiría.
10. Fracción prevenida. Una fórmula alternativa para FP es FP=p(1-RR) donde p es la prevalencia de la exposición. Si se hubiese vacunado a menos niños, la FP disminuiría.
11. Fracción prevenida. La FP es la proporción de casos potenciales en la comunidad prevenidos por la vacunación.

LECCIÓN 7

VALIDEZ

7-1 Validez

*El objetivo principal de la mayoría de las investigaciones epidemiológicas es obtener un **cálculo válido** de una **medida de efecto** de interés. En esta actividad ilustramos tres tipos generales de problemas de **validez**. También distinguimos la validez de la **precisión,** presentamos el término **sesgo** y discutimos cómo efectuar el ajuste por sesgo.*

Ejemplos de problemas de validez

La validez en los estudios epidemiológicos se ocupa de las fallas metodológicas que pueden derivar en conclusiones falsas sobre la relación entre la exposición y la enfermedad. Se describen brevemente varios ejemplos de problemas de validez.

La validez de un estudio epidemiológico se relaciona con la existencia de alguna falla en el diseño del estudio o en los métodos de recolección o de análisis de datos que pudiera derivar en conclusiones falsas sobre la relación entre la exposición y la enfermedad. Si no existen tales fallas, decimos que el estudio es **válido**. Si existen fallas, la medida en que los resultados difieren de las conclusiones correctas se llama **sesgo**. Lo que se busca es la validez de un estudio; el sesgo es lo que no nos permite obtener resultados válidos.

En 1946, Berkson demostró que los estudios de casos y controles llevados a cabo exclusivamente en hospitales están sujetos a un tipo de **sesgo de "selección"**, acertadamente denominado **sesgo de Berkson**. El sesgo de Berkson surge porque los pacientes con dos enfermedades o comportamientos de alto riesgo tienen más probabilidades de ser hospitalizados que los que tienen una única enfermedad. Dichos pacientes tenderán a estar sobrerrepresentados en la población de estudio, en comparación con la población de la comunidad. En particular, se ha demostrado que existe una asociación entre las enfermedades respiratorias y las enfermedades óseas en los pacientes internados, pero no en la población en general. Asimismo, debido a que el tabaquismo está fuertemente asociado a las enfermedades respiratorias, esperaríamos que un estudio llevado a cabo en hospitales sobre la relación entre el tabaquismo y las enfermedades óseas demuestre dicha relación, incluso si ésta no se da en la población en general.

En las décadas de los ochenta y noventa, los investigadores de las fuerzas aéreas estadounidenses evaluaron los efectos asociados con la exposición al herbicida Agente Naranja en la salud de los veteranos de la guerra de Vietnam. El Agente Naranja contenía vestigios de un contaminante altamente tóxico conocido como TCDD. Inicialmente, la exposición al TCDD se clasificó según las descripciones de las funciones de los veteranos seleccionados para el estudio. Luego se determinó que esto generaba clasificaciones erróneas sustanciales del TCDD. El problema de validez en este caso se llama **sesgo de información**. El sesgo podría evitarse utilizando técnicas de laboratorio desarrolladas para medir el TCDD en el suero sanguíneo. El uso de dichos **marcadores biológicos** en la investigación epidemiológica está aumentando rápidamente como medio para reducir la clasificación errónea y, en forma más general, para mejorar la precisión de las mediciones de los estudios.

Como ejemplo final, volvamos al Estudio de usuarios de las playas de Sydney descrito anteriormente. El

problema de validez en este estudio radicaba en si se habían tenido en cuenta todas las variables relevantes, aparte de la condición de bañista y los niveles de contaminación. Dichas variables incluían la edad, el sexo, el tiempo durante el cual la persona se había bañado y los días adicionales en que volvió a hacerlo. La razón principal de tener en cuenta estas variables adicionales es garantizar que ninguno de los efectos observados en el resultado de salud por el hecho de bañarse pudiera explicarse por medio de estas otras variables. Una distorsión en los resultados por no considerar dichas variables adicionales se llama **sesgo de confusión**.

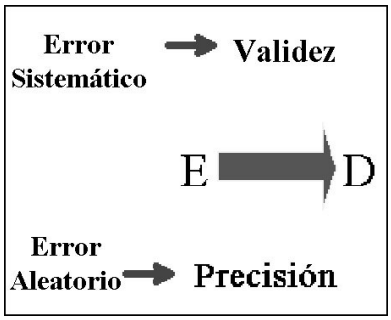
Resumen

- Validez: problema general de si existe alguna falla en el diseño del estudio o en los métodos de recolección o de análisis de datos que pueda derivar en conclusiones falsas sobre la relación entre la exposición y la enfermedad.
- Sesgo: medida de la distorsión de las conclusiones sobre la relación entre la exposición y la enfermedad.
- Los problemas de validez están ejemplificados mediante:
  - Estudios de casos y controles basados en hospitales (**sesgo de selección** de Berkson).
  - Clasificaciones erróneas de funciones laborales para evaluar la exposición al TCDD (**sesgo de información**).
  - Control de variables relevantes en el Estudio de usuarios de las playas de Sydney (**sesgo de confusión**).

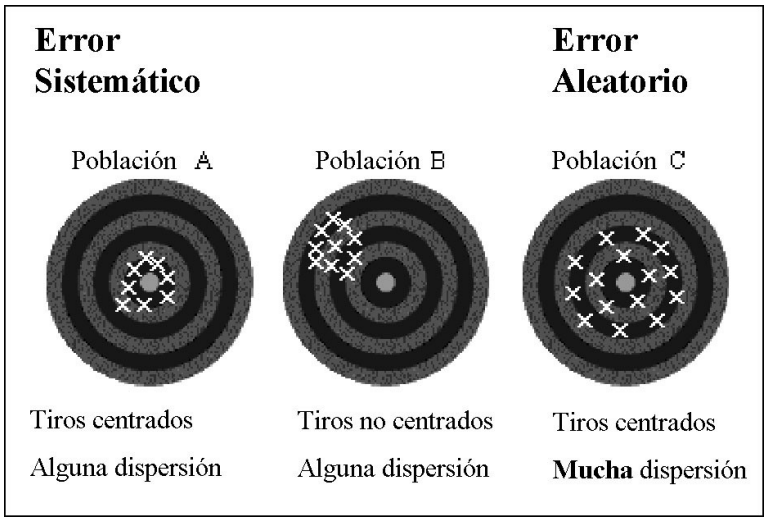
Validez versus precisión

*La validez y la precisión se refieren a dos fuentes diferentes de la inexactitud que puede ocurrir cuando se calcula la relación entre la exposición y la enfermedad: el **error sistemático** (problema de validez) y el **error aleatorio** (problema de precisión). El error sistemático y el error aleatorio pueden distinguirse en términos de tiros al blanco.*

La validez y la precisión están influenciadas por dos tipos diferentes de error que pueden presentarse cuando se calcula la relación entre la exposición y la enfermedad. El **error sistemático** afecta la **validez** y el **error aleatorio** afecta la **precisión**.

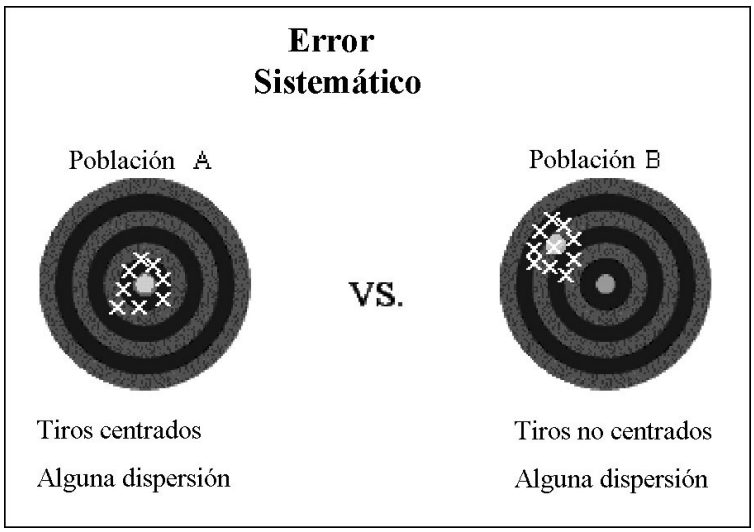


Estos dos tipos de error pueden distinguirse viendo un estudio epidemiológico como un tiro al blanco. El punto azul en el medio del blanco simboliza la **medida verdadera de efecto** que se calcula en una población de interés. (Nota: para ser coherentes con el CD, el uso del término "punto azul" en este texto se refiere al centro del blanco o diana). Cada tiro representa un **cálculo** del verdadero efecto obtenido de uno de quizás muchos estudios en cada una de las tres poblaciones.



Para el Blanco A, los tiros se centran alrededor del punto azul, aunque ninguno de ellos, en verdad, lo tocaron, sino que todos impactaron en una parte distinta del blanco. Para el Blanco B, los tiros dieron todos lejos del centro, pero tienen aproximadamente la misma dispersión que los tiros del Blanco A. Para el Blanco C, los tiros se centran alrededor del punto azul, pero a diferencia del Blanco A, están más dispersos uno del otro.

El error sistemático se ejemplifica comparando el Blanco A con el Blanco B. Los tiros al Blanco A están apuntados al punto azul, mientras que los tiros al Blanco B no lo están, más bien, se encuentran centrados alrededor del punto rojo. (Nota: para ser coherentes con el CD, el uso del término "punto rojo" se referirá al punto que se encuentra arriba y a la izquierda de la diana en la Población B). La distancia entre el punto azul y el punto rojo mide el error sistemático asociado al Blanco B. Por el contrario, no hay error sistemático asociado al Blanco A.



El error sistemático se verifica cuando hay una diferencia entre la medida verdadera de efecto y lo que se calcula realmente. Decimos que el estudio es válido cuando no hay error sistemático. Por lo tanto, la validez se refiere a si un estudio apunta a la medida de efecto correcta, representada por la diana. Lamentablemente, en los

estudios epidemiológicos y de otro tipo generalmente se desconoce la diana. En consecuencia, es difícil determinar el sesgo y su evaluación es siempre subjetiva, en cierta medida.

Todos los blancos ilustran el error aleatorio, que ocurre cuando hay una diferencia entre un valor calculado a partir de los datos del estudio y la medida de efecto que, en realidad, se calcula. Los Blancos A y B muestran la misma cantidad de error aleatorio porque, básicamente, existe la misma dispersión de tiros alrededor del punto azul del Blanco A que alrededor del punto rojo del Blanco B. Por el contrario, el Blanco C, en el que los tiros están mucho más dispersos, muestra más error aleatorio que los Blancos A o B.

Por lo tanto, cuanto más dispersos estén los tiros, mayor será el error aleatorio y menor será la precisión de cualquiera de aquéllos. Por consiguiente, la precisión se refiere a cuánta variación individual haya entre tiro y tiro, dado el lugar real al que se está apuntando. En otras palabras, la precisión refleja la **variabilidad de muestreo**.

Los problemas de precisión generalmente se refieren a la inferencia estadística sobre los parámetros de la población a la que en realidad se apunta. Por el contrario, los problemas de validez se refieren a las fallas metodológicas del diseño del estudio o el análisis que pueden influir en si el estudio está apuntando al parámetro de población correcto, representado por el punto azul en cada blanco.

Preguntas de estudio (P7.1)

Considere un estudio transversal para evaluar la relación entre la ingesta de calcio (alta versus baja) en la dieta de las personas y la prevalencia de artritis de cadera en las mujeres de entre 45 y 69 años que viven en la ciudad de Atlanta. Se selecciona una muestra de pacientes mujeres en hospitales a partir de los registros hospitalarios de 1989 y se determina la presencia o ausencia de artritis en cada una de ellas, así como una medida de ingesta promedio de calcio en su dieta antes de ingresar al hospital.

1. ¿Cuál es la población objetivo o meta de este estudio?
2. ¿Qué representa el centro del blanco (es decir, la diana) en términos epidemiológicos?
3. ¿Qué significa error aleatorio en relación con este estudio?
4. ¿Qué significa error sistemático en relación con este estudio?

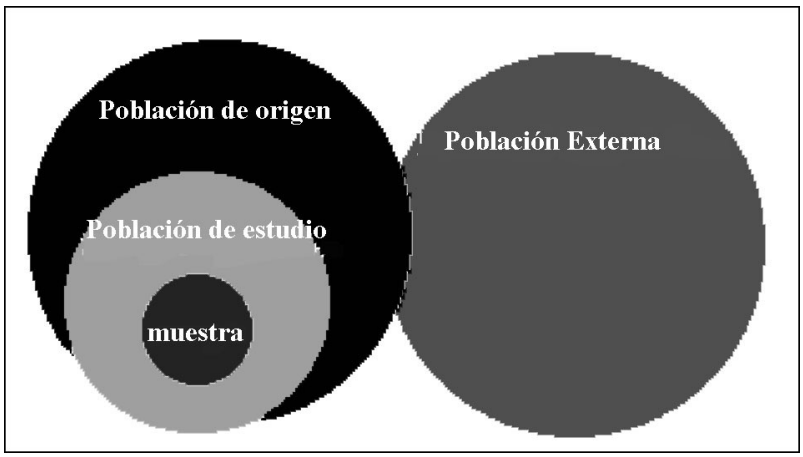
Resumen

- La validez se refiere al error sistemático, mientras que la precisión se refiere al error aleatorio.
- El error sistemático y el error aleatorio pueden distinguirse en términos de tiros al blanco.
- Error sistemático: diferencia entre lo que un estimador está realmente calculando y la medida de efecto de interés.
- Error aleatorio: diferencia entre cualquier valor calculado a partir de los datos del estudio y la medida de efecto que en verdad se está calculando.
- La validez no considera la inferencia estadística, sino las fallas metodológicas del diseño del estudio o análisis.
- La precisión se refiere a la inferencia estadística sobre los parámetros de la población a la que en realidad se apunta.

Jerarquía de poblaciones

Para clarificar aún más la diferencia entre validez y precisión, ahora describimos la jerarquía de poblaciones que se consideran en cualquier estudio epidemiológico.

Normalmente, identificamos poblaciones diferentes cuando pensamos en la validez de un estudio epidemiológico. Estas poblaciones pueden estar contenidas dentro de las demás o simplemente pueden superponerse.

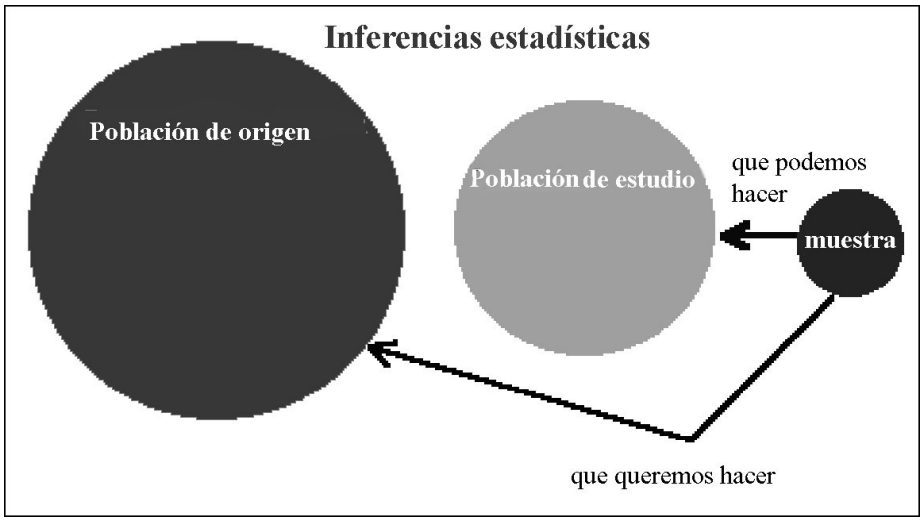


Llamamos **muestra** al conjunto de individuos de los cuales se han obtenido los datos del estudio. Utilizamos los resultados de la muestra para hacer inferencias sobre poblaciones más grandes. ¿Pero sobre qué poblaciones podemos hacer estas inferencias? ¿Qué población representa la muestra?

La **población de estudio** es el conjunto de individuos que representa realmente nuestra muestra e incluye normalmente a aquellos individuos a los cuales nos resulta posible estudiar. Podemos estar limitados a realizar muestreos en hospitales o en lugares o momentos específicos. La población de estudio se define según lo que sea práctico, que puede no ser lo que idealmente nos gustaría.

La **población de origen** es el conjunto de individuos de interés limitado, por ejemplo, en una ciudad, comunidad u ocupación específica, que están en riesgo de ser un caso. Claramente todos los casos deben provenir de la población de origen (si no estuviesen en riesgo, no se habrían convertido en casos). También es probable que la población de origen incluya individuos que, aunque estén en riesgo, pueden no convertirse en casos. La población de origen se ha denominado **base del estudio** o **población objetivo (o meta)**.

Podemos hacer inferencias estadísticas a partir de la muestra de la población de estudio, pero nos gustaría poder hacer inferencias a partir de la muestra de la población de origen. Lamentablemente, la **población de estudio**, que en verdad está representada por nuestra muestra, puede no ser representativa de la **población de origen**.



Preguntas de estudio (P7.2)

Considere un estudio epidemiológico realizado en la ciudad de Nueva York (NY) para evaluar si la obesidad está asociada a la hipertensión en los adultos jóvenes. Los investigadores decidieron que no era viable considerar la toma de una muestra de entre todos los adultos jóvenes de la ciudad. Se decidió que los gimnasios brindarían una gran fuente de adultos jóvenes de la ciudad de NY. Se toma una muestra de individuos en varios gimnasios de toda la ciudad seleccionados en forma aleatoria y se toma su presión arterial para determinar la condición de hipertensión.

1. ¿Cuál es la población de origen en este estudio?
2. ¿Cuál es la población de estudio en este estudio?
3. ¿La población de estudio está representada por la muestra?
4. ¿La población de origen está representada por la población de estudio?

En un caso simple, cada miembro de la población de estudio también está en la población de origen, es decir que sólo estamos estudiando individuos que en verdad están en riesgo. Si la población de estudio es representativa de la población de origen y la muestra es representativa de la población de estudio, entonces no hay sesgo al hacer inferencias de la muestra a la población de origen.



Preguntas de estudio (P7.2) (continuación)

Recuerde el estudio epidemiológico realizado en la ciudad de Nueva York para evaluar si la obesidad está asociada con la hipertensión en los adultos jóvenes. Suponga que los investigadores decidieron que era importante tomar una muestra de entre todos los adultos jóvenes de la ciudad de NY. Utilizaron la información del censo del año 2000 para obtener un listado de todos los adultos jóvenes de la ciudad de NY. Se toma una muestra de individuos de varias manzanas de la ciudad seleccionadas en forma aleatoria y se toma su presión arterial para determinar la condición de hipertensión.

5. ¿Cuál es la población de origen en este estudio?
6. ¿Cuál es la población de estudio en este estudio?
7. ¿La población de estudio está representada por la muestra?
8. ¿La población de origen está representada por la población de estudio?

A veces, inclusive una muestra tomada de una población de estudio que representa a la población de origen puede volverse sesgada. Por ejemplo, en un típico estudio de cohorte, aunque todos los miembros de la muestra de la población de estudio inicial también sean miembros de la población de origen, la muestra inicial puede cambiar durante el curso de la investigación. Dicha muestra inicial puede sufrir exclusiones, abandonos, falta de respuestas o pérdida de seguimiento. De esta forma, la población de estudio final puede resultar incluyendo sólo a aquellos individuos que deseen y puedan quedarse en dicho estudio y no ser representativa de la población de estudio inicial o de la población de origen como hubiésemos deseado.

Preguntas de estudio (P7.2) (continuación)

9. Verdadero o falso. En un estudio de cohorte, la población de ***estudio*** incluye a las personas que reúnen los requisitos para ser elegidas, pero cuyo seguimiento se habría perdido si en verdad se hubiese realizado.

10. Verdadero o falso. En un estudio de cohorte, la población de ***origen*** incluye a las personas que reúnen los requisitos para ser elegidas, pero cuyo seguimiento se habría perdido si en verdad se hubiese realizado.
11. Verdadero o falso. En un estudio de casos y controles, la población de estudio puede incluir a personas que reúnen los requisitos para ser controles y que no estaban en riesgo de ser un caso.
12. Verdadero o falso. En un estudio transversal, la población de estudio puede incluir a personas que desarrollaron el resultado de salud, pero que habían fallecido antes del momento en que se seleccionó la muestra.

En general, es posible que los miembros de la población de estudio ni siquiera estén en riesgo y no podamos distinguirlo. Por ejemplo, podemos obtener nuestra muestra de una población de estudio compuesta por personas que asisten a una clínica para enfermedades de transmisión sexual. Estas personas pueden padecer o no una enfermedad de transmisión sexual o pueden estar exponiéndose o no a una afección de este tipo. Por ejemplo, algunos de estos sujetos pueden tener parejas que no están infectadas y, por lo tanto, pueden no estar en riesgo. Si no están en riesgo, no son parte de la población de origen, pero quizás no lo sepamos.

También puede suceder que la población de estudio no incluya a individuos que están en riesgo (y, por lo tanto, sean parte de la población de origen), ya sea porque no sabemos que están en riesgo o porque no es práctico contactarlos. Por ejemplo, cuando poco se sabía sobre el SIDA, es posible que un estudio sobre hombres homosexuales en riesgo de padecer dicha enfermedad no haya incluido a los usuarios de drogas intravenosas, que también estaban en riesgo.

Finalmente, muchas veces nos gustaría generalizar nuestras conclusiones a una **población externa** diferente. Se entiende por población externa a aquélla que difiere de la población de estudio pero que con respecto a la cual, de todas maneras, nos gustaría generalizar los resultados (p. ej., una ciudad, comunidad u ocupación diferente). En un entorno de salud pública, siempre nos preocupa la salud del público en general, aunque debamos estudiar subpoblaciones más pequeñas por razones prácticas. A fin de obtener conclusiones estadísticas que se basen en una muestra para generalizarlas a una población externa, la población de estudio debe ser representativa de la población externa, aunque a menudo ésto es difícil de lograr.

Preguntas de estudio (P7.2) (continuación)

Considere un estudio epidemiológico realizado en la ciudad de Nueva York para evaluar si la obesidad está asociada a la hipertensión en los adultos jóvenes. Suponga que fue de interés determinar si los resultados del estudio se podían generalizar a la población de todo el estado de Nueva York.

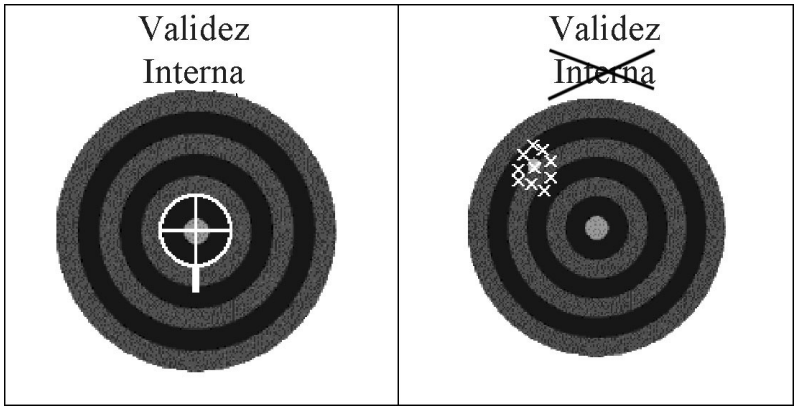
13. Considere la variedad de las poblaciones descritas, ¿qué tipo de población se está considerando? Desarrolle su respuesta brevemente.

Resumen

- En cualquier estudio epidemiológico hay una variedad de poblaciones a considerar.
- La **muestra** es el conjunto de individuos de los cuales se han obtenido los datos del estudio.
- La **población de estudio** es el conjunto de individuos que representa en verdad nuestra muestra e incluye normalmente a aquellos individuos a los cuales nos es factible estudiar.
- La **población de origen** es el grupo de interés limitado sobre el cual los investigadores desean evaluar una relación entre la exposición y la enfermedad.
- La **población externa** es un grupo al que no se ha restringido el estudio pero al cual, sin embargo, el investigador desea generalizar los resultados del estudio.

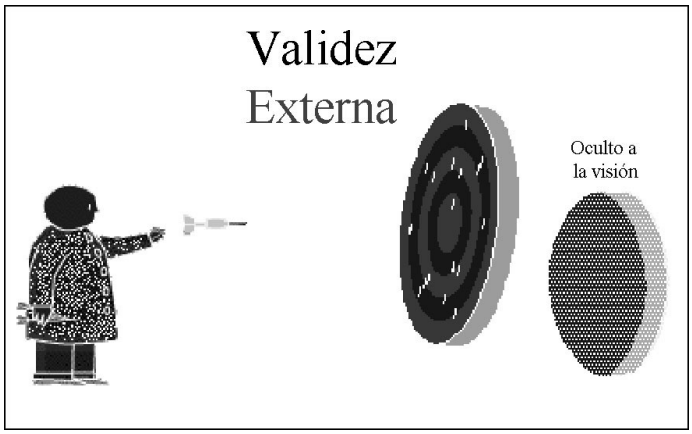
Validez interna versus externa

El tiro al blanco brinda un ejemplo que ilustra la diferencia entre la **validez interna** y **externa**. La validez interna considera si estamos apuntando al centro del blanco. Si apuntamos a este punto rojo (a la izquierda y arriba de la diana) en lugar de a la diana, entonces nuestro estudio **no** es internamente válido.



La **validez interna** implica sacar conclusiones acerca de la población de origen sobre la base de la información de la población de estudio. Dichas inferencias no se extienden más allá de la población de origen de interés limitado.

La **validez externa** se refiere a un blanco distinto; en especial, uno al que no pretendemos tirar, cuya diana, en verdad, no podemos ver. Podríamos imaginarnos este blanco externo como oculto a nuestra visión.



Suponga que este blanco oculto está alineado con el blanco al que le estamos tirando. Entonces, al apuntar a la diana del blanco que podemos ver, también le estamos apuntando a la diana del blanco externo. En este caso, los resultados de nuestra población de estudio pueden generalizarse a esta población externa y, por lo tanto, tenemos validez externa. Si el blanco externo no está alineado con nuestro blanco, nuestro estudio no tiene validez externa, y sus resultados no deberían extenderse a esta población externa.

La validez externa involucra aplicar nuestras conclusiones a una población externa más allá del interés limitado del estudio. Dichas inferencias requieren juicios sobre otros resultados y su conexión con los resultados del estudio, la conceptualización del proceso de enfermedad y los procesos biológicos relacionados, al igual que las características comparativas de la población de origen y de la población externa. La validez externa, por lo tanto, es más subjetiva y menos cuantificable que la validez interna.

Preguntas de estudio (P7.3)

Considere un estudio epidemiológico realizado en la ciudad de Nueva York para evaluar si la obesidad está asociada a la hipertensión en los adultos jóvenes. Se obtiene una muestra de individuos en varios gimnasios de toda la ciudad y se toma su presión arterial para determinar la condición de hipertensión.

1. ¿Qué se requiere para que este estudio sea internamente válido?

Suponga que fue de interés determinar si los resultados del estudio se podían generalizar a todo el estado de Nueva York.

2. ¿Esto se refiere a la validez interna o a la validez externa? Desarrolle su respuesta brevemente.

Los resultados del Ensayo de prevención primaria realizado por las Clínicas de investigación de lípidos y publicado en 1984 (JAMA, vol. 251) demostraron una reducción significativa de la mortalidad cardiovascular en hombres blancos de 35 a 59 años a los que se les hizo seguir una dieta de reducción del colesterol y se les suministró medicación.

3. ¿Qué pregunta relacionada con la validez externa podría formularse sobre los resultados de este estudio?
4. ¿Qué pregunta(s) referida(s) a la validez interna podría(n) formularse sobre los resultados de este estudio?

Resumen

- La validez interna se refiere a si estamos apuntando al centro del blanco al que sabemos que estamos tirando.
- La validez externa se refiere a un blanco al que no pretendemos tirar, cuya diana, en verdad, no podemos ver.
- La validez interna se refiere a sacar conclusiones sobre la población objetivo o meta en base a la información de la población de estudio.
- La validez externa se refiere a aplicar conclusiones a una población externa más allá del interés limitado del estudio.

Auto-evaluación (P7.4)

Coloque **verdadero** o **falso** en cada una de las siguientes afirmaciones; para las preguntas 8 a 11, una opción adicional es **depende**.

1. El error aleatorio ocurre cuando existe alguna diferencia (no cero) entre el valor del *odds ratio* en la población de estudio y el cálculo del *odds ratio* obtenido de la muestra analizada. . . **???**

2. El error sistemático se presenta cuando existe alguna diferencia (no cero) entre el valor de la medida de efecto en la población de origen y el cálculo a partir de la muestra. . . . **???**

3. En un estudio válido, no hay ni error sistemático ni aleatorio. . . . . **???**

4. La población de estudio siempre es un subconjunto de la población de origen. . . **???**

5. La muestra siempre es un subconjunto de la población de estudio. . . **???**

6. La muestra siempre es un subconjunto de la población de origen. . . . **???**

7. La medida de efecto calculada en la muestra siempre es igual a la medida de efecto correspondiente en la población de estudio. . . . . **???**

8. Suponga que la razón de riesgos en la población de origen es 3,0, mientras que el cálculo de la razón de riesgos en la muestra (de estudio) es 1,2. Entonces, el estudio no es internamente válido. . **???**

9. Suponga que la razón de riesgos en la población de origen es 3,0, mientras que la razón de riesgos en la población de estudio es 1,2. Entonces, el estudio no es internamente válido. . . **???**

10. Suponga que la razón de riesgos en la población de origen es 3,0, mientras que el cálculo de la razón de riesgos en la población de estudio es 3,1. Entonces, el estudio no es internamente válido. . **???**

11. Suponga que la razón de riesgos en la población de origen es 3,0, mientras que el cálculo de la razón de riesgos en la población de estudio es 3,1. Entonces, el estudio no es externamente válido. . **???**

7-2 Validez

Definición cuantitativa del sesgo

*El **sesgo** en un estudio epidemiológico puede definirse cuantitativamente en términos del **parámetro objetivo** de interés y la medida de efecto que realmente se está calculando en la **población de estudio**.*

Se dice que un estudio tiene sesgo cuando no es internamente válido. Cuantifiquemos lo que queremos decir por sesgo. La medida de efecto en la población de origen es nuestro parámetro objetivo.

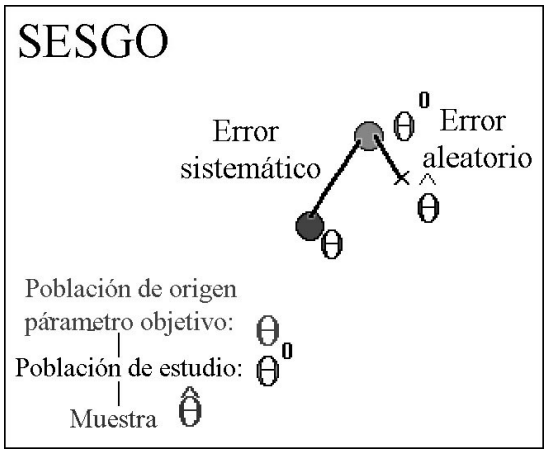


La elección de este parámetro depende de las características del diseño del estudio, sus objetivos y el tipo de sesgo que se está considerando. Denotamos el parámetro objetivo con la letra griega **θ** (Theta). Queremos calcular el valor de **θ** en la población de origen.

Recordando la jerarquía de las poblaciones asociadas a un estudio dado, indicamos como $\boldsymbol{\theta^0}$ la medida de efecto en la población de estudio. La estimación $\hat{\theta}$ (theta con acento circunflejo) denota nuestra medida de efecto calculada a partir de la muestra que en verdad se está analizando. Por supuesto, $\hat{\theta}$, $\boldsymbol{\theta^0}$ y **θ** pueden tener todas valores diferentes.

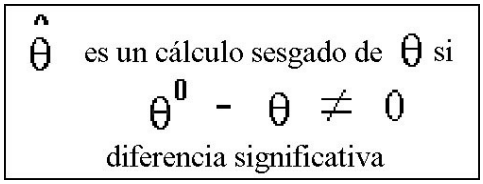


Cualquier diferencia entre $\hat{\theta}$ y $\boldsymbol{\theta^0}$ es generada por un error aleatorio. Cualquier diferencia entre $\boldsymbol{\theta^0}$ y **θ** es producida por un error sistemático.



Utilizamos $\hat{\theta}$ para calcular **θ**. Decimos que $\hat{\theta}$ es un cálculo sesgado de **θ** si θ^0 no es igual a **θ**, y definimos el sesgo como la diferencia entre estos dos parámetros: sesgo ($\hat{\theta}$, **θ) = θ^0 - θ**

Por lo tanto, el sesgo ocurre si una medida de efecto para la población de estudio difiere sistemáticamente del valor del parámetro objetivo. El signo *desigual* que se muestra a continuación no debería interpretarse estrictamente como cualquier diferencia de cero, sino como una diferencia relevante de cero. Es necesaria dicha interpretación flexible de la definición porque el sesgo raramente se puede cuantificar, precisamente porque el parámetro objetivo siempre es desconocido.



Preguntas de estudio (P7.5)

Considere un estudio de cohorte para evaluar si el mayor consumo de alcohol durante el embarazo genera bebés con bajo peso al nacer. Aunque es normalmente desconocida, suponga que la razón de riesgos en la población de origen (es decir, la cohorte completa) es 3,5. Suponga, además, que la muestra de estudio es representativa de la población de origen al comienzo del seguimiento, pero que hay una migración considerable hacia afuera de la ubicación de la población objetivo o meta. Como resultado, la razón de riesgos encontrada en la muestra final es de 1,5.

1. ¿Cuáles son los valores del parámetro de la población objetivo (es decir, de origen o meta) y del parámetro de la población de estudio, en este escenario?
2. ¿Hay sesgo en este estudio? Desarrolle su respuesta brevemente.
3. ¿Cuál es el valor del sesgo en este estudio?
4. Si el efecto verdadero que se está calculando es *alto* (por ejemplo, RR>3,5), pero los datos del estudio básicamente no muestran efecto, ¿esto indica un sesgo? Desarrolle su respuesta brevemente.
5. Si el efecto verdadero calculado indica que *no* hay asociación, pero los datos del estudio muestran una asociación alta, ¿esto indica un sesgo? Desarrolle su respuesta brevemente.

Resumen

- El sesgo mide la medida en que los resultados del estudio están distorsionados con respecto a los resultados correctos que se hubieran encontrado en un estudio válido.
- El sesgo ocurre si una medida de efecto para la población de estudio difiere sistemáticamente del valor del parámetro objetivo.
- El sesgo raramente se puede cuantificar, principalmente porque el parámetro objetivo siempre es desconocido.
- **θ** (Theta) es el parámetro objetivo, $\boldsymbol{\theta^0}$ es el parámetro de la población de estudio y $\hat{\theta}$ es el cálculo de la muestra.
- $\hat{\theta}$ es un cálculo sesgado de **θ,** siempre que $\boldsymbol{\theta^0}$ no sea igual a **θ**.
- Sesgo ($\hat{\theta}$, **θ**) = $\boldsymbol{\theta^0 - \theta}$

Relación entre el parámetro objetivo y el tipo de sesgo

El parámetro objetivo en la mayoría de los estudios epidemiológicos es, normalmente, una medida de efecto, por ejemplo, algún tipo de razón de riesgos, *odds ratio* o razón de tasas, apropiada para el diseño del estudio que se está considerando. La elección del parámetro depende de si el tipo de sesgo de interés es un sesgo de selección, de información o de confusión, o si más de uno de estos tres tipos de sesgo resulta de interés.

Cuando el **sesgo de selección** es el único tipo de sesgo que se está considerando, el parámetro objetivo normalmente es el valor de la medida de efecto de interés en la población de origen de donde se derivan los casos.

Sin embargo, si el **sesgo de información** es el único tipo de sesgo considerado, entonces el parámetro objetivo es la medida de efecto que corrige el posible error de clasificación o que resultaría de la ausencia de dicho error.

Si sólo hay sesgo debido a la **confusión**, entonces el parámetro objetivo es la medida de efecto calculada cuando se controla la confusión.

Si es posible tener más de un sesgo, entonces el parámetro objetivo es el valor de la medida de efecto después de la corrección de todas las fuentes que aportan sesgo.

Sesgo epidemiológico versus sesgo estadístico

La definición de sesgo que hemos dado utilizando la expresión sesgo = $\boldsymbol{\theta^0 - \theta}$ identifica implícitamente a $\hat{\theta}$ como un **estimador estadísticamente coherente**, es decir, $\hat{\theta}$ converge a **θ** en probabilidad a medida que se aumenta hipotéticamente el tamaño del estudio sin cambiar ninguna característica esencial del diseño del estudio en sí mismo.

Se recomienda esta definición para describir el **sesgo epidemiológico** porque la coherencia estadística básicamente refleja sólo aspectos metodológicos del diseño del estudio que podrían estar asociados con el error sistemático en lugar del error aleatorio.

Por el contrario, una definición de **sesgo estadístico** estaría basada en el concepto del **valor estadístico esperado**, es decir, sesgo = valor esperado de $\hat{\theta} - \theta$, donde el "valor esperado" es el promedio a largo plazo de $\hat{\theta}$ a medida que aumenta ilimitadamente el tamaño de la muestra.

Utilizando el **sesgo estadístico**, permitiríamos que el sesgo estuviese presente sólo debido a las propiedades matemáticas del estimador. Por ejemplo, el valor esperado de un cálculo de la razón de riesgos no es igual a la razón de riesgos de la población que, en verdad, está calculando, incluso si no hay fallas metodológicas. Esto no es lo que se supone que describe el **sesgo epidemiológico** y, entonces, no es deseable el uso de la teoría del valor

esperado para describir el sesgo epidemiológico.

Dirección del sesgo

Aunque nunca puede cuantificarse realmente la magnitud precisa del sesgo, a menudo puede determinarse la ***dirección del sesgo****. Ésta indica si el parámetro objetivo está sobrestimado o subestimado sin especificar la magnitud del sesgo.*

Hemos definido al sesgo como la diferencia entre el valor de la medida de efecto en nuestra población objetivo (es decir, de origen o meta) y el valor de la medida de efecto que, en verdad, se está calculando en la población de estudio. Ya que el parámetro objetivo siempre es desconocido y el efecto que se está calculando en la población de estudio posee error aleatorio, es virtualmente imposible cuantificar con precisión la magnitud del sesgo en un estudio epidemiológico dado. Sin embargo, el investigador puede, a menudo, determinar la dirección del sesgo. Dicha evaluación requiere normalmente un juicio subjetivo basado en el conocimiento del investigador sobre las variables que se están estudiando y las características del diseño del estudio que son las fuentes de un posible sesgo. Se entiende por dirección a la determinación de si el parámetro objetivo está **sobrestimado** o **subestimado** sin especificar la magnitud del sesgo. Si el parámetro objetivo está **sobrestimado**, decimos que la dirección del sesgo **se aleja del valor nulo**. Si el parámetro objetivo está **subestimado**, decimos que la dirección del sesgo es **hacia el valor nulo**.

$$\text{SESGO}(\hat{\theta},\theta) = \theta^0 - \theta$$

Dirección del sesgo

$\theta^0 >^{\text{objetivo}} \theta$ Sobrestimado → se aleja del valor nulo

$\theta^0 < \theta$ Subestimado → hacia el valor nulo

Por ejemplo, suponga que el parámetro objetivo es una razón de riesgos cuyo valor es 1,5, pero la razón de riesgos que, en verdad, estimada a partir de nuestro estudio es 4. Entonces, el verdadero efecto ha sido sobrestimado, ya que el efecto del estudio parece ser más fuerte de lo que es en realidad. Tanto el parámetro objetivo como el parámetro de la población de estudio en este ejemplo son mayores que el valor nulo de 1 para una razón de riesgos. Por lo tanto, el sesgo se aleja del valor nulo, ya que el valor incorrecto de 4 está más alejado de 1 que el valor correcto de 1,5.

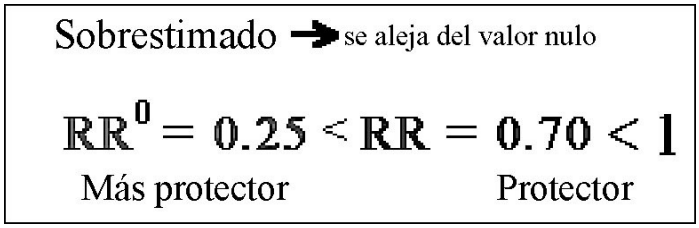
Sobrestimado → se aleja del valor nulo

$$RR^0 = 4 > RR = 1.5 > 1$$

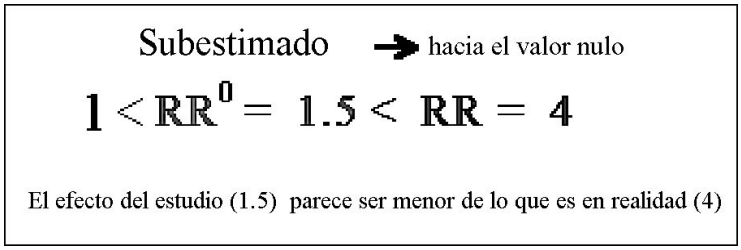
valor nulo

En forma similar, si el parámetro objetivo es la razón de riesgos cuyo valor es menor que el valor nulo de 1 (p. ej., 0,70) y la razón de riesgos estimada en nuestro estudio es 0,25, el verdadero efecto también está sobrestimado. En este caso, el verdadero efecto es protector, ya que es menor que 1, y el efecto calculado de 0,25

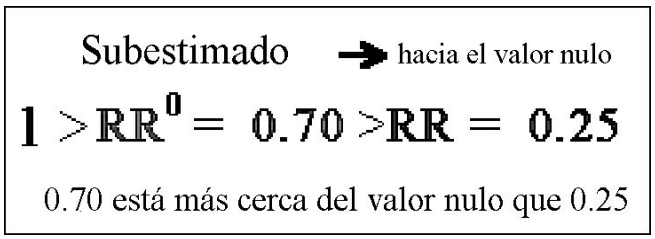
es aún más protector que el efecto real. Nuevamente, el valor incorrecto de 0,25 está más alejado de 1 que el valor correcto de 0,70; por lo tanto, el sesgo se aleja del valor nulo.



Para describir la **subestimación**, o **sesgo hacia el valor nulo**, suponga que el objetivo es una razón de riesgos igual a 4, pero la razón de riesgos estimada es 1,5. Entonces, el verdadero efecto está subestimado, ya que el efecto del estudio parece ser menor de lo que es en realidad. Asimismo, el valor incorrecto de 1,5 está más cercano al valor nulo que el valor correcto de 4; por lo tanto, el sesgo es **hacia el valor nulo**.



Si el parámetro objetivo es una razón de riesgos igual a 0,25, pero la razón de riesgos estimada es 0,70, entonces, nuevamente, el valor correcto está subestimado y el sesgo es hacia el valor nulo. En este caso, el valor incorrecto de 0,70 está más cercano al valor nulo que el valor correcto de 0,25.



Suponga, sin embargo, que la razón de riesgos de objetivo es 0,50, pero la razón de riesgos que, en verdad, se está calculando es 2. Estos dos valores están en lados opuestos del valor nulo, por lo que no podemos sostener que el sesgo es hacia el valor nulo o se aleja de éste. En este caso, denominamos a este tipo **sesgo de transición**. En otras palabras, un sesgo de cambio puede ocurrir si en los datos la exposición parece tener un efecto dañino en la enfermedad cuando, en verdad, es protectora. En forma alternativa, un sesgo de cambio puede ocurrir si la exposición parece ser protectora cuando, en verdad, es dañina.

Sesgo de cambio

$$RR^0 = 2 > 1 > RR = 0.50$$

La exposición parece tener un efecto dañino cuando, en verdad, es protectora.

O

$$RR^0 = 0.7 < 1 < RR = 1.5$$

La exposición parece ser protectora cuando, en verdad, es dañina.

Preguntas de estudio (P7.6)

Conocer la dirección del sesgo puede tener una importancia práctica para el investigador.

1. Suponga que un investigador encuentra un efecto muy fuerte, digamos que la RR calculada (en la población de estudio) es 6, y también puede argumentar de forma persuasiva que cualquier posible sesgo debe ser hacia el valor nulo. Entonces, ¿qué conclusión puede sacarse sobre el valor correcto (es decir, objetivo) de la RR?
2. Suponga que un investigador encuentra una RR calculada (en la población de estudio) muy débil, digamos de 1,3, y puede argumentar que cualquier sesgo debe alejarse del valor nulo. Entonces, ¿qué conclusión puede sacarse sobre el valor correcto de la RR?

Resumen

- La dirección del sesgo indica si el parámetro objetivo está sobrestimado o subestimado sin especificar la magnitud del sesgo.
- Si el parámetro objetivo está sobrestimado, la dirección del sesgo se aleja del valor nulo.
- Si el parámetro objetivo está subestimado, la dirección del sesgo es hacia el valor nulo.
- El sesgo de cambio se presenta si el parámetro objetivo está del lado opuesto del valor nulo del parámetro que, en verdad, se está calculando en el estudio.

Sesgo positivo versus negativo

Podemos caracterizar la dirección del **sesgo** como **positiva** o **negativa** basándonos en nuestra definición cuantitativa de sesgo. La definición de sesgo la hemos dado utilizando la expresión **sesgo = $\theta^0 - \theta$**, donde **θ^0** es la medida de efecto que, en verdad, se calcula mediante la muestra de estudio y **θ** es el valor de la medida de efecto en la población objetivo o meta de interés.

Para ejemplificar, suponga que el parámetro objetivo es la razón de riesgos, cuyo valor es, por ejemplo, 1,5, y el valor real del parámetro que se está calculando es 4, es decir:

RR = 1,5 y **RR^o = 4**.

Entonces, utilizando la fórmula anterior, el sesgo sería positivo. En forma similar, si:

RR = 0,70 y **RR^o = 0,25**

el sesgo también sería positivo. Por lo tanto, utilizando nuestra definición, si el parámetro objetivo está **sobrestimado**, como sucede en estos dos ejemplos, entonces el sesgo sería **positivo**.

Por el contrario, suponga que el parámetro objetivo es nuevamente la razón de riesgos, cuyo valor es, por ejemplo, 4, y el valor real del parámetro que se calcula es 1,5, es decir:

RR = 4 y **RR^o = 1,5**

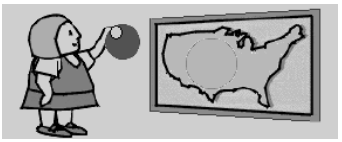
Entonces, utilizando nuestra fórmula del sesgo, el sesgo sería **negativo**. En forma similar, si:

RR = 0,25 y **RR^o = 0,70**
el sesgo también sería **negativo**. Por lo tanto, utilizando nuestra definición, si el parámetro objetivo está **subestimado**, como sucede en estos dos últimos ejemplos, entonces el sesgo sería **negativo**.

Lamentablemente, si deseamos caracterizar el sesgo como positivo o negativo, debemos recordar que la fórmula que usamos sustrae el valor correcto del parámetro del valor incorrecto. Sin embargo, si intercambiamos estos dos parámetros, la conclusión sobre positivo o negativo también estará intercambiada.

En consecuencia, para evitar esta posible confusión, la dirección del sesgo se describe más comúnmente como "se aleja del valor nulo" (es decir, sobrestimación) en contraposición a "hacia el valor nulo" (es decir, subestimación).

Auto-evaluación (P7.7)



Coloque **verdadero** o **falso** en cada una de las siguientes afirmaciones:

1. Si la RR calculada es igual a 2,7 en la muestra y se determina que hay un sesgo que se aleja del valor nulo, entonces la RR en la población objetivo (es decir, de origen) es mayor que 2,7. . **???**
2. Si la RR calculada en la muestra es 1,1 y se determina que hay un sesgo hacia el valor nulo, entonces básicamente no hay asociación en la población objetivo (medida por la RR). . . **???**
3. Si la RR calculada en la muestra es 1,1 y se determina que hay un sesgo que se aleja del valor nulo, entonces básicamente no hay asociación en la población objetivo. . . . **???**
4. Si la RR calculada es igual a 0,4 en la muestra y se determina que hay un sesgo que se aleja del valor nulo, entonces la RR en la población objetivo (es decir, de origen) es menor que 0,4. . **???**
5. Si la RR calculada en la muestra es 0,4 y se determina que hay un sesgo hacia el valor nulo, entonces básicamente no hay asociación en la población objetivo (medida por la RR). . . **???**
6. Si la RR calculada en la muestra es 0,98 y se determina que hay un sesgo que se aleja del valor nulo, entonces básicamente no hay asociación en la población objetivo. . . . **???**

Complete los espacios vacíos:

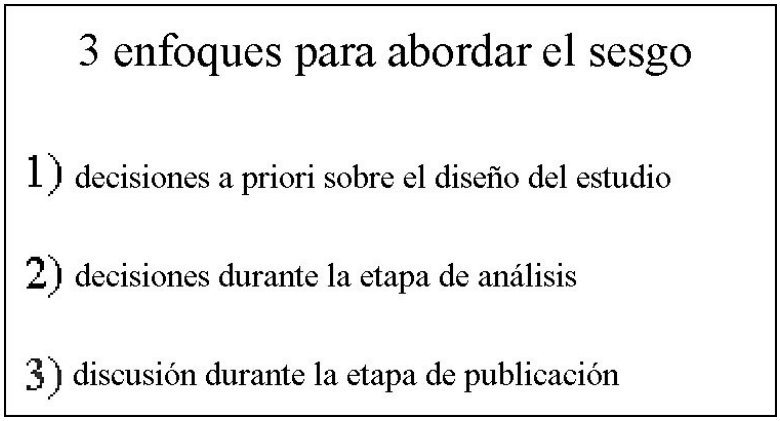
7. Si el OR es igual a 3,6 en la población objetivo y 1,3 en la población de estudio, entonces el sesgo . . . . . **???**
8. Si la RDI es igual a 0,25 en la población objetivo y 0,95 en la población de estudio, entonces el sesgo . . . . . **???**
9. Si la RR es igual a 1 en la población objetivo y 4,1 en la población de estudio, entonces el sesgo . . . . . **???**
10. Si la RR es igual a 0,6 en la población objetivo y 2,1 en la población de estudio, entonces el sesgo . . . . . **???**
11. Si la RR es igual a 1 en la población objetivo y 0,77 en la población de estudio, entonces el sesgo . . . . . **???**
12. Si la RR es igual a 4,0 en la población objetivo y 0,9 en la población de estudio, entonces el sesgo . . . . . **???**

Opciones
se aleja del valor nulo **es de cambio** **es hacia el valor nulo**

¿Qué se puede hacer con respecto al sesgo?

La evaluación del sesgo es normalmente subjetiva e involucra un juicio sobre su presencia, dirección o, mucho más raramente, su magnitud. Sin embargo, hay formas de abordar el problema del sesgo, inclusive el ajuste del cálculo de la muestra para efectuar la "corrección" por el sesgo. A continuación se describen tres enfoques generales.

Aquí hay tres enfoques generales para abordar el sesgo: 1) decisiones *a priori* sobre el **diseño** del estudio; 2) decisiones durante la etapa de **análisis**; y 3) **discusión** durante la etapa de publicación.



Cuando se **diseña** un estudio, se pueden tomar decisiones para minimizar o incluso evitar el sesgo en los respectivos resultados. Se puede evitar el **sesgo de selección** mediante la inclusión o exclusión de los individuos que cumplen con las características, con la selección de la población de origen o el grupo de comparación, por ejemplo, el grupo de control en un estudio de casos y controles.

Preguntas de estudio (P7.8)

1. ¿Qué tipo de sesgo puede evitarse teniendo especial cuidado de medir con precisión la exposición, la enfermedad y las variables de control que se están estudiando, inclusive el uso de estudios pilotos para identificar los problemas de medición que pueden corregirse en el estudio principal?
    A. Sesgo de selección
    B. Sesgo de información
    C. Sesgo de confusión
2. ¿Qué tipo de sesgo puede evitarse asegurándose de medir u observar variables en la etapa de diseño que puedan tener que explicarse en la etapa de análisis?
    A. Sesgo de selección
    B. Sesgo de información
    C. Sesgo de confusión

En la etapa de **análisis**, el investigador quizás puede determinar la presencia o dirección de un posible sesgo aplicando un razonamiento lógico sobre las características metodológicas del diseño del estudio que se está utilizando realmente.

Preguntas de estudio (P7.8) (continuación)

3. En el Estudio de usuarios de las playas de Sydney, tanto la condición de bañista como el resultado de salud se determinaron mediante informes personales de los individuos y sus recuerdos. ¿Esto indica la necesidad de evaluar en la etapa de análisis la presencia o dirección de qué tipo de sesgo?
    A. Sesgo de selección
    B. Sesgo de información
    C. Sesgo de confusión

4. Asimismo, en el Estudio de usuarios de las playas de Sydney, los individuos tenían que ser excluidos del análisis si no completaban la entrevista de seguimiento. Esta falta de respuesta puede afectar la representatividad de la muestra. ¿Este es un ejemplo de qué tipo de sesgo?
    A. Sesgo de selección
    B. Sesgo de información
    C. Sesgo de confusión

En la etapa de **análisis**, el sesgo también puede reducirse o eliminarse *ajustando* el cálculo de la muestra haciendo conjeturas sobre el grado de sesgo. Dicho ajuste se efectúa normalmente para el sesgo de confusión, representando cuantitativamente los efectos de las variables de confusión mediante métodos de **análisis estratificado** o **modelación matemática**.

El ajuste del sesgo de selección y el sesgo de información está limitado por la disponibilidad de la información necesaria para medir el alcance del sesgo. Una fórmula simple para un cálculo "corregido" involucra manipular la ecuación del sesgo, moviendo el parámetro objetivo al lado izquierdo de la ecuación. Esta fórmula no es tan fácil de aplicar como parece. En vez de eso, la mayoría de los investigadores se tendrán que contentar con exponer los argumentos a favor de la dirección del sesgo. El sesgo calculado depende de la disponibilidad de más parámetros fundamentales, que a menudo son difíciles de determinar. Discutimos estos parámetros en más detalle en las lecciones siguientes.

Etapa de análisis

Ajuste

Sesgo de selección/información:

Limitado por la disponibilidad de la información necesaria para medir el **alcance** del sesgo

$$\hat{\theta}^{C} = \hat{\theta} - \hat{SESGO}$$

no es tan fácil de aplicar

θ y θ^{0} nunca se saben

$\hat{\theta}^{C}$ - cálculo corregido

El enfoque final para abordar el sesgo es cómo se informa el estudio. Normalmente, se brinda una descripción de los sesgos potenciales en la sección "Discusión" de una publicación. Esta discusión, especialmente cuando se refiere a un posible sesgo de **selección** o de **información**, es bastante subjetiva, pero se espera un juicio dada la dificultad inherente a la cuantificación de los sesgos. Raramente, si alguna vez ocurre, el investigador admite en la reseña que el sesgo pone en duda seriamente las conclusiones del estudio. Por lo tanto, el lector debe revisar esta sección con mucho cuidado.

Resumen: ¿qué se puede hacer con respecto al sesgo?

- La respuesta depende del tipo de sesgo que se está considerando: de selección, información o confusión.
- Los enfoques para abordar el sesgo son: las decisiones en la etapa de diseño del estudio, en la etapa de análisis y en la etapa de publicación.
- En la etapa de diseño del estudio, pueden tomarse decisiones para evitar el sesgo en los respectivos resultados.
- En la etapa de análisis, se puede utilizar el razonamiento lógico sobre las características metodológicas del diseño del estudio que en verdad se utiliza.
- También en la etapa de análisis, el sesgo de **confusión** puede reducirse o eliminarse si ajusta cuantitativamente el cálculo de la muestra.
- Una fórmula simple para un cálculo corregido: $\grave{e}^{c} = \hat{\grave{e}} - \hat{S}esgo$. Esta fórmula no es tan fácil de aplicar como parece.
- Los sesgos potenciales se describen en la sección **Discusión** de una publicación. Tenga cuidado.

Nomenclatura

θ	Theta, parámetro de la población objetivo o meta
$\hat{\theta}$	"Theta con acento circunflejo", cálculo del parámetro de la muestra que realmente se está analizando
θ^0	Parámetro de la población de estudio
RR	Razón de riesgos de la población objetivo o meta
$\mathbf{RR}^0$	Razón de riesgos de la población de estudio

Referencias

Berkson J. *Limitations of the application of fourfold table analysis to hospital data*. Biometrics Bulletin 1946; 2: 47-53.

Corbett SJ, Rubin GL, Curry GK, Kleinbaum DG. *The health effects of swimming at Sydney Beaches*. The Sydney Beach Users Study Advisory Group. Am J Public Health. 1993; 83(12): 1701-6.

Greenberg RS, Daniels SR, Flanders WD, Eley JW, Boring JR. *Medical Epidemiology* (3rd Ed). Lange Medical Books, Nueva York, 2001.

Hill H, Kleinbaum DG. *Bias in Observational Studies*. In Encyclopedia of Biostatistics, pág. 323-329, Oxford University Press, 1999.

Horwitz RI, Feinstein AR. *Alternative analytic methods for case-control studies of estrogens and endometrial cancer*. N Engl J Med 1978; 299(20):1089-94.

Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic Research: Principles and Quantitative Methods*. John Wiley and Sons Publishers, Nueva York, 1982.

Perera FP, et al. *Biologic markers in risk assessment for environmental carcinogens*. Environ Health Persp 1991; 90:247.

The Lipid Research Clinics Coronary Primary Prevention Trial results: II. The relationship of coronary heart disease to cholesterol lowering. JAMA 1984; 251(2):365-74.

Warner L, Clay-Warner J, Boles J, Williamson J. *Assessing condom use practices. Implications for evaluating method and user effectiveness*. Sex Transm Dis 1998; 25(6):273-7.

Ejercicios
EPI-1. Problemas generales de validez

a. Describa la diferencia entre error sistemático y error aleatorio en términos del tiro al blanco.
b. ¿Por qué debería tener prioridad la consideración de validez sobre la consideración de precisión?
c. Describa un ejemplo en el que la población de estudio no sea un subconjunto de la población de origen. (Sugerencia: muestra de la clínica para enfermedades de transmisión sexual).
d. Describa un ejemplo en el que la población de estudio sea un subconjunto de la población de origen. (Sugerencia: pérdida de seguimiento en un estudio de cohorte).
e. Si la población de origen difiere de la población de estudio, ¿es todavía posible que la muestra de estudio pueda brindar un cálculo del efecto cercano al parámetro objetivo en la población de origen? Desarrolle su respuesta.
f. ¿Por qué no se puede evaluar la validez externa basándose en el diseño/resultados de un único estudio?
g. En terminología estadística, el valor esperado de un estadístico de muestra es el promedio a largo plazo de los valores del estadístico que se obtiene de muestreos repetidos. Explique por qué la definición de "sesgo epidemiológico" en términos de la diferencia entre los valores de los parámetros de la población de estudio y de origen no es equivalente a decir que el "valor esperado del parámetro de la población de estudio es igual al parámetro de la población de origen".
h. Suponga que usted es un investigador interesado en demostrar que determinada variable de exposición está asociada a un resultado de salud de interés. Usted realiza un estudio de casos y controles y encuentra que su

cálculo del *odds ratio* es 3,5. Si puede razonar justificablemente que la dirección de cualquier sesgo posible en su estudio es hacia el valor nulo, explique por qué se sentiría satisfecho con los resultados de su estudio.

i. Suponga que es un científico que trabaja para el gobierno y que está revisando un estudio financiado por su entidad, a fin de evaluar si determinado fármaco que ya está disponible en el mercado tiene consecuencias nocivas para la salud. El estudio encuentra un efecto estadísticamente significativo de 3,5 que indica que la droga es dañina. Sin embargo, basándose en su análisis, usted puede argumentar en forma convincente que cualquier sesgo que pueda existir se aleja del valor nulo. ¿Debería recomendar que se retire la droga del mercado? Desarrolle su respuesta.

EPI-2. Sesgo: genes y cáncer de vejiga

Un grupo de investigadores tenía interés en estudiar la relación potencial entre un gen recientemente descubierto y el desarrollo de cáncer de vejiga. Suponga que se descubrió que las personas cuyas pruebas genéticas resultaron positivas y desarrollaron cáncer de vejiga tendían a morir rápidamente y, por lo tanto, tenían menos probabilidad de estar disponibles para su inclusión en un estudio de casos y controles sobre esta relación. ¿Qué efecto tendría esto sobre el cálculo observado del *odds ratio*? (Una de las siguientes opciones es correcta).

a. Sesgo hacia el valor nulo.
b. Sesgo que se aleja del valor nulo.
c. Sesgo, pero no se puede determinar la dirección.
d. Sin sesgo, pero existe un problema potencial con la validez externa.

EPI-3. Tipo de sesgo: genes y cáncer de vejiga (continuación)

¿Qué tipo de sesgo se está considerando en la pregunta 2 anterior?

EPI-4. Dirección del sesgo: multivitaminas vs. defectos de nacimiento

Se realizó un estudio de casos y controles para evaluar la relación entre la autodeclaración que una mujer brinda sobre el uso de suplementos multivitamínicos (exposición) y el nacimiento subsecuente de un bebé con defectos de nacimiento (resultado). El cuadro siguiente resume los resultados de este estudio:

	Uso de multivitaminas		
	Uso periconcepcional	**Sin uso de vitaminas**	**Total**
Defectos de nacimiento	164	166	330
Sin defectos de nacimiento	228	122	350

Cuando se publicó el estudio, sus autores recibieron críticas por haber elegido un grupo de control con una prevalencia inadecuadamente alta de uso de multivitaminas. ¿Qué impacto tendría el uso de dicho grupo de control en el *odds ratio* observado? (Suponga que la exposición es verdaderamente protectora). Seleccione la mejor respuesta.

a. Crearía un sesgo hacia el valor nulo.
b. Crearía un sesgo que se aleja del valor nulo.
c. Crearía un sesgo, pero no puede predecirse la dirección.
d. Aumentaría la probabilidad de un error de tipo II.

Respuestas a las preguntas de estudio y auto-evaluaciones

P7.1

1. Mujeres entre 45 a 69 años que residen en la ciudad de Atlanta.
2. Una medida de efecto, ya sea una razón de prevalencia o un *odds ratio* de prevalencia (es decir, el punto azul o la diana), para la asociación entre la ingesta de calcio y la prevalencia de artritis de cadera en la población objetivo.
3. El error aleatorio se refiere a si un *odds ratio* calculado en la muestra tomada en hospitales (es decir, el tiro al blanco) difiere del *odds ratio* en la población de pacientes de hospitales (es decir, el punto rojo o el centro de los tiros reales) de donde se selecciona la muestra.
4. El error sistemático se refiere a si el *odds ratio* en la población de pacientes de hospitales de la muestra (es decir, el punto rojo) es diferente del *odds ratio* en la población de origen (es decir, el punto azul).

P7.2

1. Todos los adultos jóvenes de la ciudad de Nueva York.
2. Todos los adultos jóvenes que concurren a gimnasios en la ciudad de Nueva York (NY) y que finalmente se quedarían en el estudio para ser analizados.
3. Sí. La muestra se selecciona de la población de estudio en forma aleatoria y, por lo tanto, es representativa.
4. Probablemente no. El grupo de adultos jóvenes de gimnasios de la ciudad de NY es diferente del grupo de todos los adultos jóvenes de la ciudad de NY. Ya que el ejercicio físico está tan fuertemente relacionado con la salud, el uso de todos los adultos jóvenes que concurren a gimnasios probablemente no es la mejor opción para este estudio.
5. Todos los adultos jóvenes de la ciudad de Nueva York.
6. Todos los adultos jóvenes de la ciudad de NY que finalmente se quedarían en el estudio para ser analizados.
7. Sí. La muestra se selecciona en forma aleatoria de la población de estudio (por definición de la población de estudio) y, por lo tanto, es representativa de ésta. Sin embargo, ni la población de estudio ni la muestra pueden ser representativas de todos los adultos jóvenes de la ciudad de NY, si no todos los que fueron seleccionados en la muestra participan del estudio.
8. Sí. Suponiendo que todos los seleccionados participan en el estudio (es decir, brindan los datos requeridos), la población de estudio es la misma que la población de origen. Sin embargo, si muchos de los que se incluyeron en la muestra (por ejemplo, un subgrupo particular) no brindan los datos de estudio necesarios, la muestra final y su correspondiente población de estudio puede no ser representativa de todos los adultos jóvenes de la ciudad de NY.
9. Falso. Las personas cuyo seguimiento se ha perdido no se encuentran en la muestra de estudio, por lo tanto, no pueden ser incluidas en la población de estudio representada por la muestra.
10. Verdadero. Las personas cuyo seguimiento se ha perdido no se encuentran en la muestra, pero, sin embargo, están incluidas en la población de origen de interés.
11. Verdadero. En la población de estudio, los controles pueden especificarse como personas sin la enfermedad, independientemente de si están en riesgo de ser un caso. No obstante, la población de origen puede contener sólo personas que están en riesgo de ser un caso.
12. Falso. La población de estudio en un estudio transversal está restringida sólo a los sobrevivientes.
13. La población de adultos jóvenes del estado de Nueva York es una población externa, porque el estudio estaba limitado a los adultos jóvenes de la ciudad de Nueva York. Extrapolar los resultados del estudio al estado de Nueva York va más allá de considerar los aspectos metodológicos del estudio real.

P7.3

1. El estudio será internamente válido siempre que la población de estudio correspondiente a la muestra que en realidad se analiza no esté sustancialmente distorsionada con respecto a la

población de origen correspondiente a adultos jóvenes de los gimnasios de la ciudad. Por ejemplo, si la muestra finalmente analizada es una población mucho más sana que la población de origen, puede cuestionarse la validez interna.
2. Validez externa. El estudio estaba limitado a personas de la ciudad de Nueva York. Extrapolar los resultados del estudio al estado de Nueva York va más allá de considerar los aspectos metodológicos del estudio de la ciudad de Nueva York.
3. ¿Los resultados del estudio también se aplican a las mujeres u hombres de diferentes edades?
4. ¿La muestra de estudio era representativa de la población de origen? ¿Los grupos de comparación fueron seleccionados adecuadamente? ¿Los individuos siguieron su dieta y régimen de medicación coherentemente? ¿Se consideraron las variables relevantes?

P7.4

1. Verdadero.
2. Falso. El error sistemático ocurre cuando hay una diferencia entre la medida del efecto verdadero y lo que se calcula en realidad, es decir, una diferencia entre las poblaciones de origen y de estudio.
3. Falso. Un estudio válido significa que no hay error sistemático, pero probablemente todavía haya error aleatorio.
4. Falso. No siempre. En forma ideal, la población de estudio sería equivalente a la población de origen. Sin embargo, puede suceder que la población de estudio y la población de origen simplemente se superpongan.
5. Verdadero. La muestra siempre se selecciona de la población de estudio.
6. Falso. Si la población de estudio no es un subconjunto de la población de origen o no es equivalente a ella, entonces la muestra puede no ser un subconjunto de la población de origen (ni estar completamente incluida en ella).
7. Falso. Pueden ser distintas debido al error aleatorio.
8. Depende. La diferencia puede ser un resultado del error aleatorio. Si la razón de riesgos en la población de estudio difiere relevantemente de 3,0, entonces el estudio no es internamente válido.
9. Verdadero. Cualquier diferencia relevante entre la población de estudio y la población de origen significa que el estudio no es internamente válido.
10. Falso. La diferencia entre 3,0 y 3,1 no se consideraría una diferencia relevante.
11. Depende. No conocemos la razón de riesgos en la población externa.

P7.5

1. El parámetro de la población objetivo es 3,5. No conocemos el valor del parámetro de la población de estudio, pero es probable que esté más cerca de 1,5 que de 3,5, porque se supone que la muestra es representativa de la población de estudio.
2. Parece haber un sesgo en el estudio porque el valor de 1,5 calculado de la muestra es distinto del valor calculado de la población y es razonable pensar que la muestra final ya no representa a la población de origen.
3. El sesgo no puede determinarse con exactitud porque 1,5 es un valor calculado de la muestra. Sin embargo, el sesgo en la razón de riesgos es aproximadamente 1,5 – 3,2 = -2.
4. Sí, siempre que la razón de la diferencia sea el error sistemático.
5. Sí, siempre que la razón de la diferencia sea el error sistemático.

P7.6

1. La RR correcta debe ser incluso mayor que 6.
2. La RR correcta debe indicar un efecto aún menor o ningún efecto.

P7.7

1. Falso. Si el sesgo se aleja del valor nulo, la RR en la población de origen debe estar entre 1 y 2,7.
2. Falso. Si el sesgo es hacia el valor nulo, entonces la RR en la población de origen debe ser mayor que 1,1. Debido a que no podemos determinar cuánto más grande es, no podemos concluir que esto básicamente implica una falta de asociación.

3. Verdadero. Si el sesgo se aleja del valor nulo, entonces la RR en la población de origen está entre 1 y 1,1. Por lo tanto, podemos concluir que básicamente no hay asociación.
4. Falso. Si el sesgo se aleja del valor nulo, entonces la RR en la población de origen debe estar entre 0,4 y 1,0.
5. Falso. Si el sesgo es hacia el valor nulo, entonces la RR en la población de origen debe ser menor que 0,4 y, por lo tanto, existe asociación.
6. Verdadero. Si el sesgo se aleja del valor nulo, entonces la RR en la población de origen está entre 0,98 y 1,0, lo que significa que básicamente no hay asociación.
7. Hacia el valor nulo.
8. Hacia el valor nulo.
9. Se aleja del valor nulo.
10. Es de cambio.
11. Se aleja del valor nulo.
12. Es de cambio.

P7.8

1. B
2. C
3. B
4. A

LECCIÓN 8

SESGO DE SELECCIÓN

8-1 Sesgo de selección

*El **sesgo de selección** se refiere al error sistemático que puede surgir por la manera en que se seleccionan los individuos para un estudio. En esta lección describimos ejemplos de sesgo de selección, brindamos un marco cuantitativo para evaluarlo, mostramos cómo puede ocurrir el sesgo de selección en diferentes tipos de diseños de estudios epidemiológicos y analizamos cómo efectuar el ajuste por el sesgo de selección o, de lo contrario, qué hacer con él.*

El sesgo de selección en diferentes diseños de estudios

El sesgo de selección es un error sistemático que surge por la forma en que se seleccionan los individuos en un estudio o porque hay pérdidas selectivas de individuos antes del análisis de los datos. El sesgo de selección puede ocurrir en cualquier tipo de estudio epidemiológico. En los estudios de casos y controles, la principal fuente de sesgo de selección es la manera en que se seleccionan los casos, los controles o ambos, y la medida en que los antecedentes de exposición influyen en dicha selección. Por ejemplo, el sesgo de selección era de interés en los estudios de casos y controles que encontraron una asociación entre el uso del suplemento de L-triptófano y el síndrome de eosinofilia-mialgia (SEM), una enfermedad caracterizada principalmente por dolores musculares incapacitantes, malestar general y elevado número de eosinófilos. Los *odds ratios* obtenidos de estos estudios pueden haber sobrestimado el verdadero efecto.

Preguntas de estudio (P8.1)

1. Suponiendo que el *odds ratio* que relaciona el L-triptófano con el SEM está sobrestimado, ¿cuál de las siguientes opciones es correcta?
    1. La dirección del sesgo es hacia el valor nulo.
    2. El L-triptófano tiene una asociación más débil con el SEM de lo que en verdad se observa.
    3. El *odds ratio* correcto es mayor que el observado.

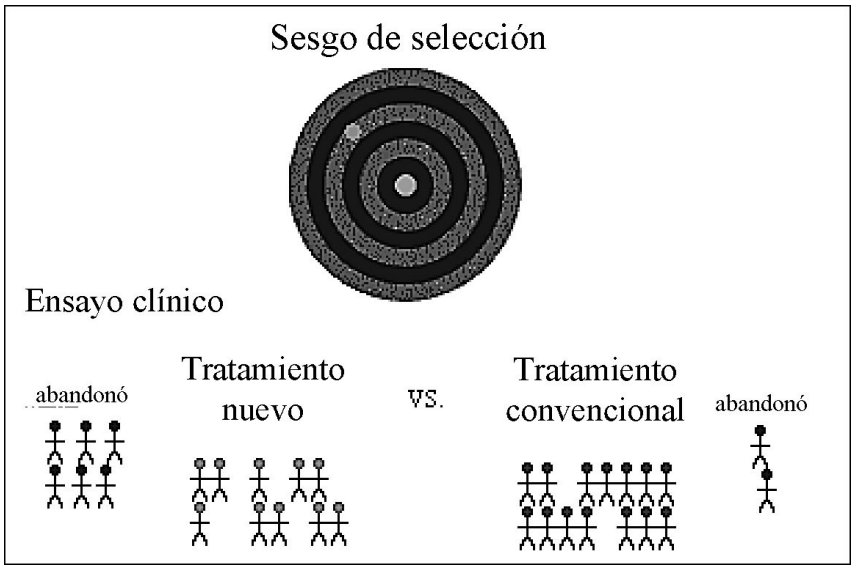
En consecuencia, el sesgo se alejaría del valor nulo.

Una crítica fundamental de estos estudios fue que la publicidad inicial sobre una asociación sospechada puede haber dado como resultado un diagnóstico preferencial del SEM entre los usuarios conocidos de L-triptófano, en comparación con los no usuarios.

Preguntas de estudio (P8.1) (continuación)

2. Suponiendo un diagnóstico preferencial del SEM a partir de la publicidad sobre el L-triptófano, ¿cuál de las siguientes opciones es correcta? La proporción expuesta entre los casos diagnosticados que se seleccionaron para ser estudiados probablemente sea **???** la proporción expuesta entre todos los casos de la población de origen.
    1. mayor que
    2. menor que
    3. igual a

En los estudios de cohorte y los ensayos clínicos, las principales fuentes de sesgo de selección son la **pérdida de seguimiento**, el **abandono del estudio** o la **falta de respuesta**. Por ejemplo, considere un ensayo clínico que compara los efectos de un tratamiento nuevo con los de otro convencional para cierto tipo de cáncer. Suponga que los pacientes a los que se les asigna el nuevo tratamiento tienen más probabilidad de desarrollar efectos colaterales que los que siguen el tratamiento convencional y, en consecuencia, es más probable que abandonen el estudio.



Preguntas de estudio (P8.2)

Ensayo clínico: nuevo tratamiento para el cáncer en contraposición al tratamiento convencional. Suponga que los pacientes con el nuevo tratamiento tienen más probabilidades de abandonar el estudio que los que siguen el tratamiento convencional.

1. ¿Por qué la anterior información sobre el abandono del estudio sugiere la posibilidad de un sesgo de selección en este estudio?
2. ¿Por qué no se solucionará este problema con un análisis por intención de tratar ?
3. ¿Cuál es la población de origen?
4. ¿Cuál es la población de estudio?

En los estudios transversales, la fuente principal del sesgo de selección es lo que se denomina **supervivencia selectiva**. Sólo los sobrevivientes se pueden incluir en los estudios transversales. Si los casos expuestos tienen más probabilidad de sobrevivir más tiempo que los casos no expuestos, o viceversa, las conclusiones obtenidas en un estudio transversal pueden ser diferentes de las obtenidas en un estudio de cohorte adecuado.

Preguntas de estudio (P8.2) (continuación)

Suponga que deseamos evaluar si hay supervivencia selectiva en un estudio transversal.

5. ¿Cuál es la población de origen?
6. ¿Cuál es la población de estudio?

Resumen

- El sesgo de selección puede presentarse a raíz del error sistemático que surge por la forma en que se selecciona a los individuos para un estudio y la manera en que se quedan para ser analizados.
- La razón principal de dicho sesgo generalmente difiere del tipo de estudio utilizado.
- En los estudios de casos y controles, la principal fuente de sesgo de selección es la manera en que se seleccionan los casos, los controles o ambos.
- En los estudios de cohorte y los ensayos clínicos, la principal fuente de sesgo de selección es la pérdida de seguimiento, el abandono del estudio o la falta de respuesta.
- En los estudios transversales, la fuente principal de sesgo de selección es lo que se denomina supervivencia selectiva.

Ejemplo de sesgo de selección en los estudios de casos y controles

En los estudios de casos y controles, debido a que el resultado de salud ya ha ocurrido, la selección de los casos, los controles o ambos podría estar influenciada por una condición de exposición anterior. En la década de los setenta se publicó un debate animado sobre el sesgo de selección en estudios que analizaba si el uso de estrógenos como reemplazo de hormonas producía cáncer de endometrio. Los estudios anteriores fueron estudios de casos y controles e indicaron un fuerte efecto nocivo de los estrógenos. Normalmente, los controles que se usaron fueron mujeres con cánceres ginecológicos distintos del cáncer de endometrio. Los críticos sostuvieron que, debido a que los estrógenos muchas veces causan sangrado vaginal, independientemente de si se padece cáncer, a las usuarias de estrógenos que padecían cáncer de endometrio se les sometería en forma selectiva a un cribado o tamizaje para el diagnóstico de dicha enfermedad, en contraposición a las que no usaban estrógenos y padecían cáncer de endometrio.

Preguntas de estudio (P8.3)

1. Si el razonamiento de los críticos fuese correcto, ¿por qué habría un problema de sesgo de selección al elegir controles que fuesen mujeres con cánceres ginecológicos distintos del cáncer de endometrio?
2. Si el razonamiento de los críticos fuese correcto, ¿esperaría que el *odds ratio* obtenido del estudio estuviese sesgado hacia el valor nulo o que se alejara de éste? Desarrolle su respuesta brevemente.

Se propuso una elección alternativa de los controles: mujeres con tumores benignos en el endometrio, ya que se postuló que dicho grupo de control tendría la misma probabilidad que los casos de ser sometido

selectivamente a un tamizaje para el respectivo diagnóstico.

Preguntas de estudio (P8.3) (continuación)

3. ¿Por qué las usuarias de estrógenos con tumores benignos en el endometrio tendrían más probabilidades de ser sometidas selectivamente a un tamizaje para el diagnóstico de sus tumores, en contraposición a las que no usaban estrógenos?
4. Suponiendo que las usuarias de estrógenos que padecen tanto tumores cancerosos como benignos en el endometrio tienen la probabilidad de ser sometidas selectivamente a un tamizaje para el diagnóstico de sus tumores, en contraposición a las no usuarias, ¿qué problema puede seguir existiendo si se elige al último grupo como control?

Sin embargo, la constante investigación y el debate han indicado que no es probable que el tamizaje selectivo para el diagnóstico de casos introduzca mucho sesgo. De hecho, la alternativa propuesta para la elección de los controles podría, en verdad, producir un sesgo.

Preguntas de estudio (P8.3) (continuación)

Los investigadores llegaron a la conclusión de que, como a casi todas las mujeres con cáncer de endometrio invasivo se les iba a diagnosticar la enfermedad en última instancia, las usuarias de estrógenos iban a estar levemente sobrerrepresentadas, si así fuera, entre una serie de mujeres con cáncer de endometrio. Para las preguntas siguientes, suponga que el tamizaje selectivo para el diagnóstico de casos no influye en la detección de casos de cáncer de endometrio.

5. Si el grupo de control estuviera compuesto por mujeres con tumores benignos en el endometrio, ¿por qué esperaría tener sesgo de selección?
6. ¿La dirección del sesgo sería hacia el valor nulo o se alejaría de éste? Desarrolle su respuesta brevemente.
7. Si el grupo de control estuviera compuesto por mujeres con cánceres ginecológicos distintos del de endometrio, ¿por qué no esperaría un sesgo de selección en el cálculo del *odds ratio*?

En la práctica médica actual, el punto de vista preponderante sostiene que tomar sólo estrógenos es potencialmente dañino para el cáncer de endometrio. En consecuencia, a las mujeres a las que se les recomienda la terapia de reemplazo de hormonas normalmente se les suministra una combinación de progesterona y estrógenos, en lugar de sólo estrógenos.

Resumen

- El sesgo de selección se refiere a una distorsión de los resultados del estudio que ocurre debido a la manera en que se selecciona a los individuos para el estudio.
- En los estudios de casos y controles, la principal preocupación es que la selección de los casos, los controles o ambos podría estar influenciada por una condición de exposición anterior.
- En la década de los setenta se publicó un debate animado sobre un posible sesgo de selección entre los investigadores que estudian si el uso de estrógenos -la exposición- como reemplazo de hormonas produce cáncer de endometrio.
- El argumento que respalda el sesgo de selección no se ha sostenido en el tiempo; la práctica médica actual con respecto a la terapia de reemplazo de hormonas normalmente involucra una combinación de

progesterona y estrógenos, en lugar de sólo estrógenos.

Ejemplo de sesgo de selección en los estudios de cohorte

El sesgo de selección se presenta en los estudios de cohorte, así como en los estudios de casos y controles. En los estudios de cohorte prospectivos, el resultado de salud, que todavía no ha ocurrido cuando se determina la condición de exposición, no puede influenciar la forma en que se seleccionan los individuos para el estudio. Sin embargo, si el resultado de salud no está determinado para todos los que fueron seleccionados inicialmente para el estudio, los resultados pueden estar sesgados. Las principales fuentes de dicho sesgo de selección son la **pérdida de seguimiento**, el **abandono del estudio** o la **falta de respuesta**. Es posible que el conjunto de individuos que se quedan para ser analizados ya no represente a la población de origen de la cual se seleccionó la muestra original.

Considere esta tabla de doble entrada que describe el seguimiento durante cinco años de la enfermedad "D" en determinada población de origen.

Estudio de seguimiento

(Población de origen)	E	no E	Total
D	150	50	200
no D	9850	9950	19800
Total	10000	10000	20000

$$RR = (150/10000) / (50/10000) = 3.0$$

Suponga que se realiza un estudio de cohorte utilizando una muestra aleatoria del 10% de esta población. ¿Cuáles serían las frecuencias de celda esperadas para esta cohorte suponiendo que **no hay** sesgo de selección?

Muestra aleatoria del 10% : ¡Llene los espacios!

(Muestra Inicial)	E	No E	
D	15	5	
No D	985	995	
Total	1000	1000	

Preguntas de estudio (P8.4)

1. Suponiendo que ésta es la muestra que se analiza, ¿hay sesgo de selección?

Suponga que la cohorte inicial se obtuvo del muestreo del 10%. Sin embargo, ahora suponga que se pierde el seguimiento del 20% de las personas expuestas y el 10% de las personas no expuestas. Asimismo, suponga que las personas expuestas tienen el mismo riesgo de enfermedad en la cohorte final que en la cohorte

inicial y que lo mismo es cierto para las personas no expuestas.

Ahora suponga:

Muestra aleatoria del 10% y pérdida al seguimiento

	Se pierde el seguimiento del 20% E	Se pierde el seguimiento del 10% No E
D	12	4.5
No D	788	895.5
Total	800	900

Preguntas de estudio (P8.4) (continuación)

2. ¿La muestra que se acaba de describir representa a la población de origen o a la población de estudio?
3. ¿Cuál es la población de origen para la cual se deriva esta muestra?
4. Para los supuestos anteriores, ¿hay sesgo de selección?

Suponga que un patrón distinto de pérdida de seguimiento da como resultado la tabla de doble entrada que se muestra aquí.

Muestra aleatoria del 10%,patrón distinto de pérdida al seguimiento

	E	No E
D	14	4
No D	786	896
Total	800	900

$\hat{RR} = 3.9$

Preguntas de estudio (P8.4) (continuación)

5. ¿Estos resultados indican un sesgo de selección?
6. ¿Las personas expuestas en la población de estudio tienen el mismo riesgo de enfermedad que las de la población de origen?
7. ¿Las personas no expuestas en la población de estudio tienen el mismo riesgo de enfermedad que las de la población de origen?
8. ¿Los ejemplos anteriores demuestran que habrá sesgo de selección en estudios de cohorte siempre que el porcentaje de pérdida de seguimiento en el grupo expuesto difiera del porcentaje de pérdida de seguimiento en el grupo no expuesto?

Resumen

- Las principales fuentes de sesgo de selección en los estudios de cohorte son la pérdida de seguimiento, el abandono del estudio y la falta de respuesta.
- En los estudios de cohorte, es posible que el conjunto de individuos que se quedan en la muestra final que es analizada ya no represente a la población de origen de la cual se seleccionó la cohorte original.

- El sesgo de selección ocurrirá si la pérdida de seguimiento da como resultado un riesgo de enfermedad en los grupos de individuos expuestos y/o no expuestos que difiere en la muestra final, cuando se compara con la cohorte original.

Algunas sutilezas sobre el sesgo de selección en los estudios de cohorte

Referencia: Hill, HA y Kleinbaum, DG, *Bias in Observational Studies*, en Encyclopedia of Biostatistics, P.A. Armitage y T Colton, ed. junio de 1998.

El sesgo de selección en los estudios de cohorte puede ocurrir incluso con una tasa de respuesta general bastante alta o con muy poca pérdida de seguimiento. Considere un estudio de cohorte en el que el 95% de todos los individuos que se reunieron originalmente en la cohorte continúan en el estudio hasta el final. Es decir, sólo el 5% de los individuos se pierden durante el seguimiento. Si las pérdidas de seguimiento se encuentran principalmente entre los individuos expuestos que desarrollaron la enfermedad, entonces la razón de riesgos correcta (es decir, objetivo) podría estar subestimada en forma considerable, a pesar de la pequeña cantidad de pérdida de seguimiento. Esto se debe a que, en la muestra que se analiza, el riesgo estimado de desarrollar la enfermedad en los individuos expuestos será menor que el de la población de origen, mientras que el riesgo correspondiente a los individuos no expuestos reflejará con precisión a la población de origen.

Es posible que no exista sesgo de selección a pesar de que haya bajas tasas de respuesta o alta pérdida de seguimiento. Suponga que sólo el 10% de todos los individuos que se seleccionaron inicialmente acuerdan participar del estudio, pero este 10% representa a una verdadera muestra aleatoria de la población de origen. Entonces, la razón de riesgos resultante no estará sesgada. La cuestión clave aquí es determinar si los riesgos para los expuestos y los no expuestos en la muestra que se analiza están modificados desproporcionadamente debido a una falta de respuesta o a una pérdida de seguimiento con respecto a los riesgos correspondientes en la población de origen de la cual se seleccionó la muestra inicial.

Básicamente, aquí estamos comparando dos cuadros de 2 x 2: uno representa a la población de origen y el otro a la muestra:

Población de origen

	E	**no E**
D	A	B
no D	C	D
Total	N_1	N_0

Muestra para análisis

	E	**no E**
D	a	b
no D	c	d
Total	n_1	n_0

Cuando se consideran estos cuadros, el sesgo de selección ocurrirá sólo si la razón de riesgos en la población de origen, es decir,

$$\frac{A/N_1}{B/N_0}$$

difiere razonablemente de la razón de riesgos en la muestra, es decir,

$$\frac{a/n_1}{b/n_0}$$

En el primer ejemplo (95% de pérdida de seguimiento), el argumento para el sesgo de selección es, básicamente, que el numerador $\mathbf{a/n_1}$ de la razón de riesgos en la muestra analizada sería menor que el respectivo numerador $\mathbf{A/N_1}$ en la población de origen, mientras que los denominadores correspondientes en estas dos razones de riesgos serían iguales.

En el segundo ejemplo (10% de falta de respuesta), el argumento para la ausencia de sesgo de selección es básicamente que, a pesar de la alta falta de respuesta, los respectivos numeradores y denominadores en la población de origen y en la muestra son iguales.

Otros ejemplos

Aquí se exponen algunos ejemplos adicionales de estudios donde es probable que se planteen cuestionamientos sobre el sesgo de selección.

Preguntas de estudio (P8.5)

Considere un estudio de cohorte retrospectivo que compare trabajadores en determinada industria química con un grupo de comparación basado en la población con respecto al desarrollo de enfermedad coronaria (EC).

1. En dicho estudio se puede presentar un sesgo de selección debido al denominado "efecto del trabajador sano". ¿Cómo podría surgir dicho sesgo?

El sesgo de selección puede generarse a partir del uso de voluntarios para un estudio.

2. Explique la afirmación anterior en términos de población de estudio y población de origen.
3. ¿Cuál es la forma alternativa de ver el problema de validez que surge cuando un estudio está limitado a voluntarios?

Al evaluar los trastornos neurológicos de largo plazo entre los niños que padecen convulsiones febriles, los estudios basados en clínicas tienden a informar una frecuencia mucho más alta de dichos desórdenes que la que se encuentra en los estudios basados en la población.

4. ¿La afirmación anterior indica que los estudios basados en clínicas pueden causar un sesgo de selección? Desarrolle su respuesta brevemente.

Resumen

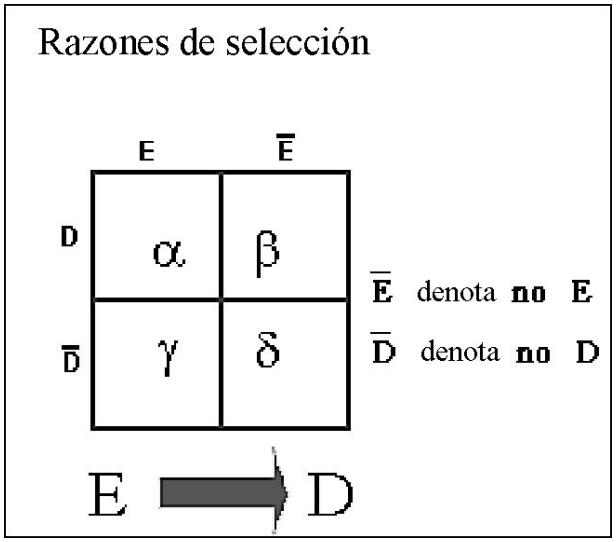
- El sesgo de selección se puede presentar debido al llamado "efecto del trabajador sano". Los trabajadores tienden a ser más saludables que los individuos en la población general y, por lo tanto, pueden tener un resultado más favorable, independientemente de su condición de exposición.
- El sesgo de selección puede surgir del uso de voluntarios, que pueden tener características distintas de las personas que no lo son.
- Los estudios basados en clínicas pueden provocar un sesgo de selección porque los pacientes de dichos establecimientos tienden a padecer una enfermedad más severa que las personas de una muestra basada en la población.

8-2 Sesgo de selección (continuación)

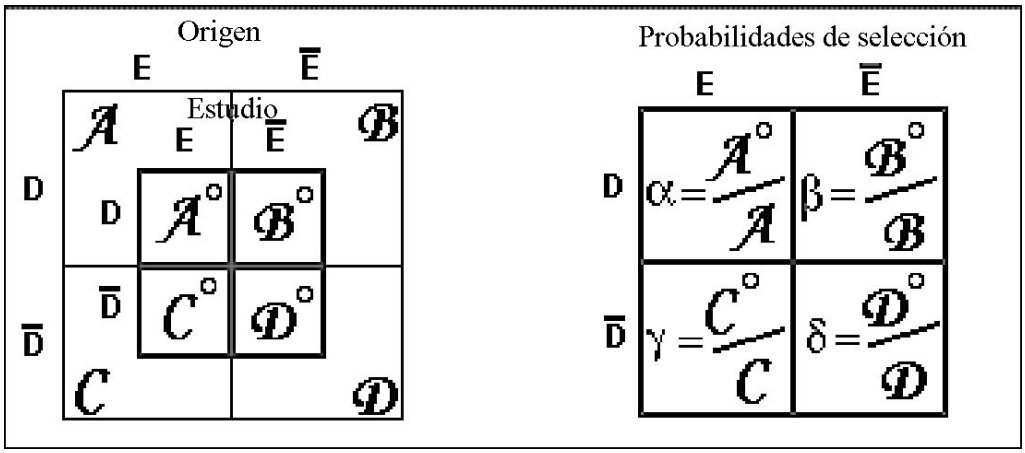
Razones de selección y probabilidades de selección

La probabilidad de selección da la probabilidad de que una persona de una de las cuatro celdas en la población de origen se convierta en miembro de la población de estudio.

Para cuantificar cómo puede ocurrir el sesgo de selección, necesitamos considerar parámetros subyacentes llamados **razones de selección**. Existen cuatro de dichas razones de selección: **alfa (α)**, **beta (β)**, **gamma (γ)** y **delta (δ)**; éstas corresponden a las cuatro celdas del cuadro de doble entrada que relaciona la exposición con la enfermedad.



Una razón de selección se define dividiendo la frecuencia de celda de una de las cuatro celdas en la población de estudio por la frecuencia de celda correspondiente en la población de origen. Utilizando este marco, **α** = A^o/A, **β** = B^o/B, etc. En muchos estudios, la población de estudio será un subconjunto de la población de origen. De ser así, **α**, **β**, **γ** y **δ** se denominan, normalmente, probabilidades de selección.



Cada probabilidad de selección da la probabilidad de que un individuo de una de las cuatro celdas en la población de origen sea candidato a ser seleccionado para la población de estudio. Por lo tanto, **α** denota la probabilidad de que un caso expuesto en la población de origen sea seleccionado para el estudio y se aplican cálculos de probabilidad similares a **β**, **γ** y **δ**.

Preguntas de estudio (P8.6)

Las tablas que aparecen a continuación muestran las poblaciones de origen y de estudio para un estudio de cohorte en el que se ha perdido el seguimiento del 20% del grupo expuesto y del 10% del grupo no expuesto. Estas tablas suponen que las personas expuestas tienen el mismo riesgo de enfermedad en la cohorte final que en la cohorte original y que lo mismo sucede para las personas no expuestas.

1. ¿Cuáles son los valores de **α**, **β**, **γ** y **δ** para estos datos?
2. ¿Hay sesgo de selección en el *odds ratio* o en la razón de riesgos?
3. ¿Cuál es el valor del producto cruzado de las probabilidades de selección, es decir, (αδ)/(βγ)?

Pobl. de estudio	E	No E
D	150	50
no D	9850	9950
Total	10000	10000

RR = 3.0

Pobl. de estudio	E	No E
D	12	4.5
no D	788	895.5
Total	800	900

$RR^0 = 3.0$

Las tablas que aparecen a continuación muestran las poblaciones de origen y de estudio para otro estudio de cohorte en el que se ha perdido el seguimiento del 20% del grupo expuesto y del 10% del grupo no expuesto. En estas tablas, las personas expuestas tienen un riesgo de enfermedad más alto en la población de estudio que en la población de origen y sucede lo contrario para las personas no expuestas.

4. ¿Cuáles son los valores de **α**, **β**, **γ** y **δ** para estos datos?
5. ¿Hay sesgo de selección en el *odds ratio* o en la razón de riesgos?
6. ¿Cuál es el valor del producto cruzado de las probabilidades de selección, es decir (αδ)/(βγ)?

Pobl. de estudio	E	No E
D	150	50
no D	9850	9950
Total	10000	10000

RR = 3.0

Pobl. de estudio	E	No E
D	14	4
no D	786	896
Total	800	900

$RR^0 = 3.9$

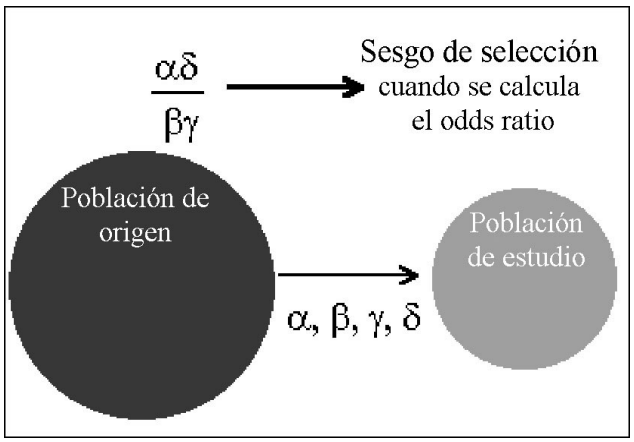
Resumen

- Para cuantificar cómo puede ocurrir el sesgo de selección, necesitamos considerar parámetros subyacentes llamados razones de selección.
- Existen cuatro razones de selección por considerar, una por cada celda del cuadro de 2 x 2 que relaciona la condición de exposición con la condición de enfermedad.

- Una razón de selección brinda la cantidad de individuos de una de las cuatro celdas en la población de estudio dividido por la cantidad correspondiente de individuos en la población de origen.
- Si la población de estudio es un subconjunto de la población de origen, la razón de selección normalmente se denomina probabilidad de selección.
- La probabilidad de selección da la probabilidad de que una persona de una de las cuatro celdas en la población de origen se convierta en miembro de la población de estudio.

Evaluación cuantitativa del sesgo de selección

El sesgo de selección puede evaluarse utilizando una expresión matemática que involucra las cuatro razones de selección o probabilidades de selección que relacionan a la población de origen con la población de estudio. Esta expresión considera el sesgo cuando se calcula el *odds ratio*.

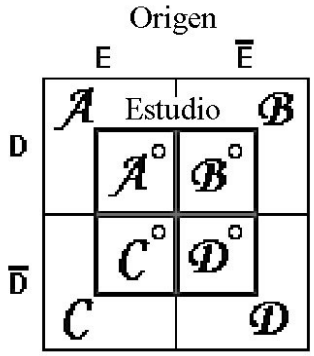


Recuerde que el cálculo del *odds ratio* ($\hat{\mathrm{OR}}$) está sesgado si el *odds ratio* en la población de estudio (OR^o) difiere razonablemente del *odds ratio* en la población de origen (OR). El sesgo es la diferencia entre estos dos parámetros, como se muestra aquí:

$$\text{Sesgo}(\hat{OR}, OR) = OR^0 - OR$$

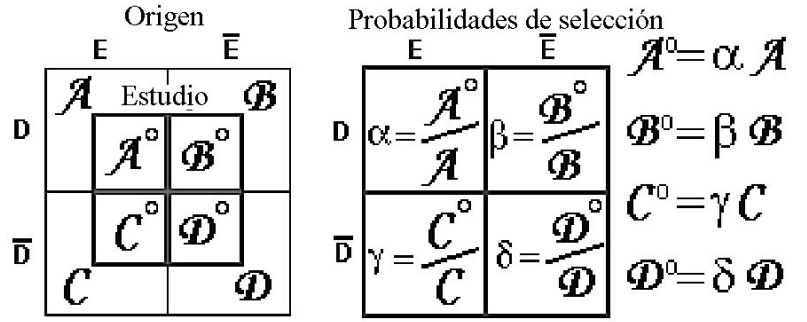
Ahora, si utilizamos el marco descrito anteriormente para las tablas de 4 celdas en las poblaciones de origen y de estudio, podemos expresar el sesgo sustituyendo las frecuencias de celda de cada tabla en la fórmula de sesgo, como se muestra a continuación:

$$\text{Sesgo}(\hat{OR}, OR) = OR^0 - OR = \frac{A^0 D^0}{B^0 C^0} - \frac{AD}{BC}$$



Debido a que las razones de selección α, β, γ y δ se definen como las frecuencias relativas para la población de estudio divididas por la población de origen, podemos reescribir en forma equivalente $A^o = \alpha\, x\, A$, $B^o = \beta\, x\, B$, $C^o = \gamma\, x\, C$, $D^o = \delta\, x\, D$.

$$\text{Sesgo}(\hat{OR}, OR) = \frac{A^0 D^0}{B^0 C^0} - \frac{AD}{BC}$$



Luego sustituimos estas expresiones por las frecuencias de celda en la expresión del sesgo para obtener el resultado que se muestra aquí.

$$\text{Sesgo}(\hat{OR}, OR) = \frac{\alpha A \delta D}{\beta B \gamma C} - \frac{AD}{BC}$$

Si igualamos esta expresión a 0, surge que el **sesgo = 0** si y sólo si la razón de productos cruzados $(\alpha\delta)/(\beta\gamma) = 1$. Por lo tanto, obtenemos la siguiente regla para evaluar el sesgo de selección en el *odds ratio*.

$$\frac{\alpha\delta}{\beta\gamma} = 1 \Longleftrightarrow \text{Sesgo}_{OR} = 0$$

El sesgo es > 0, = 0 ó < 0, dependiendo si la razón de productos cruzados de las probabilidades de selección es > 1, = 1 ó < 1.

$$\text{Sesgo}_{OR} \begin{cases} > 0 \text{ si } \dfrac{\alpha\,\delta}{\beta\,\gamma} > 1 \\ = 0 \text{ si } \dfrac{\alpha\,\delta}{\beta\,\gamma} = 1 \\ < 0 \text{ si } \dfrac{\alpha\,\delta}{\beta\,\gamma} < 1 \end{cases}$$

La parte "= 0" de esta regla establece condiciones para la **ausencia de sesgo**, mientras que las partes > 0 y < 0 establecen condiciones para la **dirección del sesgo**. Sin embargo, al interpretar esta regla, debemos recordar que el signo igual indica "razonablemente igual" en lugar de "exactamente igual", y deberían aplicarse interpretaciones similares a los signos de mayor y menor.

Preguntas de estudio (P8.7)

Las tablas que aparecen a continuación muestran las poblaciones de origen y de estudio para un estudio de cohorte en el que se ha perdido el seguimiento del 20% del grupo expuesto y del 10% del grupo no expuesto. Las razones de selección para la población de estudio 1 son: $\alpha = 12/150 = 0{,}08$; $\beta = 4{,}5/50 = 0{,}09$; $\gamma = 788/9850 = 0{,}08$; y $\delta = 895{,}5/9950 = 0{,}09$.

1. ¿Hay sesgo de selección en el *odds ratio* para la población de estudio 1? Explique brevemente en términos de la razón de productos cruzados de las razones de selección.

Pobl. de origen	E	no E
D	150	50
no D	9850	9950
Total	10000	10000

Pobl. de estudio	E	no E
D	12	4.5
no D	788	895.5
Total	800	900

Las tablas que se muestran a continuación consideran la misma población de origen que la de la pregunta anterior, pero una población de estudio distinta (n.° 2). Las razones de selección para la población de estudio 2 son: $\alpha = 14/150 = 0{,}093$; $\beta = 4/50 = 0{,}08$; $\gamma = 786/9850 = 0{,}08$; y $\delta = 896/9950 = 0{,}09$.

2. ¿Hay sesgo de selección en el *odds ratio* para la población 2? Explique brevemente en términos de la razón de productos cruzados de las razones de selección.
3. Si su respuesta a la pregunta anterior fue positiva, ¿cuál es la dirección del sesgo?

Pobl. de origen	E	no E	Pobl. de estudio	E	no E
D	150	50	D	14	4
no D	9850	9950	no D	786	896
Total	10000	10000	Total	800	900

Resumen

- El sesgo de selección puede evaluarse utilizando una expresión matemática que involucra las cuatro probabilidades de selección que relacionan la población objetivo con la población de estudio.
- En el cálculo del *odds ratio*, el sesgo = 0 si y sólo si la razón de productos cruzados $(\alpha\delta)/(\beta\gamma) = 1$.
- El sesgo es > 0, $= 0$ ó < 0, dependiendo si la razón de productos cruzados de las probabilidades de selección es > 1, $= 1$ ó < 1.

Sesgo de selección para la razón de riesgos

Ahora analizaremos cómo evaluar el sesgo de selección para la razón de riesgos. Hemos mostrado que la razón de riesgos calculada a partir de un estudio de seguimiento puede aproximarse mediante el *odds ratio* si la enfermedad es poco frecuente. Asimismo, el *odds ratio* muchas veces no es tan diferente de la razón de riesgos, incluso si la enfermedad no es poco frecuente. En consecuencia, esta aproximación de enfermedad poco frecuente sugiere que la evaluación del sesgo de selección en la razón de riesgos muchas veces puede involucrar la misma regla sobre el producto cruzado de razones de selección $[(\alpha\delta)/(\beta\gamma)]$ que se aplica a la evaluación del sesgo de selección en el *odds ratio*.

Preguntas de estudio (P8.8)

Las siguientes tablas brindan las poblaciones de origen y de estudio para un estudio de cohorte. Aquí no se pierde el seguimiento de ningún individuo que desarrolló la enfermedad. Sin embargo, se pierde el seguimiento del 20% de los individuos que no padecen la enfermedad, independientemente de la condición de exposición.

1. Calcule las 4 razones de selección para este estudio.
2. ¿Las razones de selección son también probabilidades de selección?
3. ¿La enfermedad en este estudio es poco frecuente tanto en la población de origen como en la población de estudio?
4. Sobre la base de las razones de selección, ¿hay sesgo de selección en el *odds ratio*?
5. ¿Hay sesgo de selección en la razón de riesgos? Desarrolle su respuesta brevemente.

Población de origen	E	No E
D	100	50
no D	9900	9950
Total	10000	10000
RR = 2.0, OR = 2.01		

Población de estudio	E	No E
D	100	50
no D	7920	7960
Total	8020	8010
$RR^0 = 2.0$, $OR^0 = 2.01$		

En el ejemplo, la aproximación de enfermedad poco frecuente se aplicó tanto a la población de origen como a la población de estudio. Sin embargo, si la aproximación de enfermedad poco frecuente no se aplica para estas dos poblaciones, la presencia o ausencia de sesgo en el *odds ratio* podría no corresponder al sesgo en la razón de riesgos calculada para los mismos datos.

Preguntas de estudio (P8.8) (continuación)

Las tablas que aparecen a continuación muestran la misma población de origen, pero una población de estudio distinta de la del ejemplo anterior. Nuevamente, no se pierde el seguimiento de ningún individuo que desarrolló la enfermedad. Sin embargo, hay una pérdida extremadamente alta de seguimiento (99%) de los individuos que **no** desarrollaron la enfermedad, independientemente de la condición de exposición.

6. Calcule las cuatro razones de selección para este estudio.
7. Sobre la base de las razones de selección, ¿hay sesgo de selección en el *odds ratio*?
8. ¿Hay sesgo de selección en la razón de riesgos? Desarrolle su respuesta brevemente.
9. ¿La enfermedad en este estudio es poco frecuente tanto en la población de origen como en la población de estudio?

Población de origen	E	No E
D	100	50
no D	9900	9950
Total	10000	10000

RR = 2.0, OR = 2.01

Población de estudio	E	No E
D	100	50
no D	99	99.5
Total	199	149.5

$RR^0 = 1.503$, $OR^0 = 2.01$

El ejemplo anterior ilustra que la presencia o ausencia de sesgo en el *odds ratio* **no** siempre corresponderá al mismo grado de sesgo en la razón de riesgos. En consecuencia, debe tenerse cuidado cuando se intenta utilizar la regla del producto cruzado de las razones de selección para evaluar si existe sesgo en la razón de riesgos.

Preguntas de estudio (P8.8) (continuación)

Las siguientes tablas muestran poblaciones de origen y de estudio diferentes de las consideradas anteriormente. Se trata de un estudio de cohorte en el que no se perdió el seguimiento de ningún individuo que desarrolló la enfermedad. Sin embargo, el 50% de los no casos expuestos se perdieron durante el seguimiento, mientras que ninguno de los no casos no expuestos se perdió durante el seguimiento.

10. Calcule las 4 razones de selección para este estudio.
11. Sobre la base de las razones de selección, ¿hay sesgo de selección en el *odds ratio*?
12. ¿Hay sesgo de selección en la razón de riesgos? Desarrolle su respuesta brevemente.
13. ¿La enfermedad en este estudio es poco frecuente tanto en la población de origen como en la población de estudio?
14. ¿El alcance del sesgo es el mismo para la razón de riesgos y para el *odds ratio*?

Población de origen	E	No E
D	40	10
no D	60	90
Total	100	100

RR = 4.0, OR = 6.0

Población de estudio	E	No E
D	40	10
no D	30	90
Total	70	100

$RR^{0} = 5.71$, $OR^{0} = 12$

Resumen

- La aproximación de enfermedad poco frecuente sugiere que la evaluación del sesgo de selección en la **razón de riesgos** puede involucrar la misma regla sobre el producto cruzado de las **razones de selección** que se aplica a la evaluación del sesgo de selección en el ***odds ratio***.
- Sin embargo, si la aproximación de enfermedad poco frecuente no aplica tanto para la población de origen como para la población de estudio, la presencia o ausencia de sesgo en el *odds ratio* podría no corresponder al sesgo en la razón de riesgos calculada para los mismos datos.
- La presencia o ausencia de sesgo en el *odds ratio* **no** siempre corresponderá al mismo grado de sesgo en la razón de riesgos.

Auto-evaluación (P8.9)

Complete los espacios vacíos

Cuadro de probabilidades de selección

	E	$\bar{E}$
D	α	β
$\bar{D}$	γ	δ

Los críticos de los estudios de casos y controles que encontraron una asociación entre el uso de estrógenos (E) y el cáncer de endometrio (D) sostuvieron que las usuarias de estrógenos que padecen la enfermedad tienen más probabilidad de ser seleccionadas como casos que las no usuarias de estrógenos que padecen la enfermedad.

1. En el cuadro de probabilidades de selección anterior, ¿cómo se traduce la afirmación de los críticos en una declaración sobre las probabilidades de selección? **???**

Estos críticos también alegaron que si las pacientes ginecológicas con tumores benignos fuesen elegidas como controles, entonces las usuarias de estrógenos que no padecen la enfermedad tendrían más probabilidad de ser seleccionadas como controles que las no usuarias de estrógenos que no padecen la enfermedad.

2. En el cuadro anterior, ¿cómo se traduce esta última afirmación en una declaración sobre las probabilidades de selección? **???**

Se refutó a los críticos de la pregunta anterior con el argumento de que a casi todas las mujeres con cáncer invasivo de endometrio se les diagnosticaría la enfermedad en última instancia, de manera que no había probabilidades de que las usuarias de estrógenos estuvieran sobreseleccionadas en estudios anteriores.

3. En el cuadro de probabilidades de selección anterior, ¿cómo se traduce este último argumento en una declaración sobre las probabilidades de selección? **???**

Aquí también se sostuvo que si las mujeres con cánceres ginecológicos fuesen elegidas como controles, este último grupo no estaría sujeto a un tamizaje selectivo de diagnóstico y sus cánceres casi siempre se detectarían sin dicho tamizaje.

4. En el cuadro anterior, ¿cómo se traduce esta última afirmación en una declaración sobre las probabilidades de selección? **???**

Opciones
alfa = beta **alfa > beta** **alfa > delta** **alfa > gamma** **gamma = delta** **gamma > delta**

8-3 Sesgo de selección (continuación)

Sesgo de selección debido a una elección inadecuada de los controles

Se recomienda ampliamente el uso de preservativos para la prevención de enfermedades de transmisión sexual (ETS). Sin embargo, muchos estudios han encontrado una cantidad sorprendentemente pequeña de pruebas sobre la efectividad de los preservativos, incluso para las ETS, donde se esperaría una eficacia alta. La aparente falta de efecto del uso del preservativo encontrada en los estudios de ETS se atribuye normalmente a una clasificación errónea de la exposición debido a que la utilización de preservativos se informa de manera poco precisa. Sin embargo, otra explicación podría ser el sesgo de selección que surge de la elección del grupo de comparación utilizado.

La población de estudio en los estudios sobre efectividad del preservativo comprende, en general, a individuos examinados en las clínicas de ETS. En consecuencia, el grupo de control en los estudios de casos y controles sobre efectividad del preservativo se elige como una muestra aleatoria de personas que no padecen ETS y que pertenecen a la población de estas clínicas. Una crítica a dichos estudios es que la población de origen correcta está compuesta por individuos que tienen compañeros sexuales infectados, ya que no se puede contraer una ETS sin tener tal compañero. Esta población de origen es más limitada que una población de individuos que visitan clínicas de ETS, pero asegura que un individuo que cumple los requisitos para ser seleccionado para el estudio en verdad está en riesgo de contraer una ETS.

Si los controles se seleccionan de ***todas*** las personas que no padecen ETS, y no sólo de aquellas que poseen parejas infectadas, entonces los controles no expuestos podrían incluir a personas que no poseen parejas infectadas y quizás no muestren un comportamiento de alto riesgo de contraer ETS. Por el contrario, los controles expuestos pueden incluir una proporción mucho más grande de personas con parejas infectadas.

Preguntas de estudio (P8.10)

Suponga que los controles se seleccionan de entre todas las personas examinadas en las clínicas de ETS que no padecen una enfermedad de este tipo. Coloque **verdadero** o **falso** en cada una de las siguientes afirmaciones:

1. Es probable que los controles no expuestos estén sobrerrepresentados en el estudio cuando se comparan con los no casos no expuestos en la población de origen (es decir, sólo aquellos en riesgo).
2. Es probable que los controles expuestos estén sobrerrepresentados en el estudio cuando se comparan con los no casos expuestos en la población de origen.

Las siguientes tablas consideran un estudio hipotético de casos y controles en el que los controles se eligen de entre todos los pacientes de las clínicas que **no** padecen ETS, incluso si la población de origen está limitada a los pacientes con riesgo de ser un caso.

Población de origen

ETS

	Sí	No	Total
Caso	100	600	700
No caso	200	600	800
Total	300	1200	1500

$OR = 0.5$

Población de estudio

ETS

	Sí	No	Total
Caso	100	600	700
Control	200	1200	1400
Total	300	1800	2100

$OR^0 = 1.0$

3. ¿Hay sesgo de selección en el *odds ratio*?
4. ¿Los no casos no expuestos están sobrerrepresentados en la población de estudio cuando son comparados con los no casos no expuestos en la población de origen?
5. ¿Cuáles son los valores de α, β, γ y δ para el estudio hipotético?
6. Utilizando las razones de selección que se acaban de calcular, ¿de qué forma este ejercicio ilustra el sesgo de selección cuando se permite que los controles incluyan a personas que tienen compañeros sexuales que no padecen ETS?
7. ¿La dirección del sesgo de selección descrito en la pregunta anterior sería hacia el valor nulo o se alejaría de éste?

Una manera de minimizar el sesgo de selección en estos estudios sería redefinir el grupo de control como los individuos no infectados que, con certeza, poseen compañeros sexuales infectados con ETS. Por ejemplo, se podría restringir a los controles para que fueran individuos a los cuales el Ministerio de Salud local les haya informado que eran un contacto sexual reciente de personas a las que se les había diagnosticado una ETS. Dicha restricción del grupo de control ayudaría a garantizar que todos los controles, independientemente del uso de preservativo, estuvieran expuestos recientemente a parejas infectadas y, por lo tanto, corrieran el mismo riesgo de contraer una ETS.

Preguntas de estudio (P8.10) (continuación)

8. Falso o verdadero. Si los controles se restringieran a personas que no padecen ETS y que tienen compañeros sexuales con dichas enfermedades, es probable que los controles no expuestos estén sobrerrepresentados en el estudio cuando se comparan con los no casos no expuestos en la población de origen.
9. Falso o verdadero. Es probable que los controles expuestos estén sobrerrepresentados en el estudio, cuando se comparan con los no casos expuestos en la población de origen.
10. ¿Qué implican las respuestas a las preguntas anteriores sobre la relación entre las razones de selección γ y δ?
11. Si se supone que las razones de selección para los casos expuestos y no expuestos son iguales, ¿por qué no sería probable que haya sesgo de selección en los estudios que restringen los controles exclusivamente a personas que tienen compañeros sexuales con ETS?

Resumen

- En los estudios de casos y controles sobre ETS, el grupo de control típicamente ha sido una muestra aleatoria de personas que no padecen tales enfermedades de la población de las clínicas.
- Una crítica a dichos estudios es que la población de origen correcta está compuesta por individuos de estas clínicas que tienen compañeros sexuales con ETS.

- En dichos estudios, es probable que los controles no expuestos estén sobrerrepresentados cuando son comparados con los no casos no expuestos en la población de origen (es decir, de clínicas).
- Es probable que haya sesgo de selección cuando se permite que los controles incluyan a personas que tienen compañeros sexuales sin ETS.
- Una manera de minimizar el sesgo de selección sería redefinir el grupo de control como los individuos no infectados que, con certeza, poseen compañeros sexuales con ETS.

Sesgo de selección en los estudios de casos y controles hospitalarios

En 1946, Berkson demostró que los estudios de casos y controles en los que se usan pacientes de hospitales están sujetos a un tipo de sesgo de "selección" llamado **sesgo de Berkson**. El sesgo de Berkson surge porque los pacientes con dos enfermedades o comportamientos de alto riesgo tienen más probabilidades de ser hospitalizados que los que tienen una única enfermedad. Dichos pacientes tenderán a estar sobrerrepresentados en la población de estudio, en comparación con la población de la comunidad.

Por ejemplo, debido a que el tabaquismo está fuertemente asociado a varios tipos de cáncer, esperaríamos que un estudio en hospitales sobre la relación entre el tabaquismo y casi cualquier cáncer demuestre una relación más fuerte que la que existiría en una población de la comunidad. Para ejemplificar el sesgo de Berkson, considere la siguiente tabla de doble entrada que describe la exposición y la enfermedad para cierta población de la comunidad.

Población de la comunidad

	¿Fuman?		
Enfermedad	Sí	No	Total
Sí	400	200	600
No	100.000	200.000	300.000
Total	100.400	200.200	300.600

$\hat{OR}$ =(400 x 200.000) / (200 x 100.000) = 4,0

Muestra de casos y controles basada en hospitales

	¿Fuman?		
	Sí	No	Total
Casos	225	75	300
Controles	100	200	300
Total	325	275	600

Suponga que de los 600 casos, 300 ó el 50% se seleccionan en varios hospitales para un estudio de casos y controles. Suponga además que 225 ó el 75% de estos 300 casos son fumadores. Suponga que hay 300 controles seleccionados como una muestra del 0,1% de todos los no casos de la comunidad. Suponiendo que no hay sesgo o error aleatorio al elegir estos controles, esperaríamos, por lo tanto, 100 controles expuestos y 200 controles no expuestos, como se mostró anteriormente.

Preguntas de estudio (P8.11)

1. ¿Cuáles son las poblaciones de origen y de estudio para este estudio?
2. ¿Cuáles son los valores de α, β, γ y δ para este estudio?
3. ¿α, β, γ y δ son probabilidades de selección?
4. ¿Por qué hay sesgo de selección en el cálculo del *odds ratio*? Explique en términos de razones de selección.
5. ¿Cuál es la dirección del sesgo?
6. ¿Por qué éste es un ejemplo del sesgo de Berkson?

En el ejemplo anterior, los casos provinieron del hospital pero los controles provinieron de la comunidad. Para compensar la elección de casos hospitalizados, las tablas que aparecen a continuación muestran los resultados de un estudio en el que se eligieron 300 no casos hospitalizados como controles.

Población de la comunidad

	¿Fuman?		
Enfermedad	Sí	No	Total
Sí	400	200	600
No	100.000	200.000	300.000
Total	100.400	200.200	300.600

Muestra de casos y controles basada en hospitales

	¿Fuman?		
	Sí	No	Total
Casos	225	75	300
Controles	150	150	300
Total	375	225	600

7. ¿Los controles en el estudio anterior están sobrerrepresentando a los no casos en la comunidad?
8. ¿Cuáles son los valores de α, β, γ y δ para este estudio?
9. En términos de probabilidades de selección, ¿por qué hay sesgo de selección en el cálculo del *odds ratio*?
10. ¿Cuál es la dirección del sesgo?
11. ¿Por qué éste es un ejemplo del sesgo de Berkson?

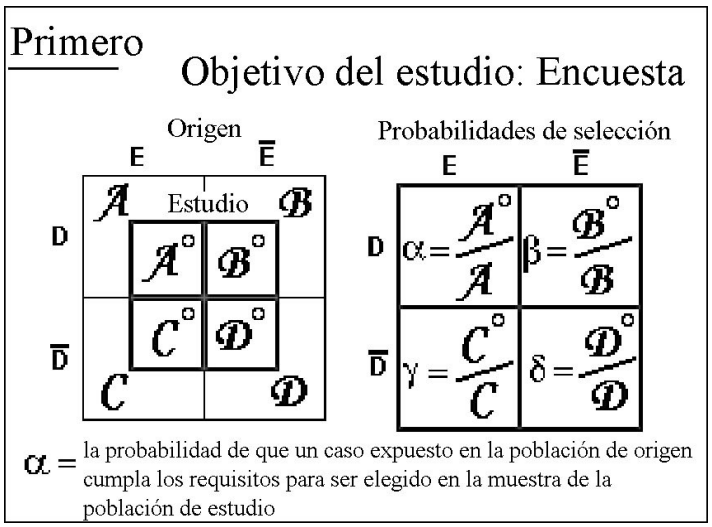
Resumen

- Berkson demostró que los estudios de casos y controles en los que se usan pacientes hospitalizados están sujetos a un tipo de sesgo de selección llamado sesgo de Berkson.
- Los pacientes con dos enfermedades o comportamientos de alto riesgo tienen más probabilidades de ser hospitalizados que los que tienen una única enfermedad.
- Los pacientes hospitalizados en una población de estudio tenderán a estar sobrerrepresentados, en comparación con una población de la comunidad.
- La dirección del sesgo de Berkson puede ser hacia el valor nulo o alejarse de éste, según en qué medida los casos hospitalizados estén más o menos sobrerrepresentados que los controles hospitalizados, en comparación con la comunidad.

El sesgo de selección en los estudios transversales

En los estudios transversales hay dos maneras distintas de considerar el sesgo de selección. Si el objetivo del estudio es principalmente analizar una población que se supone estable en un punto del tiempo o en un intervalo corto, entonces la población de origen es el subconjunto de una población transversal supuestamente estable que estaba en riesgo de padecer la enfermedad antes del comienzo del estudio. Entonces, el sesgo de selección puede ocurrir cuando algún aspecto del proceso de selección, como una falta de respuesta, distorsiona suficientemente la población de estudio como para sesgar el cálculo.

Aquí, la muestra del estudio es un subconjunto de la población de origen, por lo cual los parámetros claves de interés son probabilidades de selección. Alfa (α) denota la probabilidad de que un caso expuesto en la población de origen transversal cumpla los requisitos para ser elegido en la muestra de la población de estudio. Dicha probabilidad está dada por A^o dividido por A, las frecuencias de celda respectivas en la población de origen y en la población de estudio. β, γ y δ se definen de manera similar.



El sesgo de selección en el *odds ratio* (la medida de efecto de interés en un estudio transversal) puede ocurrir cuando el producto cruzado de las probabilidades de selección [(αδ)/(βγ)] difiere relevantemente de la unidad.

Sin embargo, si el objetivo del estudio es determinar si hay una relación etiológica entre la exposición y la enfermedad, entonces, la población de origen de interés ya no es la población transversal que se está analizando. En lugar de ello, la población de origen adecuada está dada por una cohorte identificada en un punto del tiempo anterior que contiene las personas en riesgo de desarrollar la enfermedad al momento del estudio transversal. En otras palabras, un estudio transversal brinda datos de prevalencia de los que deseamos hacer inferencias sobre datos de incidencia. En esta situación, la fuente principal del sesgo de selección es lo que se denomina **supervivencia selectiva**. (Obsérvese que la supervivencia selectiva es un problema importante en las enfermedades con alto riesgo de mortalidad, como, por ejemplo, la enfermedad coronaria, pero es un problema menor en las enfermedades con bajo riesgo de mortalidad, como, por ejemplo, la artritis). Los individuos en los estudios transversales son sobrevivientes, especialmente aquellos que padecen la enfermedad, ya sean expuestos o no expuestos.

Las probabilidades de selección α, β, γ y δ dan las probabilidades de sobrevivir desde el momento T_0 hasta el momento T_1. Aquí puede ocurrir un sesgo de selección si la probabilidad de sobrevivir el tiempo necesario para ser incluido en el estudio transversal es diferente para las cuatro celdas de la cohorte de la población de origen.

Para ejemplificar la supervivencia selectiva, la siguiente tabla (en la parte superior izquierda) describe la incidencia de la enfermedad desde el momento T_0 hasta el momento T_1 para el seguimiento de una cohorte fija de

1000 personas que no padecían la enfermedad en el momento T_0. Comparamos los resultados de los datos de incidencia anteriores con los resultados esperados de un estudio transversal realizado sobre los sobrevivientes. La población de estudio de los sobrevivientes en el momento T_1 se muestra en la siguiente tabla en la parte inferior derecha.

Incidencia para cohorte fija de T_0 a T_1

	E	no E	Total
D	50	80	130
no D	150	720	870
Total	200	800	1000

RR = 2.5, ROR = 3

Resultados esperados en el momento T1

	E	No E	Total
D	25	40	65
no D	142.5	540	682.5
Total	162.5	580	747.5

PR = 2.2, POR = 2.4

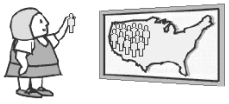
Preguntas de estudio (P8.12)

1. Para los individuos expuestos y no expuestos en forma separada, ¿qué proporción de personas que desarrollan la enfermedad se espera que sobreviva hasta el momento T_1?
2. ¿Qué proporción de personas expuestas que no desarrollan la enfermedad se espera que sobreviva hasta el momento T_1?
3. ¿Qué proporción de personas no expuestas que no desarrollan la enfermedad se espera que sobreviva hasta el momento T_1?
4. ¿Cuáles son las 4 probabilidades de selección para la supervivencia selectiva?
5. ¿Cuál es el producto cruzado de probabilidades de selección?
6. ¿Hay sesgo de selección tanto en el *odds ratio* como en la razón de riesgos?

Resumen

- En los estudios transversales hay dos maneras distintas de considerar el sesgo de selección, según el objetivo del estudio.
- Si el objetivo del estudio es analizar una población que se supone estable en un punto del tiempo o en un intervalo corto, entonces puede ocurrir sesgo de selección debido a una falta de respuesta y/o alguna otra distorsión selectiva de la población de estudio.
- Si el objetivo del estudio es determinar si hay una relación etiológica entre la exposición y la enfermedad, entonces la fuente principal del sesgo de selección es la supervivencia selectiva.
- La supervivencia selectiva puede ocurrir si la probabilidad de sobrevivir el tiempo necesario para ser incluido en el estudio transversal es diferente para las cuatro celdas de la cohorte de la población de origen.

Auto-evaluación (P8.13)



1. Si el objetivo del estudio es principalmente analizar una población que se supone estable en un punto del tiempo o en un intervalo corto, entonces la población de origen es el subconjunto de una población transversal que se supone estable y que corría riesgo de padecer la enfermedad **???** del comienzo del

estudio.

2. Entonces, el sesgo de selección puede ocurrir cuando algún aspecto del proceso de selección, como **???**, distorsiona suficientemente la **???** como para sesgar el cálculo.
3. El sesgo de selección en el *odds ratio* (la medida de efecto de interés en un estudio transversal) puede ocurrir cuando el producto cruzado de las probabilidades de selección difiere relevantemente de **???**.

Opciones
100 después falta de respuesta uno antes supervivencia selectiva población de origen población de estudio cero

No obstante, si el objetivo del estudio es determinar si hay una relación etiológica entre la exposición y la enfermedad, la población de origen de interés ya no es la población transversal que se está analizando.

4. En lugar de ello, la población de origen adecuada está dada por una cohorte identificada en un punto del tiempo **???** que contiene a las personas en riesgo de desarrollar la enfermedad **???** en el momento del estudio transversal.
5. En otras palabras, un estudio transversal brinda **???** de los que deseamos hacer inferencias sobre **???**.

Opciones
después al anterior datos de incidencia posterior datos de prevalencia

Considere nuevamente un estudio transversal que tenga como objetivo determinar si existe una relación etiológica entre la exposición y la enfermedad.

6. En esta situación, la fuente principal del sesgo de selección es lo que se denomina **???**.
7. Los individuos en los estudios transversales son **???**, especialmente aquellos que padecen la enfermedad, ya sean expuestos o no expuestos.
8. Los que no **???** no son incluidos en el estudio.

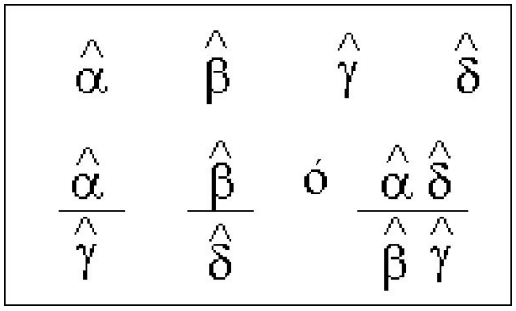
Opciones
casos mueren falta de respuesta supervivencia selectiva sobreviven sobrevivientes

8-4 Sesgo de selección (continuación)

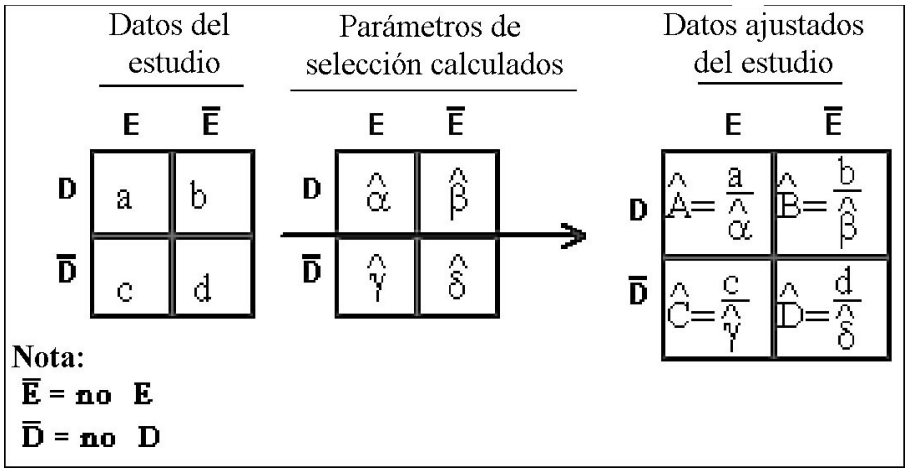
¿Qué se puede hacer cuantitativamente con respecto al sesgo de selección?

Suponga que está convencido de que un estudio ya diseñado y realizado tiene imperfecciones debido a un sesgo de selección. ¿Es posible corregir cuantitativamente dicho sesgo? Por ejemplo, ¿puede ajustarse el valor de un *odds ratio* sesgado para brindar un valor modificado que no posea sesgo de selección?

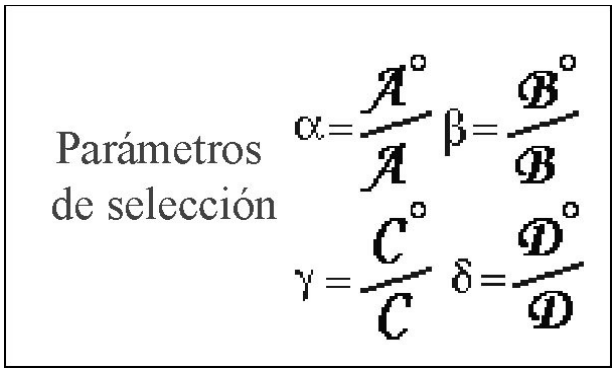
La respuesta aquí depende de si pueden determinarse cálculos confiables de los parámetros de las razones de selección subyacentes. En verdad, incluso si estos parámetros no pueden calcularse en forma individual, es posible corregir el sesgo de selección si podemos calcular razones de estos parámetros de selección como, por ejemplo, α sobre γ, β sobre δ o simplemente la razón de productos cruzados de los parámetros de selección.



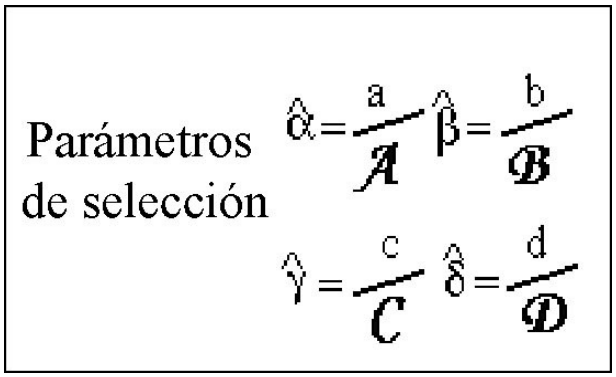
Aquí se muestran tres cuadros de 2 x 2 que describen los datos del estudio, los parámetros de selección calculados y los datos ajustados del estudio que debemos considerar si queremos ajustar el sesgo de selección.



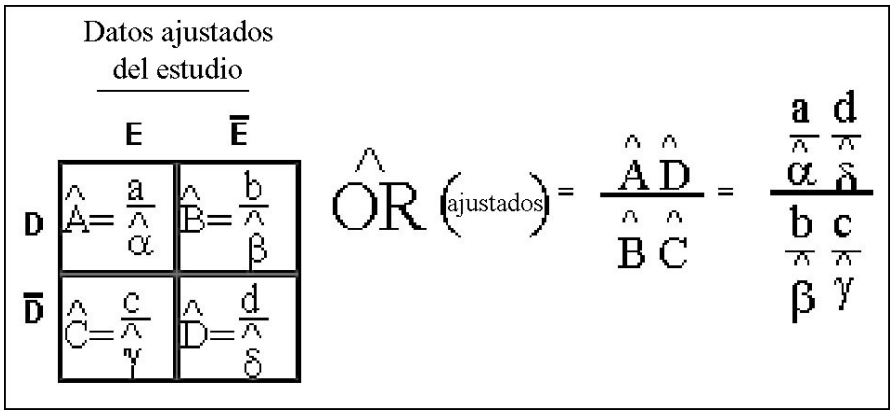
El cuadro de los datos ajustados del estudio se deriva utilizando definiciones previas de los parámetros de selección como razones correspondientes a las frecuencias de celda de las poblaciones de estudio y de origen.



Este cuadro se obtiene sustituyendo A^o por a, B^o por b, etc. y luego resolviendo las letras mayúsculas A, B, C y D en términos de las letras minúsculas a, b, c y d, dividido por sus valores correspondientes de los parámetros de selección.



Un cálculo ajustado del *odds ratio* que corrige el sesgo de selección está dado, entonces, por el producto cruzado de las frecuencias de celda ajustadas, que pueden escribirse en términos de los datos del estudio y los valores calculados de los parámetros de selección como se muestra aquí.



Por lo tanto, el *odds ratio* que ajusta el sesgo de selección está dado por el *odds ratio* calculado a partir de los datos del estudio dividido por el producto cruzado de los parámetros de selección.

$$\hat{OR}_{\text{ajustado}} = \frac{\hat{OR}}{\left(\frac{\hat{\alpha}\,\hat{\delta}}{\hat{\beta}\,\hat{\gamma}}\right)}$$

Esta fórmula de ajuste puede escribirse, alternativamente, mediante las razones de los parámetros de selección como se muestra a continuación.

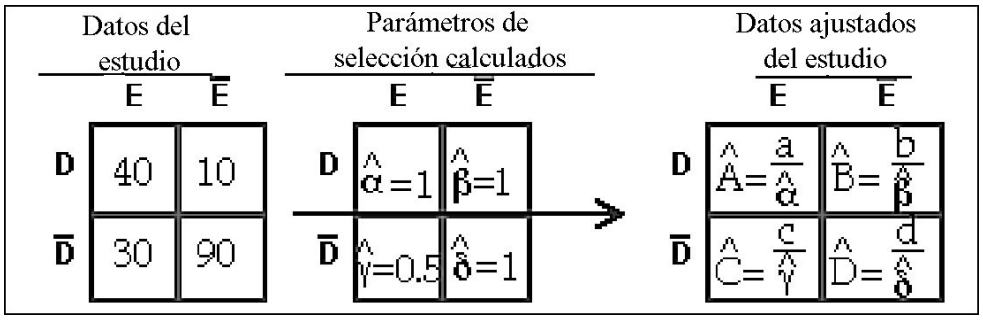
$$= \hat{OR} \times \frac{\hat{r}_{\bar{D}}}{\hat{r}_{D}} = \hat{OR} \times \frac{\hat{r}_{\bar{E}}}{\hat{r}_{E}}$$

donde $\hat{r}_{D} = \frac{\hat{\alpha}}{\hat{\beta}}$ $\hat{r}_{E} = \frac{\hat{\alpha}}{\hat{\gamma}}$

$\hat{r}_{\bar{D}} = \frac{\hat{\gamma}}{\hat{\delta}}$ $\hat{r}_{\bar{E}} = \frac{\hat{\beta}}{\hat{\delta}}$

Las fórmulas alternativas permiten la corrección del sesgo incluso si sólo pueden determinarse las razones de los parámetros de selección, en lugar de los parámetros de selección individuales.

Preguntas de estudio (P8.14)



Considere los tres cuadros anteriores de **datos del estudio, valores calculados de los parámetros de selección** y **datos ajustados del estudio**.

1. Calcule las frecuencias ajustadas de celda $\hat{A}, \hat{B}, \hat{C}$ y $\hat{D}$.
2. ¿Cuál es el *odds ratio* ajustado que corrige el sesgo de selección?
3. Calcule las razones de los parámetros de selección $\hat{r}_{D} = \hat{\alpha}/\hat{\beta}$ y $\hat{r}_{\bar{D}} = \hat{\lambda}/\hat{\delta}$.
4. Basándose en $\hat{r}_{D}$ y $\hat{r}_{\bar{D}}$, utilice la fórmula de corrección alternativa para calcular el *odds ratio* que corrige

el sesgo de selección.

También es posible un cálculo ajustado de la razón de riesgos si pueden identificarse los valores calculados de los parámetros de selección. Del cuadro de 2 x 2 de frecuencias ajustadas de celda podemos obtener una razón de riesgos ajustada, como se muestra a continuación.

$$\hat{RR}_{(ajustado)} = \frac{\frac{\hat{A}}{\hat{A}+\hat{C}}}{\frac{\hat{B}}{\hat{B}+\hat{D}}}$$

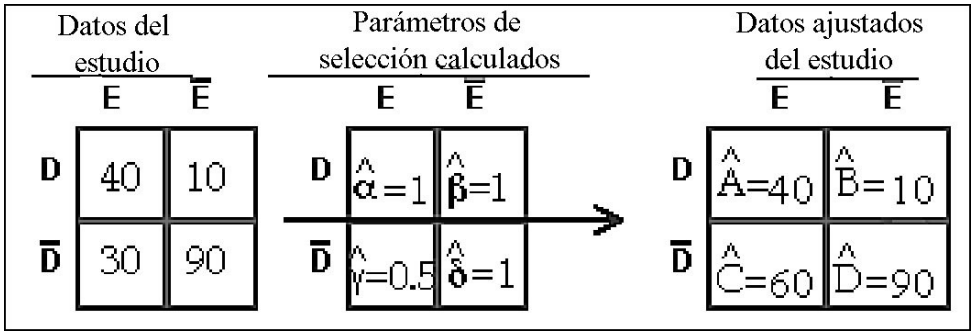
Si luego sustituimos $\hat{A}, \hat{B}, \hat{C}$ y $\hat{D}$ F por sus expresiones correspondientes que involucran datos del estudio dividido por el valor calculado del parámetro de selección, y realizamos algunas manipulaciones algebraicas, obtenemos la fórmula de la razón de riesgos ajustada que se muestra a continuación.

$$RR_{ajustado} = \frac{\left(\frac{a}{a + c\frac{\hat{\alpha}}{\hat{\gamma}}}\right)}{\left(\frac{b}{b + d\frac{\hat{\beta}}{\hat{\delta}}}\right)}$$

Puede escribirse una fórmula alternativa incluyendo la razón de los parámetros de selección $\hat{r}_E$ y $\hat{r}_{\bar{E}}$, como se muestra aquí.

$$\hat{RR}_{(ajustado)} = \frac{\frac{a}{a + c\hat{r}_E}}{\frac{b}{b + d\hat{r}_{\bar{E}}}}$$

Preguntas de estudio (P8.14) (continuación)



Considere nuevamente los tres cuadros anteriores de **datos del estudio, valores calculados de los parámetros de selección** y **datos ajustados del estudio**.

5. Si tomamos en cuenta la información anterior, ¿cuál es la razón de riesgos ajustada que corrige el sesgo de selección?
6. ¿Cuáles son los valores de las razones $\hat{\alpha}/\hat{\gamma}$ y $\hat{\beta}/\hat{\delta}$?
7. Utilice la fórmula para corregir el sesgo de selección en la razón de riesgos a fin de obtener la razón de riesgos ajustada.

Resumen

- El sesgo de selección puede corregirse cuantitativamente si se pueden calcular los cuatro parámetros de selección o las razones de estos parámetros.
- Una fórmula del *odds ratio* ajustado es:

$$\hat{OR}_{ajustado} = \frac{\hat{OR}}{\left(\frac{\hat{\alpha}\,\hat{\delta}}{\hat{\beta}\,\hat{\gamma}}\right)}$$

- Las fórmulas alternativas son:

* Fórmulas Alternativas

$$\hat{OR}_{ajustado} = \hat{OR}\,\frac{\hat{r}_{\bar{D}}}{\hat{r}_{D}} = \hat{OR}\,\frac{\hat{r}_{\bar{E}}}{\hat{r}_{E}}$$

- El sesgo de selección puede corregirse cuantitativamente si pueden calcularse los cuatro parámetros de selección.
- A la izquierda se presenta una fórmula de la razón de riesgos ajustada y a la derecha se observa una fórmula alternativa para la $RR_{ajustada}$:

$$RR_{\text{ajustado}} = \frac{\left(\frac{a}{a + c\frac{\hat{\alpha}}{\hat{\gamma}}}\right)}{\left(\frac{b}{b + d\frac{\hat{\beta}}{\hat{\delta}}}\right)} \qquad RR_{\text{ajustado}} = \frac{\left(\frac{a}{a + c\hat{r}_E}\right)}{\left(\frac{b}{b + d\hat{r}_{\bar{E}}}\right)}$$

¿Qué se puede hacer cualitativamente con respecto al sesgo de selección?

La actividad anterior brindó fórmulas cuantitativas para corregir el sesgo de selección. Lamentablemente, la información requerida para aplicar estas fórmulas, concretamente los parámetros de las razones de selección o sus razones, es conceptualmente complicada, pocas veces está disponible y no es fácil de cuantificar. La forma ideal de abordar el sesgo de selección es prevenir o, al menos, minimizar dicho sesgo cuando se diseña un estudio, en lugar de tratar de corregir el sesgo una vez que se hayan recabado los datos.

En los estudios de casos y controles, los controles deberían seleccionarse cuidadosamente para representar a la población de origen que produjo los casos. También se debería considerar el uso de dos o más grupos de control.

Preguntas de estudio (P8.15)

1. Suponga que se utilizan dos grupos de control en un estudio de casos y controles, y que los efectos calculados resultantes difieren considerablemente según qué grupo de control se utilice. ¿Cómo se puede determinar cuál de los dos cálculos es correcto?
2. Suponga que se utilizan dos grupos de control y que los efectos calculados resultantes son básicamente iguales, independientemente de qué grupo de control se utilice. ¿Esto significa que no hay sesgo de selección?

Los estudios de casos y controles que utilizan casos incidentes, al igual que los estudios de casos y controles anidados, se prefieren a los estudios que utilizan casos prevalentes o a los estudios basados en hospitales. También se puede evitar el sesgo de selección asegurando una igualdad de oportunidades para la detección de la enfermedad entre los individuos expuestos y no expuestos.

En los estudios de cohorte, se debería intentar lograr una alta respuesta y una baja pérdida de seguimiento.

Preguntas de estudio (P8.15) (continuación)

3. Suponga que sólo el 1% de una cohorte de estudio se pierde durante el seguimiento. ¿Todavía es posible el sesgo de selección debido a una pérdida de seguimiento?
4. Suponga que se pierde el seguimiento del 20% de los individuos expuestos y del 30% de los individuos no expuestos en un estudio de cohorte. ¿Habrá sesgo de selección?

Se deberían evitar los estudios observacionales que involucren voluntarios, aunque son factibles los ensayos clínicos que incluyen voluntarios debido a la aleatoriedad. Los estudios de cohorte ocupacionales deberían evitar el sesgo por trabajador sano, garantizando que los individuos no expuestos sean tan saludables como los individuos expuestos.

En la etapa de análisis es posible determinar la dirección del sesgo, incluso si la magnitud del sesgo no

puede calcularse o no es factible una corrección numérica del mismo.

Preguntas de estudio (P8.15) (continuación)

Suponga que se pudiera plantear un argumento razonable indicando que alfa es menor que beta, mientras que gamma es igual a delta, aunque no se pudieran realizar cálculos aproximados particulares de cada parámetro de selección o de las razones de los parámetros de selección. Basándose en esta información, responda las siguientes preguntas:

5. ¿Es posible evaluar la magnitud del sesgo de selección en el *odds ratio*?
6. ¿Es posible corregir cuantitativamente el sesgo de selección en el *odds ratio*?
7. ¿Es posible determinar la dirección del sesgo?

Suponga que se pudiese argumentar que α es mayor que γ, pero que β también es mayor que γ, aunque no se pudieran realizar cálculos aproximados particulares de cada parámetro de selección o de las razones de los parámetros de selección.

8. ¿Es posible evaluar si hay sesgo de selección?
9. Suponga que se pudiese argumentar que α / γ es igual a β / δ. ¿Es posible evaluar si hay sesgo de selección en el *odds ratio*?
10. Suponga que el OR correcto es mayor que 1 y que se pudiese argumentar que α / γ es mayor que β / δ. ¿Es posible determinar la dirección del sesgo?

En la reseña del estudio se presenta un enfoque final para abordar el sesgo de selección, así como los sesgos de información y confusión. Normalmente, se brinda una descripción de los sesgos potenciales en la sección "Discusión" de una publicación.

Esta discusión, especialmente cuando se refiere a un posible sesgo de selección, es bastante subjetiva, pero se espera un juicio, dada la dificultad inherente a la cuantificación de los sesgos.

Resumen

- La información requerida para evaluar el sesgo de selección es conceptualmente complicada, está pocas veces disponible y no es fácil de cuantificar.
- En la etapa de diseño del estudio se pueden tomar decisiones para evitar el sesgo de selección en los resultados del estudio.
- En la etapa de análisis es posible determinar la dirección del sesgo de selección sin poder corregirlo cuantitativamente.
- En la etapa de la publicación, los sesgos potenciales normalmente se abordan cualitativamente en la sección "**Discusión**" de un trabajo.

El abordaje de una situación "en el peor de los casos"

Este es un abordaje práctico para evaluar la **dirección** del **sesgo de selección**; considera, en términos realistas, los cambios más extremos posibles en el cálculo de efecto como resultado de la forma en que se seleccionan los individuos. Mediante dicho abordaje, puede ser posible mostrar que el peor grado de sesgo posible tendrá un impacto mínimo o irrelevante en las conclusiones del estudio.

Por ejemplo, consideremos un estudio de cohorte que involucra a 1000 fumadores y 1000 no fumadores de

más de 40 años de edad, que no padecen cáncer de pulmón y a los que se les realiza un seguimiento durante 10 años. Suponga además que durante el período de acompañamiento, se pierde el seguimiento de 200 fumadores y 100 no fumadores. Asimismo, suponga que entre los 800 fumadores y 900 no fumadores que continúan en el estudio, el cuadro de 2 x 2 que relaciona la condición de tabaquismo al comienzo del estudio con el desarrollo de cáncer de pulmón (CP) durante los 10 años de seguimiento es el siguiente:

	Fumadores	**No fumadores**
CP	80	10
Sin CP	720	890
Total	800	900

La razón de riesgos para estos datos es (80/800)/(10/900) = 9, lo que sugiere una relación muy fuerte entre la condición de tabaquismo y el desarrollo de cáncer de pulmón. Una situación "en el peor de los casos" podría determinar cuáles serían los valores de la razón de riesgos si los 200 fumadores que se perdieron durante el seguimiento no desarrollaran cáncer de pulmón y/o los 100 no fumadores que se perdieron durante el seguimiento sí lo desarrollaran. A continuación se incluye una comparación de los cálculos para distintas situaciones de esta naturaleza:

Situación	**Razón de riesgos**
1. Datos reales observados	9
2. 1/10 de los 200 fumadores que se perdieron durante el seguimiento desarrollan CP y 1/90 de los 100 no fumadores que se perdieron durante el seguimiento desarrollan CP.	9
3. 1/10 de los 200 fumadores que se perdieron durante el seguimiento desarrollan CP y 2/90 de los 100 no fumadores que se perdieron durante el seguimiento desarrollan CP.	8,2
4. Ninguno de los 200 fumadores que se perdieron durante el seguimiento desarrolla CP y 1/90 de los 100 no fumadores que se perdieron durante el seguimiento desarrollan CP.	7,3
5. Ninguno de los 200 fumadores que se perdieron durante el seguimiento desarrolla CP y 2/90 de los 100 no fumadores que se perdieron durante el seguimiento desarrollan CP.	6,6
6. Ninguno de los 200 fumadores que se perdieron durante el seguimiento desarrolla CP y los 100 no fumadores que se perdieron durante el seguimiento desarrollan CP.	0,7

Observe que en la situación 6, el tabaquismo pasa de ser dañino a ser protector. Sin embargo, ésta no es una circunstancia muy realista. La situación 2 no es realmente una situación "en el peor de los casos" porque supone que aquellos que se perdieron durante el seguimiento corren el mismo riesgo que los que en realidad tuvieron un seguimiento durante los 10 años completos. Las otras tres situaciones, es decir, de la 3 a la 5, son "realistas" y situaciones "en el peor de los casos", ya que todas muestran que aunque la razón de riesgos se redujo, todavía es alta. La dificultad con este abordaje, como se ejemplificó anteriormente, se refiere a la medida en la que el investigador puede identificar una situación "en el peor de los casos" que sea la más realista entre todas las situaciones posibles.

Auto-evaluación (P8.16)



1. Un cálculo ajustado del *odds ratio* que corrige el sesgo de selección está dado por el producto cruzado de **???**, que puede expresarse en términos de los datos del estudio y los valores calculados de **???**.
2. Por lo tanto, el *odds ratio* que ajusta el sesgo de selección está dado por la **???** calculada de los datos del estudio dividida por el **???**.

Opciones
las frecuencias de celda ajustadas **cálculo del producto cruzado de los parámetros de selección**
odds ratio **los datos de la población** **sesgo de selección** **parámetro de selección**

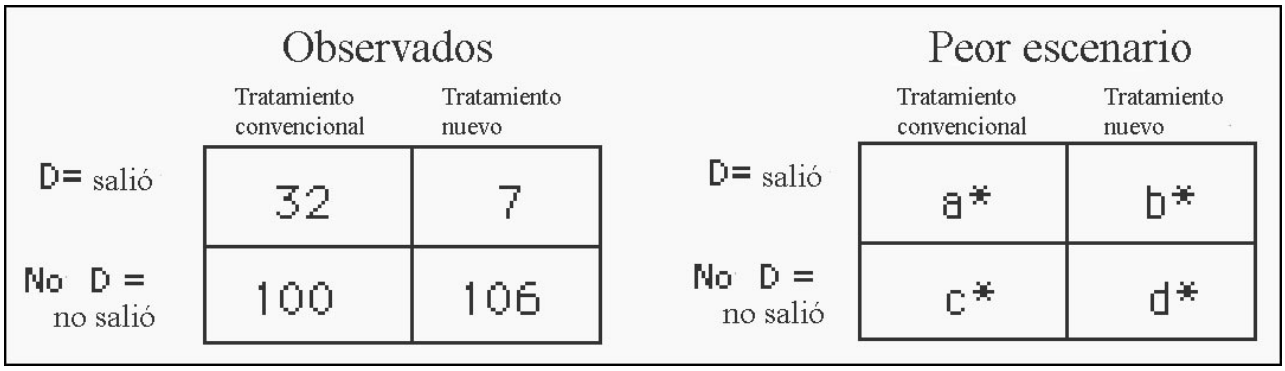
3. Los estudios de casos y controles que utilizan **???** se prefieren a los estudios que utilizan **???**.
4. Y los estudios **???** se prefieren a los estudios **???**.
5. En los estudios de cohorte se debería intentar lograr **???** y evitar **???**.

Opciones
casos **controles** **una alta respuesta** **basados en hospitales** **casos incidentes**
pérdida de seguimiento **de casos y controles anidados** **casos prevalentes** **sesgo de selección**

En la etapa de análisis se puede evaluar la medida del posible sesgo de selección utilizando lo que muchas veces se denomina análisis de una situación "en el peor de los casos". Dichos análisis consideran los cambios más extremos posibles en el efecto calculado que se deben al sesgo de selección. Determine si cada una de las siguientes afirmaciones es **verdadera** o **falsa**:

6. Este abordaje es útil, ya que podría demostrar que el peor grado de sesgo posible tendrá un impacto mínimo o irrelevante en las conclusiones del estudio. **???**
7. Este abordaje puede descartar el sesgo de selección, pero no puede confirmarlo. **???**

Los siguientes cuadros muestran los resultados observados y una situación "en el peor de los casos" para un ensayo clínico. Se perdió el seguimiento de diez individuos que recibían tratamiento convencional y 15 individuos que recibían un nuevo tratamiento. El resultado de salud bajo observación era si un individuo había salido de remisión o no (D = salió, no D = no salió) al final del ensayo. En una situación "en el peor de los casos", los 10 individuos que seguían un tratamiento convencional, y cuyo seguimiento se perdió, continuaban en remisión; por otro lado, los 15 individuos que seguían un tratamiento nuevo, y cuyo seguimiento se perdió, salieron de remisión.



8. ¿Cuáles son los valores de a*? **???**, b*? **???**, c*? **???**, d*? **???**

Opciones
106 110 121 22 32 42 90 91

9. Utilice los datos de los cuadros anteriores. ¿Cuál es el valor de la razón de riesgos sobre la base de los datos observados? **???**
10. ¿Cuál es el valor de la razón de riesgos sobre la base de la situación "en el peor de los casos"? **???**
11. ¿Hay sesgo de selección "en el peor de los casos"? **???**
12. ¿Un análisis de "en el peor de los casos" como el que se ejemplifica aquí "prueba" que hay sesgo de selección? **???**

Opciones
1,3 1,4 3,9 4,8 No Sí

Nomenclatura

Población de origen vs. población de estudio

De origen

De estudio		E	$\overline{E}$
A			**B**
D	D	**A°**	**B°**
$\overline{D}$	$\overline{D}$	**C°**	**D°**
C			**D**

Probabilidades de selección

	E	$\overline{E}$
D	**α = A°/ A**	**β = B°/ B**
$\overline{D}$	**γ = C°/ C**	**δ = D°/ D**

Datos del estudio

	E	$\overline{E}$
D	a	b
$\overline{D}$	c	d

Valores calculados de los parámetros de selección

	E	$\overline{E}$
D	$\hat{\alpha}$	$\hat{\beta}$
$\overline{D}$	$\hat{\gamma}$	$\hat{\delta}$

Datos ajustados del estudio

	E	$\overline{E}$
D	$\hat{A} = \frac{a}{\hat{\alpha}}$	$\hat{B} = \frac{b}{\hat{\beta}}$
$\overline{D}$	$\hat{C} = \frac{c}{\hat{\lambda}}$	$\hat{D} = \frac{d}{\hat{\gamma}}$

$$\text{Sesgo}\,(\hat{OR}, OR) = OR^{o} - OR = \frac{A^{o}D^{o}}{B^{o}C^{o}} - \frac{AD}{BC} = \frac{\alpha A \delta d}{\beta B \gamma C} - \frac{AD}{BC}$$

$$\hat{OR}_{\text{ajustada}} = \frac{\hat{A}\hat{D}}{\hat{B}\hat{C}} = \frac{\frac{a}{\hat{\alpha}}\frac{d}{\hat{\delta}}}{\frac{b}{\hat{\beta}}\frac{c}{\hat{\gamma}}} = \frac{\hat{OR}}{\frac{\hat{\alpha}\hat{\delta}}{\hat{\beta}\hat{\gamma}}}$$

$$\hat{RR}_{\text{ajustada}} = \frac{\frac{\hat{A}}{\hat{A}+\hat{C}}}{\frac{\hat{B}}{\hat{B}+\hat{D}}} = \frac{\frac{a}{a + c\frac{\hat{\alpha}}{\hat{\gamma}}}}{\frac{b}{b + d\frac{\hat{\beta}}{\hat{\delta}}}}$$

D	Verdaderamente padece la enfermedad
E	Verdaderamente expuesto
No D *o* $\overline{D}$	Verdaderamente no padece la enfermedad
No E *o* $\overline{E}$	Verdaderamente no expuesto
$\hat{\mathrm{OR}}$	*Odds ratio* basado en los datos observados
OR^{o}	*Odds ratio* basado en la población de estudio
$\hat{\mathrm{OR}}_{\text{ajustado}}$	*Odds ratio* basado en datos corregidos o ajustados
$\hat{\mathrm{RR}}$	Razón de riesgos basada en los datos observados
OR^{o}	Razón de riesgos basada en la población de estudio
$\hat{\mathrm{RR}}_{\text{ajustada}}$	Razón de riesgos basada en datos corregidos o ajustados

Referencias

Referencias sobre sesgo de selección

Berkson J. *Limitations of the application of fourfold table analysis to hospital data*. Biometrics Bulletin 1946; 2: 47-53.

Greenberg RS, Daniels SR, Flanders WD, Eley JW, Boring JR. *Medical Epidemiology* (3rd Ed). Lange Medical Books, Nueva York, 2001.

Hill H, Kleinbaum DG. *Bias in Observational Studies*. In Encyclopedia of Biostatistics, pág. 323-329, Oxford University Press, 1999.

Horwitz RI, Feinstein AR. *Alternative analytic methods for case-control studies of estrogens and endometrial cancer*. N Engl J Med 1978; 299(20):1089-94.

Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic Research: Principles and Quantitative Methods*. John Wiley and Sons Publishers, Nueva York, 1982.

Referencias sobre estudios acerca de la efectividad del preservativo

Crosby RA. *Condom use as a dependent variable: measurement issues relevant to HIV prevention programs*. AIDS Educ Prev 1998; 10(6): 548-57.

Fishbein M, Pequegnat W. *Evaluating AIDS prevention interventions using behavioral and biological outcome measures*. Sex Transm Dis 2000; 37 (2):101-10.

Peterman TA, Lin LS, Newman DR, Kamb ML, Bolan G, Zenilman J, Douglas JM Jr, Rogers J, Malotte CK. *Does measured behavior reflect STD risk? An analysis of data from a randomized controlled behavioral intervention study*. Project RESPECT Study Group. Sex Transm Dis 2000; 27(8):446-51.

National Institute of Allergy and Infectious Disease, *Workshop Summary: Scientific Evidence on Condom Effectiveness for Sexually Transmitted Disease (SID) Prevention*, 2001.

Stone KM, Thomas E, Timyan J. *Barrier methods for the prevention of sexually transmitted diseases*. En: Holmes KK, Sparling PF, Mardh P-A, eds. Sexually Transmitted Diseases, 3rd ed., McGraw-Hill, Nueva York, 1998.

Warner L, Clay-Warner J, Boles J, Williamson J. *Assessing condom use practices. Implications for evaluating method and user effectiveness*. Sex Transm Dis 1998; 25(6):273-7.

Warner DL, Hatcher RA. *A meta-analysis of condom effectiveness in reducing sexually transmitted HIV*. Soc Sci Med 1994; 38(8):1169-70.

Ejercicios

EPI-1. Sesgo de selección: grasa en la dieta vs. cáncer de colon

Suponga que un investigador está interesado en descubrir si el consumo de grasa está asociado al desarrollo de cáncer de colon. El investigador identifica varios cientos de casos incidentes de cáncer de colon y un número comparable de controles de la comunidad. Cada individuo llena un cuestionario de frecuencia de alimentación, describiendo sus hábitos alimenticios durante el año anterior. La información del cuestionario se utiliza para dividir a los individuos en los que tuvieron una dieta alta en grasas (expuestos) y los que tuvieron una dieta baja en grasas (no expuestos).

Suponga lo siguiente:

- Entre los casos, la posibilidad de aceptar participar en el estudio NO está relacionada con la exposición.
- Entre los controles, aquellos con dieta alta en grasas tienen la mitad de probabilidad de participar, comparados con los que consumen una dieta baja en grasas.

a. Exprese cada uno de los supuestos anteriores en términos de probabilidades de selección.
b. Calcule los *odds* de selección para los que padecen la enfermedad (r_D) y los *odds* de selección para los que no la padecen ($r_{no\ D}$).
c. Basándose en sus respuestas a los ítems a) y b), ¿espera encontrar sesgo en el *odds ratio* (OR) calculado sobre la base de este estudio? Justifique su respuesta.

Suponga que la siguiente tabla refleja los datos que se habrían observado si hubiese habido 100% de participación entre los casos y los controles (es decir, éstos son datos no sesgados):

	Dieta alta en grasas	**Dieta baja en grasas**	**Total**
Casos	200	175	375
Controles	150	225	375
Total	350	400	750

d. Basándose en la información sobre las probabilidades de selección descritas anteriormente para este estudio, especifique el cuadro de 2 x 2 que brinda la población de estudio.
e. Calcule y compare el *odds ratio* en la población de origen (es decir, 100% de participación) con el *odds ratio* en la población de estudio. ¿Hay sesgo?, y si lo hay, ¿cuál es su dirección?

EPI-2. Sesgo de selección: alergias a los alimentos vs. desarrollo de enfermedad coronaria (EC)

Considere un seguimiento durante diez años de una cohorte fija que no padece EC al inicio del seguimiento. El objetivo del estudio es determinar si las personas que padecen alergias a los alimentos (AA) antes del comienzo del seguimiento tienen un mayor riesgo de desarrollar EC. Suponga que un análisis simple de los resultados del estudio brinda la siguiente información.

	AA		
EC	**Sí**	**No**	**Total**
Sí	49	286	335
No	1378	5816	7194
Total	1427	6102	7529

RR = 0,72, OR = 0,72

¿Cuál de las siguientes afirmaciones sobre un posible sesgo de selección es adecuada para este diseño del estudio? (Elija sólo una respuesta).

a. Si las personas que tienen presión arterial alta (factor de riesgo conocido en la EC) antes del comienzo del seguimiento fueron excluidas del grupo de comparación (sin alergia), pero NO del grupo de AA (con alergia), entonces debería preocupar la presencia de un sesgo de selección que SE ALEJARÍA DEL VALOR NULO.
b. Si la probabilidad de que no se pierda el seguimiento de las personas con alergias a los alimentos es exactamente la misma probabilidad de que no se pierda el seguimiento de las personas sin alergias a los alimentos, entonces no puede haber sesgo de selección debido a la pérdida de seguimiento.
c. El sesgo por supervivencia selectiva debería causar preocupación en este estudio.
d. El sesgo de Berkson debería causar preocupación en este estudio.
e. Si las personas que no padecían ninguna enfermedad (excepto AA) antes del comienzo del seguimiento fueron excluidas tanto del grupo expuesto como del grupo de comparación (sin alergia), entonces debería preocupar la presencia de un sesgo de selección que sería HACIA EL VALOR NULO.

EPI-3. Supervivencia selectiva

a. ¿Qué significa "sesgo debido a la supervivencia selectiva" en los estudios transversales? (En su respuesta, asegúrese de definir los parámetros de probabilidades de selección adecuados).
b. ¿Bajo qué circunstancias podría no haber sesgo por supervivencia selectiva incluso si las probabilidades de selección no son todas iguales?
c. Suponga que puede evaluar que la dirección del posible sesgo por supervivencia selectiva en su estudio es hacia el valor nulo. Si los datos de su estudio dieron un *odds ratio* no estadísticamente significativo de 1,04, ¿sería correcto concluir que no había asociación entre la exposición y la enfermedad en su población de origen? Desarrolle su respuesta.

EPI-4. Elección de los casos y los controles

Se le pide que aconseje a un médico clínico interesado en realizar dos estudios de casos y controles para evaluar la relación entre el consumo de café y:

i. el cáncer primario de ovario.
ii. la enfermedad arterial coronaria.

El médico clínico desea utilizar las admisiones hospitalarias como fuente de casos. Indique cuánto más válida es la propuesta, especialmente considerando posibles sesgos de selección. ¿Cómo lo aconsejaría con respecto a la elección de los casos y los controles?

EPI-5. Mortalidad y clofibrato: ¿sesgo de selección?

Se realizó un ensayo aleatorizado controlado con placebo para calcular el efecto del clofibrato (droga que hace disminuir el colesterol) en la mortalidad a 5 años en hombres de 30 a 64 años de edad que habían sufrido un infarto de miocardio en forma reciente. La mortalidad en el grupo al que se le administró clofibrato fue 18,2% y en el grupo de control fue 19,4%. Esta diferencia no fue estadísticamente significativa. El análisis posterior mostró que aproximadamente un tercio de los hombres a los que se les suministró clofibrato no tomaron su tratamiento tal como fue indicado. Si se omite a estos hombres del análisis, hay un efecto estadísticamente significativo del clofibrato, con una mortalidad de 15,0% en el grupo con clofibrato.

a. ¿Cuál de los dos análisis alternativos anteriores es el más apropiado? ¿Por qué?
b. ¿Éste es un problema de sesgo de selección? Desarrolle su respuesta.

EPI-6. Identificación de los individuos del estudio

En un estudio titulado *Antidepressant Medication and Breast Cancer Risk* (Medicación antidepresiva y riesgo de padecer cáncer de mama, Amer. J. of Epi, fines de la década de los noventa), se indicó en la sección de metodología que "Los casos se componían de una muestra aleatoria estratificada por edad (< 50 y ≥ 50 años) de mujeres de 25 a 74 años a las que se les diagnosticó cáncer primario de mama durante 1995 y 1996 (informe de patología confirmado) y que se encontraban inscritas en el Registro de cáncer de Ontario con base poblacional. Como la supervivencia por 1 año al cáncer de mama es de 90%, no se utilizaron encuestados sustitutos. Se tomó una muestra aleatoria de controles de la población de 25 a 74 años de las listas de evaluación de propiedades del Ministerio de Finanzas de Ontario; esta base de datos incluye todos los propietarios de viviendas y los inquilinos, con indicación de la edad, sexo y domicilio".

a. Analice el abordaje de los autores con respecto a la identificación de los casos en términos del potencial de sesgo de selección.
b. Analice el abordaje de los autores con respecto a la identificación de los controles en términos del potencial de sesgo de selección.
c. ¿La decisión de no permitir encuestados sustitutos afecta potencialmente el sesgo de selección?

EPI-7. Sesgo de selección en los estudios observacionales

Las siguientes preguntas están relacionadas con la evaluación del sesgo de selección en los estudios observacionales donde se investiga la supuesta asociación entre el consumo de sal y la hipertensión. Se plantea como hipótesis que las personas que consumen mayor cantidad de sal tienen mayor riesgo de desarrollar hipertensión crónica. Se supone que ambas variables están divididas en dos partes: alto y bajo consumo de sal; hipertenso (presión arterial diastólica ≥95 y presión arterial sistólica ≥160) y normotenso. Al considerar las siguientes preguntas, ignore la posibilidad de sesgo de información o de confusión. Asimismo, suponga que la medida de efecto en la población de origen es mayor que 1.

a. Considere un estudio de casos y controles en el cual se incluye igual número de hipertensos (casos) y normotensos (controles). Suponga que los hipertensos del estudio surgen de una clínica de pacientes externos dedicada a la realización de cribado o tamizaje para el diagnóstico de hipertensos en la población, utilizando la clínica en forma voluntaria. Los normotensos se seleccionan en forma aleatoria de la población circundante a la que atiende la clínica. Suponga que los hipertensos con alto consumo de sal tienen una posibilidad cuatro veces mayor de que se les realice un tamizaje que los hipertensos con bajo consumo de sal en la clínica, y que la muestra de normotensos de la comunidad circundante es, de hecho, una muestra representativa. ¿Qué puede decirse sobre el sesgo de selección en el cálculo del *odds ratio* de exposición (EOR, por sus siglas en inglés)? (Formule su respuesta en términos de las probabilidades de selección α, β, γ y δ, o de los *odds* de selección r_D y $r_{no\ D}$).
b. En un tipo de estudio de casos y controles similar al descrito en el ítem a), suponga que los hipertensos con alto y bajo consumo de sal tienen la misma probabilidad de que se les realice un tamizaje, pero que los normotensos de la comunidad con alto consumo de sal tienen más probabilidad de cooperar en el estudio que los normotensos con bajo consumo de sal. ¿Qué puede decirse sobre el sesgo de selección en el cálculo del EOR? (Formule su respuesta en términos de las probabilidades de selección α, β, γ y δ, o de los *odds* de selección r_D y $r_{no\ D}$).

Respuestas a las preguntas de estudio y auto-evaluaciones

P8.1

1. 2
2. 1. Es probable que los casos expuestos estén sobrerrepresentados en el estudio, cuando son comparados con los casos no expuestos.

P8.2

1. Los abandonos del estudio pueden distorsionar la muestra final, en comparación con la muestra aleatoria obtenida al inicio del ensayo.
2. Los que abandonaron el estudio tienen un resultado desconocido y, por lo tanto, no se pueden analizar.
3. Una población general de pacientes con el cáncer especificado y que cumple los requisitos para recibir el tratamiento convencional o los tratamientos nuevos.
4. La muestra (esperada), ignorando el error aleatorio que se obtendría después de los abandonos de la muestra aleatoria seleccionada inicialmente para el ensayo.
5. La población de origen es la cohorte de la población de la cual se derivarían los casos si se hubiese realizado un estudio de cohorte adecuado.
6. La población de estudio es la muestra esperada obtenida de la muestra transversal que se mantiene para el análisis. En forma alternativa, la población de estudio es la población estable de la cual se obtiene la muestra transversal que se estudiará.

P8.3

1. No es probable que a las mujeres en este grupo de control (por ejemplo, con cáncer cervical o de ovarios) se les realice un tamizaje en forma selectiva para el diagnóstico del cáncer a partir del sangrado vaginal causado por el uso de estrógenos.
2. Se aleja del valor nulo debido al tamizaje selectivo de los casos, pero no de los controles. Esto daría una proporción demasiado alta de usuarias de estrógenos entre los casos, pero un cálculo correcto de la proporción de usuarias de estrógenos entre los controles.
3. Porque las que tienen sangrado vaginal por el uso de estrógenos tendrán mayor probabilidad de que se les detecte el tumor benigno en el endometrio que las no usuarias con dichos tumores endometriales.
4. Se espera que el empleo de tumores benignos de endometrio como grupo de control compense el tamizaje selectivo de los casos. Sin embargo, no queda claro que el alcance del tamizaje selectivo sería el "mismo", tanto para los casos como para los controles.
5. Un tumor benigno en el endometrio no se detecta fácilmente sin sangrado vaginal. Por lo tanto, es más probable que se detecten los tumores en los controles con tumores benignos de endometrio que utilizan estrógenos que en los controles que no los utilizan.
6. Hacia el valor nulo debido a que habría tamizaje selectivo de los controles, pero no de los casos. Esto daría una proporción demasiado alta de usuarias de estrógenos entre los controles, pero un cálculo correcto de la proporción de usuarias de estrógenos entre los casos.
7. Porque no es probable que haya tamizaje selectivo para la detección de los casos de control (que padecen otros cánceres ginecológicos), cuando se compara a las usuarias de estrógenos con las no usuarias.

P8.4

1. No, ya que la razón de riesgos para la muestra esperada es 3, que es igual a la razón de riesgos en la población de origen.
2. La población de estudio, porque la muestra que se acaba de describir brinda la cantidad esperada de individuos que se obtuvieron en la muestra final.
3. La fuente es la población de la cual se seleccionó la muestra inicial del 10% obtenida antes del seguimiento. Esta es la población de individuos de la cual se derivaron los casos.
4. No hay sesgo de selección porque RR=3 en la población de estudio, igual que en la población de origen.
5. Sí, porque la razón de riesgos en la población de estudio de 3,9 es un poco más alta que (3) en la población de origen, como resultado de

los individuos que se pierden durante el seguimiento.
6. No. El riesgo para las personas expuestas es 150/10.000 ó 0,0150 en la población de origen y 14/800 = 0,0175 en la muestra.
7. No. El riesgo para las personas no expuestas es 50/10.000 ó 0,0050 en la población de origen y 4/900 = 0,0044 en la muestra.
8. No. Sólo habrá sesgo de selección si la pérdida de seguimiento da como resultado riesgos de enfermedad en los grupos expuestos y/o no expuestos que difiere en la muestra final, cuando se le compara con la cohorte original.

P8.5

1. Los trabajadores tienden a estar más sanos que los individuos de la población general y, por lo tanto, pueden tener un resultado más favorable, independientemente de su condición de exposición.
2. Los voluntarios pueden tener características distintas de quienes no son voluntarios. La población de estudio aquí está restringida a los voluntarios, mientras que la población de origen está basada en la población, por ejemplo, una comunidad.
3. Hay falta de validez externa al aplicar las conclusiones de una población de origen de voluntarios a una población externa que está basada en la población.
4. Sí. Los estudios basados en clínicas pueden conducir a conclusiones falsas, porque los pacientes de dichos establecimientos tienden a padecer una enfermedad más severa que las personas en una muestra basada en la población.

P8.6

1. $\boldsymbol{\alpha}$ = 12/150 = 0,08; $\boldsymbol{\beta}$ = 4,5/50 = 0,09; $\boldsymbol{\gamma}$ = 788/9850 = 0,08; y $\boldsymbol{\delta}$ = 895,5/9950 = 0,09.
2. No. El *odds ratio* tanto en la población de origen como en la población de estudio es de 4,03. La razón de riesgos tanto en la población objetivo como en la población de estudio es de 3,0.
3. $(\alpha \times \delta) / (\beta \times \gamma) = (0,08 \times 0,09) / (0,09 \times 0,08) = 1$. Este resultado ilustra una regla (descrita en la próxima actividad) que dice que la razón de productos cruzados de las probabilidades de selección es igual a 1 si no hay sesgo en el *odds ratio*.
4. $\boldsymbol{\alpha}$ = 14/150 = 0,093; $\boldsymbol{\beta}$ = 4/50 = 0,08; $\boldsymbol{\gamma}$ = 786/9850 = 0,08; y $\boldsymbol{\delta}$ = 896/9950 = 0,09.
5. Sí. El *odds ratio* en la población de origen es de 3,03 y en la población de estudio es de 4,00. La razón de riesgos en la población de origen es de 3,0 y en la población de estudio es de 3,9.
6. $(\alpha \times \delta) / (\beta \times \gamma) = (0,093 \times 0,09) / (0,08 \times 0,08) = 1,3$. Este resultado ilustra una regla (descrita en la próxima actividad) que dice que la razón de productos cruzados de las probabilidades de selección será distinta de 1 si hay sesgo en el *odds ratio*.

P8.7

1. No. La razón de productos cruzados de las razones de selección es igual a 1.
2. Sí. La razón de productos cruzados de las razones de selección es igual a (0,093 x 0,09)/(0,08 x 0,08) = 1,3, que es mayor que 1.
3. Como el sesgo se define como OR^o – OR, esto significa que el sesgo se aleja del valor nulo (es decir, OR^o está más alejado del valor nulo que OR).

P8.8

1. α = 100/100 = 1,0; β = 50/50 = 1,0; γ = 7920/9900 = 0,8; y δ = 7960/9950 = 0,8.
2. Sí. La población de estudio es un subconjunto de la población de origen.
3. Sí, más o menos. La incidencia de la enfermedad en la población de origen general es 150/20.000=0,0075 y 150/16.030=0,0093 en la población de estudio general. Dentro de cada grupo de exposición, los riesgos son 100/10.000=0,01 y 50/110.000=0,005, respectivamente, en la población de origen; y 100/8.020=0,0124 y 50/8.010=0,0062, respectivamente, en la población de estudio.
4. No. La razón de productos cruzados de las razones de selección es igual a 1.
5. No. La razón de riesgos en la población de estudio es 1,998, que no es idéntica pero es básicamente igual a la razón de riesgos en la población de origen.

6. $\alpha = 100/100 = 1{,}0$; $\beta = 50/50 = 1{,}0$; $\gamma = 99/9.900 = 0{,}01$; y $\delta = 99{,}5/9950 = 0{,}01$.
7. No. La razón de productos cruzados de las razones de selección es igual a 1.
8. Sí. La razón de riesgos en la población de estudio es 1,503, que es menor que la razón de riesgos de 2,0 en la población de origen. La dirección del sesgo es hacia el valor nulo.
9. No. La incidencia de la enfermedad no es para nada poco frecuente en la población de estudio. En general, la incidencia de la enfermedad es 150/348,5 = 0,4304 en la población de estudio, y 100/199 = 0,5025 y 50/149,5 = 0,3344 dentro de los grupos de expuestos y no expuestos, respectivamente.
10. $\alpha = 40/40 = 1{,}0$; $\beta = 10/10 = 1{,}0$; $\gamma = 30/60 = 0{,}5$; y $\delta = 90/90 = 1{,}0$.
11. Sí. La razón de productos cruzados de las razones de selección es igual a (1 x 1)/(1 x 0,5) = 2, que claramente es distinto de 1. El valor incorrecto de OR = 12, que es mayor que el valor correcto de OR = 6, lo que indica que el sesgo se aleja del valor nulo.
12. Sí. La razón de riesgos en la población de estudio es 5,71, que es mayor que la razón de riesgos de 4,0 en la población de origen. El sesgo se aleja del valor nulo.
13. La enfermedad no es poco frecuente ni en la población de origen ni en la de estudio. Especialmente, no es poco frecuente para los individuos expuestos en la población de estudio.
14. No. El sesgo en el *odds ratio* ($OR^o = 12$ mientras que OR = 6) es mucho más grande que el sesgo en la razón de riesgos ($RR^o = 5{,}71$ mientras que RR = 4,0).

P8.9

1. $\alpha > \beta$
2. $\gamma > \delta$
3. $\alpha = \beta$
4. $\gamma = \delta$

P8.10

1. Verdadero. Los controles no expuestos en el estudio incluirán a las personas con compañeros sexuales con y sin ETS. Los no casos no expuestos en la población de origen no incluirán a compañeros sexuales sin ETS.
2. Falso. Los controles expuestos irán a clínicas para enfermedades de transmisión sexual, principalmente, por razones de prevención de la enfermedad, más que por anticoncepción. Es más probable que tengan compañeros sexuales con ETS que los controles no expuestos y, por lo tanto, es más probable que representen a la población de origen.
3. Sí. En la población de origen las probabilidades para los usuarios de preservativos representan la mitad de las probabilidades para los no usuarios, mientras que el *odds ratio* es uno en la población de estudio.
4. Sí. La razón (es decir, las probabilidades) de los no casos no expuestos sobre los expuestos en la población de estudio es de 6 a 1, mientras que las probabilidades de los no casos no expuestos sobre los expuestos en la población de origen es de 3 a 1.
5. $\alpha = 100/100 = 1$; $\beta = 600/600 = 1$; $\gamma = 200/200 = 1$; y $\delta = 1200/600 = 2$.
6. En este ejemplo, α (1) es igual a β (1), pero γ (1) es menor que δ (2). En consecuencia, la razón de productos cruzados de las razones de selección debería ser mayor que 1. Por lo tanto, hay sesgo de selección en el *odds ratio*.
7. El sesgo sería hacia el valor nulo, ya que el OR (incorrecto) en la población de estudio es igual a 1 mientras que el OR correcto es distinto de 1 (en este caso, menor) en la población de origen.
8. Falso. Tanto los controles no expuestos en el estudio como los no casos no expuestos se restringirán a quienes tengan compañeros sexuales con ETS.
9. Falso. Los controles expuestos y los no casos expuestos también se restringirán a quienes tengan compañeros sexuales con ETS. En consecuencia, los controles expuestos en el estudio reflejarán a la población de origen de no casos expuestos; todas esas parejas padecen ETS.
10. γ, la razón de selección para los controles expuestos, es igual a δ, la razón de selección de los controles no expuestos.
11. La razón de productos cruzados de las razones de selección debería ser igual a 1.

P8.11

1. La población de origen es la población de la comunidad y la población de estudio es la muestra de casos y controles esperada bajo las condiciones de selección descritas para el estudio.
2. α = 225/400 = 0,5625; β = 75/200 = 0,375; γ = 100/100.000 = 0,001; y δ = 200/200.000 = 0,001.
3. Sí, porque cada una se deriva como una razón en la cual el numerador es un subconjunto del denominador.
4. α = 0,5625 es mayor que β = 0,375, mientras γ = δ. En consecuencia, la razón de productos cruzados de las probabilidades de selección = (0,5625 x 0,001)/(0,375 x 0,001) = 1,5, que es diferente de 1.
5. El OR en la población de estudio es 6 mientras que el OR en la población de origen es 4. El *odds ratio* incorrecto (es decir, sesgado) de 6 es una sobrestimación del *odds ratio* correcto y, por lo tanto, se aleja del valor nulo.
6. Los pacientes que son tanto casos como fumadores están sobrerrepresentados en el estudio, en comparación con la población de la comunidad.
7. Sí. Los controles se dividen 50:50 entre los expuestos y los no expuestos, mientras que los no casos en la comunidad se dividen 1:2.
8. α = 225/400 = 0,5625; β = 75/200 = 0,375; γ = 150/100.000 = 0,0015; y δ = 150/200.000 = 0,00075.
9. α = 0,5625 es mayor que β = 0,375, mientras γ = 0,0015 es menor que δ = 0,00075. La razón de productos cruzados de las probabilidades de selección = (0,5625 x 0,00075)/(0,375 x 0,0015) = 0,75, que es diferente de 1.
10. El OR en la población de estudio es 3 mientras que el OR en la población de origen es 4. El *odds ratio* incorrecto (es decir, sesgado) de 3 es una subestimación del *odds ratio* correcto y, por lo tanto, su dirección es hacia el valor nulo.
11. Los pacientes de hospitales que son tanto casos como fumadores están más sobrerrepresentados que los pacientes de hospitales que son controles y fumadores, en comparación con la población de la comunidad.

P8.12

1. 25/50 = 0,5 para los expuestos y 40/80 = 0,5 para los no expuestos.
2. 142,5/150 = 0,95
3. 540/720 = 0,75
4. α = β = 0,5; γ = 0,95 y δ = 0,75.
5. (0,5 x 0,75) / (0,95 x 5) = 0,79
6. Sí, un leve sesgo. Por el *odds ratio*, ROR = 3 mientras que POR (= OR en la población de estudio) = 2,4. Por la razón de riesgos, RR = 2,5 mientras que PR (= efecto en la población de estudio) = 2,2. La dirección del sesgo es hacia el valor nulo.

P8.13

1. antes
2. falta de respuesta, población de estudio
3. uno
4. anterior, al
5. datos de prevalencia, datos de incidencia
6. supervivencia selectiva
7. sobrevivientes
8. sobreviven

P8.14

1. $\hat{A} = 40, \hat{B} = 10, \hat{C} = 60, \hat{D} = 90$
2. $OR_{ajustado} = 6$
3. $\hat{r}_D = 1, \ \hat{r}_{\overline{D}} = 0{,}5$
4. $OR_{ajustado} = \hat{OR} \square (\hat{r}_{\overline{D}} / \hat{r}_D) = 12 \square (0{,}5/1) = 6$
   $RR_{ajustada} = (40/100)/(10/100) = 4$
5. $\hat{r}_E = \hat{\alpha}/\hat{\gamma} = 2, \hat{r}_{\overline{E}} = \hat{\beta}/\hat{\delta} = 1$
6. $RR_{ajustada}$ = [40/(40 + [30 x 2])]/{10/(10 + [90 x 1])} = 4

P8.15

1. Deberá tomarse una decisión con respecto a cuál de los dos grupos de control es el más adecuado, por ejemplo, cuál de los grupos de control es el más representativo de la población de origen.

2. Cualquiera de las siguientes opciones es posible: 1) no hay sesgo de selección porque los dos efectos calculados son correctos; o 2) hay sesgo de selección porque los dos efectos calculados están sesgados.
3. Sí. Como la presencia de sesgo de selección depende de los parámetros de selección dentro del cuadro de 2 x 2, todavía puede haber sesgo de selección, incluso con excelentes tasas de respuesta y mínima pérdida durante el seguimiento.
4. No necesariamente. Al igual que en la pregunta anterior, la presencia de sesgo de selección depende de los parámetros de selección dentro del cuadro de 2 x 2, mientras que la información brindada sólo considera la pérdida de seguimiento en los bordes (es decir, el total de expuestos y de no expuestos) del cuadro de 2 x 2.
5. No. La magnitud del sesgo de selección no se puede determinar sin cálculos aproximados de los parámetros de selección o sus razones.
6. No. Un *odds ratio* corregido requeriría cálculos aproximados de los parámetros de selección o sus razones.
7. Debido a que alfa es menor que beta y gamma es igual a delta, el producto cruzado $(\alpha \times \delta) / (\beta \times \gamma)$ debe ser menor que 1. Por lo tanto, OR^o debe ser menor que OR, de forma tal que el sesgo sería hacia el valor nulo si el OR fuese mayor que 1 y se alejaría del valor nulo si el OR fuese menor que 1.
8. No. Sin saber si α sobre γ es igual, menor o mayor que β sobre δ, no es posible determinar si el producto cruzado $(\alpha \times \delta)/(\beta \times \gamma) = (\alpha / \gamma)/(\beta / \delta)$ es igual, menor o mayor que 1.
9. Sí. No hay sesgo de selección en el *odds ratio* porque el producto cruzado $(\alpha \times \delta) / (\beta \times \gamma)$ es igual a 1.
10. Sí. El sesgo debe alejarse del valor nulo porque el producto cruzado $(\alpha \times \delta)/(\beta \times \gamma)$ es mayor que 1, de forma que el OR^o sesgado es mayor que el OR correcto, que es mayor que 1.

P8.16

1. las frecuencias de celda ajustadas, parámetro de selección
2. *odds ratio*, cálculo del producto cruzado de los parámetros de selección
3. casos incidentes, casos prevalentes
4. de casos y controles anidados, basados en hospitales
5. una alta respuesta, pérdida de seguimiento
6. Verdadera
7. Verdadera. Como el análisis de situación "en el peor de los casos" puede demostrar que el peor grado de sesgo tendrá un impacto mínimo o irrelevante en las conclusiones, se podría descartar el sesgo de selección. Sin embargo, ya que el análisis de situación "en el peor de los casos" nos brinda los "peores resultados posibles", no podemos confirmar el sesgo de selección. No podemos estar seguros de que nuestros resultados sean tan extremos como los resultados "en el peor de los casos".
8. 32, 22, 110, 106
9. 3,9
10. 1,3
11. Sí
12. No. El análisis de situación "en el peor de los casos" brinda los "peores resultados posibles". Por lo tanto, no podemos estar seguros de que los resultados de pérdida de seguimiento que "realmente" ocurren sean tan extremos como "en el peor de los casos".

LECCIÓN 9

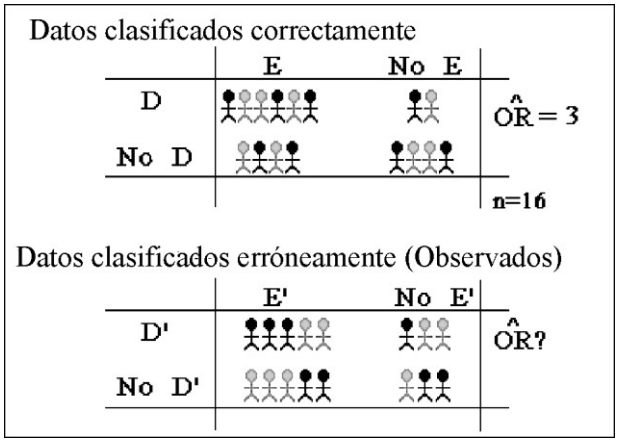
SESGO DE INFORMACIÓN

9-1 Sesgo de información

*El **sesgo de información** es un **error sistemático** que surge cuando se obtiene información incorrecta sobre una o más variables medidas inadecuadamente en un estudio. En este capítulo, nos enfocamos en las consecuencias de poseer información imprecisa sobre variables de exposición y enfermedad que son dicotómicas, es decir, cuando hay **clasificación errónea, o mala clasificación,** de la exposición y la enfermedad que produce un **sesgo** en la **medida de efecto** resultante. Las situaciones más generales, como, por ejemplo, varias categorías de exposición o enfermedad, variables continuas de exposición o enfermedad, ajuste por covariables, datos equiparados y abordajes de modelación matemática, están más allá del alcance de las actividades que se brindan en este capítulo y se discuten más adelante.*

¿Qué es el sesgo de clasificación errónea o mala clasificación?

El siguiente cuadro 2 x 2 muestra la clasificación correcta de 16 individuos según su verdadera condición de exposición y enfermedad. Veamos lo que le podría suceder a este cuadro y al *odds ratio* correspondiente si algunos de estos individuos se clasificaran erróneamente.



Suponga que 3 de los 6 casos expuestos (sombreados en gris) en realidad fueron diagnosticados erróneamente como no casos expuestos. Suponga además que uno de los dos casos no expuestos también fue clasificado erróneamente como un no caso no expuesto. Para completar el escenario hipotético de clasificación errónea, suponemos que dos de los cuatro no casos verdaderamente expuestos y dos de los no casos verdaderamente no expuestos fueron clasificados erróneamente como casos.

Los datos clasificados erróneamente son aquellos que realmente se analizarían, dado que son los que se observan en el estudio. Entonces, ¿cuál es el *odds ratio* para estos datos y cómo se compara con el *odds ratio* correcto? El *odds ratio* para los datos erróneamente clasificados es 1; el *odds ratio* correcto es 3. Claramente, hay

un sesgo debido a una clasificación errónea. Los datos clasificados erróneamente sugieren que no hay efecto de la exposición sobre la enfermedad, pero el verdadero efecto de la exposición es bastante fuerte.

Resumen

- Si los individuos se clasifican erróneamente según la condición de exposición o enfermedad, la medida de efecto, por ejemplo, el OR, puede volverse sesgada.
- El sesgo de clasificación errónea puede ocurrir si la medida de efecto para los datos clasificados correctamente es relevantemente diferente del efecto calculado que realmente se observa en los datos clasificados erróneamente.
- Los individuos se clasifican erróneamente si su ubicación en una de las cuatro celdas de los datos clasificados correctamente cambia a una ubicación de celda diferente en los datos (erróneamente clasificados) que realmente se observan.

Formulación general del sesgo de clasificación errónea

El marco general utilizado en la Lección 7 (Validez) para describir lo que se entiende por sesgo, incorpora el sesgo de clasificación errónea, además del sesgo de selección. Para el sesgo de clasificación errónea, la población objetivo está representada por un cuadro de 2 x 2 (el cuadro 1 a continuación) que supone que no hay clasificación errónea de ningún tipo; mientras que la población de estudio está representada por un cuadro de 2 x 2 (el cuadro 2 a continuación) que presenta una reorganización (en lugar de un subconjunto, como en el sesgo de selección) de la población objetivo que surge de una clasificación errónea.

Cuadro 1. Población objetivo (sin clasificación errónea)

	E	no E
D	**A** $(=A_{11}^o + B_{11}^o + C_{11}^o + D_{11}^o)$	**B** $(=A_{12}^o + B_{12}^o + C_{12}^o + D_{12}^o)$
no D	**C** $(=A_{21}^o + B_{21}^o + C_{21}^o + D_{21}^o)$	**D** $(=A_{22}^o + B_{22}^o + C_{22}^o + D_{22}^o)$

Cuadro 2. Población de estudio (clasificada erróneamente)

	E´	no E´
D´	$\mathbf{A^o}$ $(=A_{11}^o + A_{12}^o + A_{21}^o + A_{22}^o)$	$\mathbf{B^o}$ $(=B_{11}^o + B_{12}^o + B_{21}^o + B_{22}^o)$
no D´	$\mathbf{C^o}$ $(=C_{11}^o + C_{12}^o + C_{21}^o + C_{22}^o)$	$\mathbf{D^o}$ $(=D_{11}^o + D_{12}^o + D_{21}^o + D_{22}^o)$

En el cuadro 1, las personas **A** que verdaderamente padecen la enfermedad y están expuestas (**DE**) se clasifican en cada una de las cuatro celdas del cuadro 2 de la siguiente manera:

A_{11}^o se clasifican como enfermos y expuestos (**D'E'**);

B_{11}^{o} se clasifican como enfermos y no expuestos (**D', no E'**);

C_{11}^{o} se clasifican como no enfermos y expuestos (**no D', E'**); y

D_{11}^{o} se clasifican como no enfermos y no expuestos (**no D', no E'**).

En otras palabras, $A = A_{11}^{o} + B_{11}^{o} + C_{11}^{o} + D_{11}^{o}$

Se muestra una reorganización similar para las personas **B** que verdaderamente padecen la enfermedad y no están expuestas (**D, no E**), las personas **C** que verdaderamente no padecen la enfermedad y están expuestas (**no D, E**), y las personas **D** que verdaderamente no padecen la enfermedad y no están expuestas (**no D, no E**).
En consecuencia, las personas **A^{o}** en el cuadro 2 que están clasificadas como enfermas y expuestas (D'E') se derivan de las cuatro celdas de la población objetivo (cuadro 1), tal como se expresa en:

$$A^{o} = A_{11}^{o} + A_{12}^{o} + A_{21}^{o} + A_{22}^{o}$$

Declaraciones similares se aplican a **B^{o}**, **C^{o}** y **D^{o}**, tal como se indica en el cuadro 2 anterior.

Clasificación errónea de la condición de enfermedad

¿Por qué razones un individuo, como un caso expuesto, podría ser clasificado erróneamente según la condición de exposición o enfermedad? En especial, ¿por qué se podría clasificar erróneamente a los individuos como enfermos en lugar de no enfermos o viceversa? En primer lugar, un individuo puede estar diagnosticado incorrectamente. Esto puede ocurrir a raíz de conocimientos limitados de la enfermedad, porque el proceso de diagnóstico es complejo, por acceso inadecuado a tecnología de diagnóstico de vanguardia, o debido a un error de laboratorio en la medición de los marcadores biológicos de la enfermedad. Además, puede no detectarse la presencia de la enfermedad si ésta se encuentra en estatus subclínico al momento del examen físico. Un error de diagnóstico se puede presentar debido a un sesgo de detección si el médico realiza un examen más completo a los pacientes que están expuestos o que poseen síntomas relacionados con la exposición.

¿Por qué la clasificación errónea de la enfermedad?

* Diagnóstico Incorrecto
  Conocimientos limitados
  El proceso de diagnóstico es complejo
  Acceso inadecuado a tecnología
  Error de laboratorio
  Enfermedad en estado subclínico
  Sesgo de detección
* Información proporcionada por los individuos
  Recordar incorrectamente una condición de enfermedad
  Reticente a decir la verdad
* Los registros pueden ser imprecisos o estar codificados en forma incorrecta en la base de datos

Otra fuente de error surge cuando la condición de enfermedad se obtiene sólo por la información proporcionada por los individuos, en lugar del examen médico. En especial, un individuo puede recordar incorrectamente una condición de enfermedad como, por ejemplo, una afección respiratoria u otra infecciosa, que puede haber ocurrido en un período de tiempo anterior. Un individuo puede ser reticente a decir la verdad sobre una enfermedad que considera social o personalmente inaceptable. Finalmente, los registros de los pacientes pueden ser imprecisos o estar codificados en forma incorrecta en la base de datos.

El siguiente cuadro resume la clasificación errónea de la condición de enfermedad. Las columnas del cuadro muestran la verdadera condición de enfermedad. Las filas del cuadro muestran la condición de enfermedad clasificada. A este cuadro lo llamamos **cuadro de clasificación errónea**.

Cuadro de clasificación errónea

		Verdadera	
		D	No D
Clasificada	D'		
	No D'		

Suponga que los siguientes números aparecen en el cuadro:

Cuadro de clasificación errónea

		Verdadera		
		D	No D	
Clasificada	D'	6	3	9
	No D'	2	7	9
		8	10	

Preguntas de estudio (P9.1)

Las siguientes preguntas se refieren al cuadro anterior.

1. ¿Cuántas personas verdaderamente enfermas están clasificadas erróneamente?
2. ¿Cuántas personas verdaderamente no enfermas están clasificadas correctamente?
3. ¿Cuál es el porcentaje de personas verdaderamente enfermas que están clasificadas correctamente?
4. ¿Cuál es el porcentaje de personas verdaderamente no enfermas que están clasificadas correctamente?

Resumen

- La clasificación errónea de la condición de enfermedad puede producirse por alguna de las siguientes fuentes:
  - diagnóstico incorrecto
  - informes personales de los individuos
  - errores de codificación
- Un cuadro de clasificación errónea brinda un resumen práctico de cómo se puede clasificar erróneamente la condición de enfermedad como verdadera en lugar de observada.

Clasificación errónea de la condición de exposición

¿Cómo es posible clasificar erróneamente a los individuos expuestos como no expuestos o a los no expuestos como expuestos? La clasificación errónea de la condición de exposición puede producirse debido a una medición imprecisa de la exposición. Esto puede ser el resultado de un cuestionario mal formulado o de un proceso de relevamiento que no formula las preguntas correctas, o de un dispositivo de medición o una técnica de observación que presenta fallas.

¿Por qué la clasificación errónea de la condición de exposición?

* Medición imprecisa
* Informes personales de los individuos
* Sesgo del entrevistador
* Codificación incorrecta de los datos de exposición

Preguntas de estudio (P9.2)

La calidad de medición de exposición individual a campos electromagnéticos (CEM) es una crítica fundamental de los estudios que evalúan si vivir cerca de líneas o cableado de alta tensión aumenta el riesgo de padecer cáncer. ¿Cuáles de las siguientes características de la medición imprecisa de la exposición individual a los CEM cree que es **verdadera** o **falsa**?

1. Normalmente, las mediciones se realizan sólo en un punto del tiempo y/o en un lugar de residencia.
2. Los instrumentos para medir la exposición a los CEM no están disponibles.
3. Los métodos para medir las distancias y la configuración del cableado de transmisión cerca de las residencias están mal desarrollados.
4. Se necesitan mejores métodos para monitorear las mediciones através de un espacio de tiempo.
5. La cuantificación de la exposición a los CEM basándose en el uso intermitente de aparatos o herramientas es difícil de medir.
6. Los patrones de movilidad de los individuos con relación a la exposición a los CEM son difíciles de medir.

Puede producirse un error de exposición cuando se determina la exposición sólo por la información brindada por los individuos, especialmente cuando recuerdan la condición de exposición previa. Esto es un problema típico en los estudios de casos y controles, ya que los casos pueden estar más motivados que los controles para recordar las exposiciones pasadas. El error de recuerdo también puede presentarse en los estudios de cohorte. Por ejemplo, en el Estudio de usuarios de las playas de Sydney descrito en la Lección 2, se les pidió a los individuos que informaran su exposición o condición de bañistas siete días después de bañarse en el mar.

Los informes personales de los individuos también pueden ser incorrectos debido a su reticencia a ser sinceros cuando informan exposiciones relacionadas con comportamientos considerados socialmente inaceptables. Este problema sobreviene a menudo en estudios que miden el consumo de alimentos, el comportamiento sexual y el uso ilegal de drogas.

Una tercera fuente de error en la clasificación de la exposición es el sesgo del entrevistador. En especial, un entrevistador puede investigar en forma más exhaustiva la exposición para los casos que para los controles.

Preguntas de estudio (P9.2) (continuación)

Considere un estudio de casos y controles que investiga el efecto del uso de anticonceptivos orales en el desarrollo de trombosis (es decir, coagulación) venosa (es decir, en la vena). (Nota: no hay preguntas numeradas del 7 al 9).

10. ¿Por qué podría haber clasificación errónea de la exposición, es decir, uso de anticonceptivos orales, en dicho estudio?
11. ¿Cuál espera que sea la dirección de dicho sesgo de clasificación errónea?
12. ¿Cómo podría evitar tal sesgo?

Finalmente, los datos de exposición pueden estar codificados en forma incorrecta sobre la base de los datos. El siguiente cuadro resume las clasificaciones erróneas de la condición de exposición. Las columnas del cuadro muestran la verdadera condición de exposición. Las filas del cuadro muestran la condición de exposición clasificada.

Cuadro de clasificación errónea

	Verdadera E	Verdadera No E
Clasificada E'		
Clasificada No E'		

Suponga que los siguientes números aparecen en este cuadro:

Cuadro de clasificación errónea

	Verdadera E	Verdadera No E	
Clasificada E'	95	10	105
Clasificada No E'	5	70	75
	100	80	

Preguntas de estudio (P9.2) (continuación)

Las preguntas están basadas en el cuadro anterior.

13. ¿Cuántas personas verdaderamente expuestas están clasificadas erróneamente?
14. ¿Cuántas personas verdaderamente no expuestas están clasificadas correctamente?
15. ¿Cuál es el porcentaje de personas verdaderamente expuestas que están clasificadas correctamente?
16. ¿Cuál es el porcentaje de personas verdaderamente no expuestas que están clasificadas correctamente?

Resumen

- La clasificación errónea de la condición de exposición puede producirse por alguna de las siguientes fuentes:
  - medición imprecisa
  - informes personales de los individuos
  - sesgo del entrevistador
  - codificación incorrecta de los datos de exposición
- Un cuadro de clasificación errónea brinda un resumen práctico de cómo se puede clasificar erróneamente la exposición como verdadera en lugar de observada.

Clasificación errónea de la condición de exposición y de la condición de enfermedad: ejemplo

A veces, la clasificación errónea puede producirse tanto para la exposición como para la enfermedad en el mismo estudio. Por ejemplo, el siguiente cuadro considera datos de cohorte hipotéticos de individuos encuestados en las playas de Sydney, Australia, durante los meses de verano de un año reciente. El objetivo del estudio era determinar si los que se bañaban en la playa tenían más probabilidad de enfermarse que los que no lo hacían.

Datos clasificados correctamente	Se bañó (E)	No se bañó (No E)	
enfermo (D)	367	233	$\hat{RR} = 3.14$
no enfermo (No D)	300	1100	
Total	667	1333	

Se les pidió a los individuos que recordaran una semana después si se habían bañado durante, al menos, media hora el día que fueron entrevistados en la playa y si habían desarrollado resfrío, tos o gripe durante la siguiente semana. Ya que tanto la información sobre la exposición como sobre la enfermedad se obtuvo por el recuerdo de los individuos, es razonable esperar que algunos individuos puedan informar incorrectamente la condición de bañista o de enfermedad, o ambas.

Suponga que de los 367 individuos que se enfermaron y se bañaron en el mar, sólo 264 informaron que se enfermaron y se bañaron en el mar, y 30 indicaron que se enfermaron pero que no se bañaron en el mar. También, 66 expresaron que no se enfermaron pero que se bañaron en el mar y 7 manifestaron que ni se enfermaron ni se bañaron en el mar.

Suponga, además, que de los 233 individuos que verdaderamente se enfermaron pero no se bañaron en el mar, sólo 130 informaron ésto correctamente. También, 56 indicaron que se enfermaron y que se bañaron en el mar, 14 expresaron que no se enfermaron pero que se bañaron en el mar y 33 manifestaron que ni se enfermaron ni se bañaron en el mar.

De esta forma, el cuadro puede corregirse aún más para describir cómo se clasificó erróneamente a los 300 individuos que verdaderamente no se enfermaron y se bañaron en el mar. El cuadro también puede corregirse para describir la clasificación errónea de los 1100 individuos que verdaderamente ni se enfermaron ni se bañaron en el mar.

Datos observados

	Se bañó'		No se bañó'	
enfermo'	264	56	30	130
	27	33	3	77
no enfermo'	66	14	7	33
	243	297	27	693

Ahora, en forma separada, podemos sumar las 4 frecuencias dentro de cada una de las 4 celdas del cuadro de datos observados para obtener un cuadro resumido de los datos observados, como se muestra aquí:

Datos observados

	Se bañó'	No se bañó'
enfermo'	380	240
no enfermo'	620	760
Total	1000	1000

Preguntas de estudio (P9.3)

1. ¿Cuál es el valor de la razón de riesgos para los datos observados?
2. ¿Por qué hay sesgo de clasificación errónea? (Sugerencia: RR = 3,14 para los datos verdaderos).
3. ¿Cuál es la dirección del sesgo?
4. Si hay clasificación errónea tanto de la exposición como de la enfermedad, ¿el sesgo siempre será hacia el valor nulo?

Resumen

- A veces, la clasificación errónea puede producirse tanto para la exposición como para la enfermedad en el mismo estudio.
- Es probable que surja un ejemplo de dicha clasificación errónea si tanto la variable de exposición como la variable de enfermedad se determinan según lo que los individuos recuerdan.
- Cuando hay clasificación errónea tanto de la exposición como de la enfermedad, los datos observados surgen de la forma en que las frecuencias de celda en cada una de las cuatro celdas del cuadro de 2 x 2 para los datos verdaderos se dividan en las cuatro celdas del cuadro de 2 x 2 para los datos observados.

9-2 Sesgo de información (continuación)

Probabilidades de clasificación errónea: sensibilidad y especificidad

El siguiente cuadro de clasificación errónea describe cómo puede estar mal diagnosticada una enfermedad **D**. Doce individuos que verdaderamente estaban enfermos fueron clasificados erróneamente como no enfermos y 14 individuos que, en realidad, no estaban enfermos fueron clasificados erróneamente como enfermos. 48 individuos que verdaderamente estaban enfermos y 126 individuos que, en realidad, no lo estaban fueron clasificados correctamente.

Clasificación errónea de la enfermedad

Clasificada	Verdadera: D	Verdadera: No D	
D'	48	14	62
No D'	12	126	138
	60	140	

En un mundo perfecto, nadie sería clasificado erróneamente, es decir, nuestro cuadro tuviera información similar a la que se consigna a continuación. Entonces, la proporción clasificada correctamente como enfermos sería 1 y la proporción clasificada correctamente como no enfermos también sería 1.

Mundo Perfecto

Clasificada	Verdadera: D	Verdadera: No D		Proporción
D'	60	0	60	$\frac{60}{60} = 1$
No D'	0	140	140	$\frac{140}{140} = 1$
	60	140		

En el mundo real, sin embargo, estas proporciones de clasificación correcta no son iguales a 1. En nuestro ejemplo, la proporción de quienes en verdad estaban enfermos, y que se clasificaron correctamente como tales, es 0,80 (48/60). La proporción de los que en verdad no estaban enfermos, y correctamente clasificados como tales, es 0,90 (126/140).

$\frac{48}{60} = .80 \rightarrow$ Sensibilidad = $\mathbf{Prob(\ D' |\ D\ )}$ (dado)

$\frac{126}{140} = .90 \rightarrow$ Especificidad = $\mathbf{Prob(no\ \ D' |no\ \ D\ )}$

La primera de estas proporciones de clasificación se llama **sensibilidad**. Generalmente, la sensibilidad de la clasificación errónea de la condición de enfermedad es la probabilidad de que un individuo sea clasificado como enfermo, dado que en realidad está enfermo.

La segunda de estas proporciones de clasificación se llama **especificidad**. En general, la especificidad de la clasificación errónea de la enfermedad es la probabilidad de que un individuo sea clasificado como no enfermo, dado que verdaderamente no está enfermo.

El valor ideal tanto de la sensibilidad como de la especificidad es 1,0 ó 100%. También podemos utilizar el cuadro de clasificación errónea para la condición de **exposición**, a fin de definir los parámetros de sensibilidad y especificidad.

Clasificación errónea de la condición de exposición

Clasificada	Verdadera: E	Verdadera: No E	
E'			
No E'			
			(dado)

Sensibilidad = $\mathbf{Prob(\ E' |\ E\ )}$

Especificidad = $\mathbf{Prob(no\ \ E' |no\ \ E\ )}$

Preguntas de estudio (P9.4)

Considere los números en el siguiente cuadro de clasificación errónea de la exposición.

Clasificación errónea de la condición de exposición

	Verdadera: E	Verdadera: No E	
Clasificada: E'	720	190	910
Clasificada: No E'	180	910	1090
	900	1100	

Sensibilidad = **Prob(E'| E)**

Especificidad = **Prob(no E'|no E)**

1. ¿Cuál es la sensibilidad de la clasificación errónea de la exposición?
2. ¿Cuál es la especificidad de la clasificación errónea de la exposición?
3. ¿Sus respuestas a las preguntas anteriores sugieren que debería haber alguna preocupación sobre el sesgo de clasificación errónea?

Resumen

- Los parámetros subyacentes que deben considerarse cuando se evalúa el sesgo de información se llaman **sensibilidad** y **especificidad**.
- La **sensibilidad** da la probabilidad de que un individuo que está verdaderamente enfermo (o expuesto) sea clasificado como enfermo (o expuesto) en un estudio.
- La **especificidad** da la probabilidad de que un individuo que está verdaderamente no enfermo (o no expuesto) sea clasificado como no enfermo (o no expuesto) en un estudio.
- El valor ideal tanto de la sensibilidad como de la especificidad es 1 ó 100%, lo que significa que no hay clasificación errónea.

Los parámetros de sensibilidad y especificidad son probabilidades condicionales

Si permitimos una posible clasificación errónea de la enfermedad y la exposición en un cuadro de 2 x 2, entonces hay varios parámetros de sensibilidad y especificidad que pueden definirse en términos de **probabilidades condicionales**. Éstos incluyen:

Sensibilidad (Se)

Clasificación errónea de la enfermedad: $\mathbf{Se_D = Pr(D'|D)}$
Clasificación errónea de la exposición: $\mathbf{Se_E = Pr(E'|E)}$

Especificidad (Sp)

Clasificación errónea de la enfermedad: $\mathbf{Sp_D = Pr(no\ D'|no\ D)}$
Clasificación errónea de la exposición: $\mathbf{Sp_E = Pr(no\ E'|no\ E)}$

donde **D** y **E** denotan a los verdaderamente enfermos y expuestos, respectivamente, y **D'** y **E'** denotan a los clasificados (posiblemente en forma errónea) como enfermos y expuestos, respectivamente.

Además, si permitimos la posibilidad de que la clasificación errónea de la enfermedad pueda diferir según la condición de exposición, y que la clasificación errónea de la condición de exposición pueda diferir según la condición de enfermedad, podemos categorizar los parámetros de sensibilidad y especificidad aún más de la siguiente manera:

Sensibilidad

1. Clasificación errónea de la enfermedad dado expuestos: $\mathbf{Se_{D|E} = Pr(D'|DE)}$
2. Clasificación errónea de la enfermedad dado no expuestos: $\mathbf{Se_{D|no\ E} = Pr(D'|D,\ no\ E)}$
3. Clasificación errónea de la exposición dado enfermos: $\mathbf{Se_{E|D} = Pr(E'|DE)}$
4. Clasificación errónea de la exposición dado no enfermos: $\mathbf{Se_{E|no\ D} = Pr(E'no\ D,\ E)}$

Especificidad

5. Clasificación errónea de la enfermedad dado expuestos: $\mathbf{Sp_{D|E} = Pr(no\ D'|no\ D,\ E)}$
6. Clasificación errónea de la enfermedad dado no expuestos: $\mathbf{Sp_{D|no\ E} = Pr(no\ D'|no\ D,\ no\ E)}$
7. Clasificación errónea de la exposición dado enfermos: $\mathbf{Sp_{E|D} = Pr(no\ D'|D,\ no\ E)}$
8. Clasificación errónea de la exposición dado no enfermos: $\mathbf{Sp_{E|no\ D} = Pr(no\ E'|no\ D,\ no\ E)}$

Los ocho parámetros anteriores son **probabilidades condicionales** y todos son potencialmente diferentes si hay clasificaciones erróneas de la exposición y de la enfermedad. Los cuadros que se presentan a continuación ilustran adónde pertenecen las probabilidades de sensibilidad y especificidad dentro de los cuadros de clasificación errónea que permiten la clasificación errónea de la condición de enfermedad (dada la condición de exposición) y de la condición de exposición (dada la condición de enfermedad):

Probabilidades de la clasificación errónea de la enfermedad (dada la condición de exposición)

Clasificados	Verdaderamente E: D	Verdaderamente E: no D
D´	$Se_{D\|E}$	$1\text{-}Sp_{D\|E}$
no D´	$1\text{-}Se_{D\|E}$	$Sp_{D\|E}$

Clasificados	Verdaderamente no E: D	Verdaderamente no E: no D
D´	$Se_{D\|no\ E}$	$1\text{-}Sp_{D\|no\ E}$
no D´	$1\text{-}Se_{D\|no\ E}$	$Sp_{D\|no\ E}$

Probabilidades de la clasificación errónea de la exposición (dada la condición de enfermedad)

Clasificados	Verdaderamente D: E	Verdaderamente D: no E
E´	$Se_{E\|D}$	$1\text{-}Sp_{E\|D}$
no E´	$1\text{-}Se_{E\|D}$	$Sp_{E\|D}$

Clasificados	Verdaderamente no D: E	Verdaderamente no D: no E
E´	$Se_{E\|no\ D}$	$1\text{-}Sp_{E\|no\ D}$
no E´	$1\text{-}Se_{E\|no\ D}$	$Sp_{D\|no\ D}$

Clasificación errónea no diferencial

El siguiente cuadro describe la verdadera condición de exposición y enfermedad para 2000 individuos en un estudio hipotético de casos y controles sobre la relación entre la dieta y la enfermedad coronaria (EC).

Estado verdadero de exposición y enfermedad
Consumo de frutas/ vegetales

	Bajo	Alto	Total
EC	600	400	1000
No EC	300	700	1000

$$\hat{OR} = \frac{600x700}{400x300} = 3.5$$

La variable de exposición es la cantidad de frutas y verduras ingeridas en una semana promedio, categorizando tal consumo como bajo y alto, según lo que recordaron los individuos del estudio. La variable de enfermedad es la presencia o ausencia de EC. Suponga que no hay clasificación errónea de la condición de enfermedad, pero que la mayoría de los individuos informan un consumo mayor de frutas y verduras que el real porque creen que las dietas con grandes cantidades de tales alimentos son más aceptables para el investigador. En otras palabras, hay una clasificación errónea de la condición de exposición.

Los dos cuadros siguientes describen separadamente cómo se clasifica erróneamente la exposición para los casos de EC y los no casos.

Clasificando erróneamente la exposición

Casos EC	Verdadero Bajo	Verdadero Alto	Total
Bajo' Clasificado	480	20	500
Alto'	120	380	500
Total	600	400	1000

No casos EC	Verdadero Bajo	Verdadero Alto	Total
Bajo' Clasificado	240	35	275
Alto'	60	665	725
Total	300	700	1000

Preguntas de estudio (P9.5)

1. ¿Cuál es la sensibilidad y especificidad para los casos de EC?
2. ¿Cuál es la sensibilidad y especificidad para los no casos?
3. ¿Qué tienen en común estos dos cuadros de clasificación errónea?

Este ejemplo ilustra **la clasificación errónea no diferencial de la exposición**. Esto ocurre cada vez que las sensibilidades y especificidades **no** varían con la condición de enfermedad. Hemos supuesto que la condición de EC no está clasificada erróneamente en este ejemplo. Las sensibilidades y especificidades de la clasificación errónea de la enfermedad son todas iguales a 1, independientemente del grupo de exposición. Por lo tanto, en este ejemplo, no hay clasificación errónea de la enfermedad.

Preguntas de estudio (P9.5) (continuación)

4. Utilice la columna de totales en ambos cuadros de clasificación errónea (es decir, 600, 400, 300 y 700) para determinar el *odds ratio* correspondiente a los datos clasificados correctamente (es decir, verdaderos).

La fila de totales de cada cuadro de clasificación errónea de la exposición nos permite determinar los datos del estudio que realmente se observarían como consecuencia de la clasificación errónea.

Datos observados (clasificados)

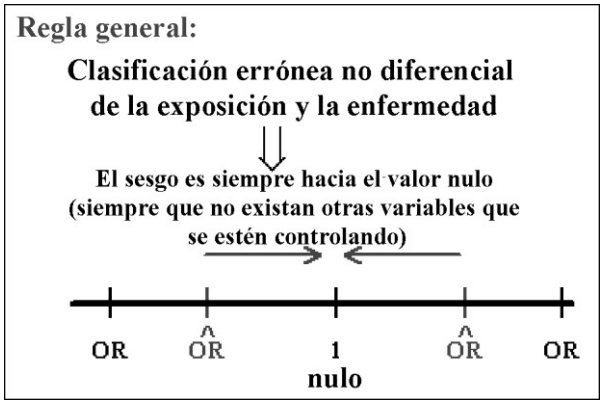
	Bajo	Alto	Total
EC	500	500	1000
No EC	275	725	1000

$$\hat{OR} = \frac{500x725}{500x275} = 2.6$$

Preguntas de estudio (P9.5) (continuación)

5. ¿Por qué hay sesgo debido a la clasificación errónea de la exposición? (Nota: el *odds ratio* correcto es 3,5).
6. ¿Cuál es la dirección del sesgo?

Este ejemplo ilustra una regla general sobre la clasificación errónea no diferencial. Cuando hay clasificación errónea no diferencial de la exposición y la enfermedad, el sesgo es siempre hacia el valor nulo, siempre que no existan otras variables que se estén controlando que también puedan estar clasificadas erróneamente.



Resumen

- **Clasificación errónea no diferencial de la enfermedad:** las sensibilidades y especificidades de la clasificación errónea de la enfermedad no difieren por niveles de exposición.
- **Clasificación errónea no diferencial de la exposición:** las sensibilidades y especificidades de la clasificación errónea de la exposición no difieren por estatus de enfermedad.
- La **clasificación errónea no diferencial tanto** de la enfermedad **como** de la exposición provoca un **sesgo hacia el valor nulo**.

Definición matemática de la clasificación errónea no diferencial

Para dar una definición general de la clasificación errónea no diferencial, debemos considerar los parámetros de sensibilidad y especificidad en los siguientes cuadros de clasificación errónea:

Probabilidades de la clasificación errónea de la enfermedad (dada la condición de exposición)

	Verdaderamente E	
Clasificados	**D**	**no D**
D´	$Se_{D\|E}$	$1\text{-}Sp_{D\|E}$
no D´	$1\text{-}Se_{D\|E}$	$Sp_{D\|E}$

	Verdaderamente no E	
Clasificados	**D**	**no D**
D´	$Se_{D\|no\ E}$	$1\text{-}Sp_{D\|no\ E}$
no D´	$1\text{-}Se_{D\|no\ E}$	$Sp_{D\|no\ E}$

Probabilidades de la clasificación errónea de la exposición (dada la condición de enfermedad)

	Verdaderamente D	
Clasificados	**E**	**no E**
E´	$Se_{E\|D}$	$1\text{-}Sp_{E\|D}$
no E´	$1\text{-}Se_{E\|D}$	$Sp_{E\|D}$

	Verdaderamente no D	
Clasificados	**E**	**no E**
E´	$Se_{E\|no\ D}$	$1\text{-}Sp_{E\|no\ D}$
no E´	$1\text{-}Se_{E\|no\ D}$	$Sp_{D\|no\ D}$

En general, hay **clasificación errónea no diferencial tanto** de la exposición **como** de la enfermedad si se satisfacen las siguientes cuatro ecuaciones:

Clasificación errónea de la enfermedad:

$\mathbf{Se_{D|E} = Se_{D|no\ E}\ (=Se_D)}$ y $\mathbf{Sp_{D|E} = Sp_{D|no\ E}\ (=Sp_D)}$

Clasificación errónea de la exposición:

$\mathbf{Se_{E|D} = Se_{E|no\ D}\ (=Se_E)}$ y $\mathbf{Sp_{E|D} = Sp_{E|no\ D}\ (=Sp_E)}$

Esta definición significa que cuando se clasifica la condición de enfermedad, las sensibilidades son iguales tanto entre el grupo expuesto (**E**) como en el no expuesto (**no E**), y lo mismo sucede con las especificidades. De forma similar, cuando se clasifica la condición de exposición, las sensibilidades son iguales tanto entre el grupo enfermo (**D**) como en el no enfermo (**no D**), y lo mismo sucede con las especificidades.

Por lo tanto, cuando hay clasificación errónea no diferencial tanto de la exposición como de la enfermedad, los cuadros de clasificación errónea presentados anteriormente se simplificarán de la siguiente manera:

Probabilidades de clasificación errónea no diferencial de la enfermedad

	Verdaderamente E	
Clasificados	**D**	**no D**
D´	Se_D	$1\text{-}Sp_D$
no D´	$1\text{-}Se_D$	Sp_D

	Verdaderamente no E	
Clasificados	**D**	**no D**
D´	Se_D	$1\text{-}Sp_D$
no D´	$1\text{-}Se_D$	Sp_D

Continúa en la página siguiente.

Definición matemática de la clasificación errónea no diferencial (continuación)

Probabilidades de la clasificación errónea no diferencial de la exposición

	Verdaderamente D	
Clasificados	**E**	**no E**
E´	Se_E	$1-Sp_E$
no E´	$1-Se_E$	Sp_E

	Verdaderamente no D	
Clasificados	**E**	**no E**
E´	Se_E	$1-Sp_E$
no E´	$1-Se_E$	Sp_D

Observe que si no hay clasificación errónea de la condición de enfermedad, entonces:

$\mathbf{Se_{D|E} = Se_{D|no\ E}\ (=Se_D) = 1}$ y $\mathbf{Sp_{D|E} = Sp_{D|no\ E}\ (=Sp_D) = 1}$

Esto significa que hay clasificación errónea no diferencial de la enfermedad, aunque no necesariamente clasificación errónea no diferencial de la exposición.

Asimismo, si no hay clasificación errónea de la condición de exposición, entonces:

$\mathbf{Se_{E|D} = Se_{E|no\ D}\ (=Se_E) = 1}$ y $\mathbf{Sp_{E|D} = Sp_{E|no\ D}\ (=Sp_E) = 1}$

Esto implica que hay clasificación errónea no diferencial de la exposición, aunque no necesariamente clasificación errónea no diferencial de la enfermedad.

¿Qué sucede si se clasifica erróneamente una variable <u>distinta</u> de la exposición o la enfermedad?

Greenland (1980) demostró que si hay clasificación errónea no diferencial de la exposición y la enfermedad, y que también existe clasificación errónea de una covariable, entonces la dirección del sesgo hacia el valor nulo es incierta. Sin embargo, si la clasificación errónea de la exposición y la enfermedad es no diferencial, y en el análisis se controla (por ejemplo, por estratificación) una covariable que no está clasificada erróneamente, entonces tanto las medidas de efecto específicas para cada estrato como las medidas de efecto de resumen ajustando por la covariable, tendrán un sesgo con dirección hacia el valor nulo.

También se abordaron otros problemas sobre la clasificación errónea de las covariables, como se expresa a continuación:

- La clasificación errónea de la exposición puede introducir falsamente una modificación de efecto (descrito en la Lección 10) por medio de una covariable.
- La clasificación errónea de un factor de confusión (también descrito en la Lección 10) puede reintroducir confusión en un cálculo de resumen que controla la confusión mediante datos clasificados erróneamente.

Clasificación errónea diferencial

El siguiente cuadro describe la verdadera condición de exposición y de enfermedad para los mismos 2000 individuos descritos en la actividad anterior en un estudio hipotético de casos y controles sobre la relación entre la dieta y la enfermedad coronaria:

Estado verdadero de exposición y enfermedad
Consumo de frutas/ vegetales

	Bajo	Alto	Total
EC	600	400	1000
No EC	300	700	1000

$$\hat{OR} = \frac{600x700}{400x300} = 3.50$$

Suponga, como antes, que no hay clasificación errónea de la condición de enfermedad, sino que los individuos informan un consumo de frutas y verduras mayor que el real porque creen que las dietas con grandes cantidades de tales alimentos son más aceptables para el investigador. Suponga también que un caso de EC, al que le preocupan las razones de su enfermedad, no tiene tantas probabilidades de sobrestimar su consumo de frutas y verduras como un control. A continuación se presentan los dos cuadros que describen cómo se clasifica erróneamente la exposición tanto para los casos de EC como para los controles.

Condición de enfermedad: No hay clasificación errónea
Condición de la exposición: Clasificación errónea

Casos EC	Verdadero Bajo	Verdadero Alto	Total
Bajo' Clasificado	580	20	600
Alto'	20	380	400
Total	600	400	1000

No casos EC	Verdadero Bajo	Verdadero Alto	Total
Bajo' Clasificado	240	35	275
Alto'	60	665	725
Total	300	700	1000

Preguntas de estudio (P9.6)

Las siguientes preguntas están basadas en los dos cuadros anteriores.

1. ¿Cuál es la sensibilidad y especificidad para los casos de EC?
2. ¿Cuál es la sensibilidad y especificidad para los no casos?
3. ¿Hay clasificación errónea no diferencial de la exposición?

Este ejemplo ilustra lo que se denomina **clasificación errónea diferencial de la exposición**. Esto ocurre porque las sensibilidades y especificidades de la clasificación errónea de la exposición varían con la condición de enfermedad.

Clasificación errónea diferencial de la exposición

580 / 600 = .97	diferente	240 / 630 = .80
Sensibilidad =97%	⟷	**Sensibilidad =80%**
380 / 400 = .95	igual	665 / 700 = .95
Especificidad =95%	⟷	**Especificidad =95%**

Los totales de las filas de cada cuadro de clasificación errónea de la exposición nos permiten determinar los datos del estudio que realmente se observarían como resultado de la clasificación errónea.

Datos observados (clasificados)

	Bajo	Alto	Total
EC	600	400	1000
No EC	275	725	1000

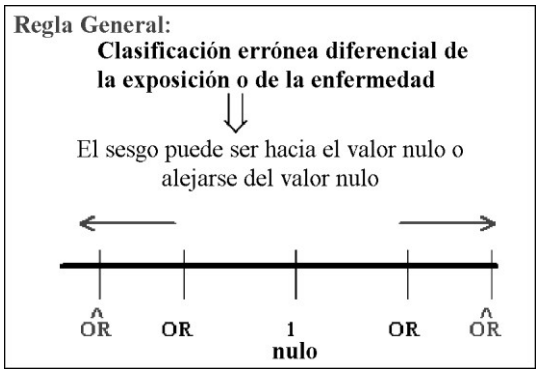
$\hat{OR} = \frac{600x725}{400x275} = 3.95$

Preguntas de estudio (P9.6) (continuación)

Las siguientes preguntas se refieren al cuadro anterior y al cuadro con la información verdadera de la exposición que se mostró previamente.

4. ¿Hay sesgo debido a la clasificación errónea de la exposición? (Note que el OR correcto es 3,5).
5. ¿Cuál es la dirección del sesgo, si lo hay?

En general, cuando hay clasificación errónea diferencial de la exposición o de la enfermedad, el sesgo puede ser **hacia el valor nulo** o **alejarse del valor nulo** (véase el siguiente ejemplo de sesgo que se aleja del valor nulo).



Resumen

- Con la clasificación errónea **diferencial**, las sensibilidades y especificidades de la clasificación errónea de **D** difieren por **E** o las sensibilidades y especificidades de la clasificación errónea de **E** difieren por **D**.
- La clasificación errónea diferencial de **D** o **E** puede producir sesgo hacia el valor nulo o alejado del valor nulo.

Definición matemática de la clasificación errónea diferencial

En general, la **clasificación errónea diferencial** se define como la clasificación errónea de la enfermedad y/o la exposición en la cual hay diferencias en las sensibilidades y especificidades de una variable (por ejemplo, la condición de enfermedad) con respecto a las categorías de la otra variable (por ejemplo, la condición de exposición). Para clarificar esta definición, nuevamente presentamos a continuación los cuadros de clasificación errónea que contienen los parámetros de sensibilidad y especificidad que estamos considerando.

Probabilidades de la clasificación errónea de la enfermedad (dada la condición de exposición)

	Verdaderamente E	
Clasificados	**D**	**no D**
D´	$Se_{D\|E}$	$1\text{-}Sp_{D\|E}$
no D´	$1\text{-}Se_{D\|E}$	$Sp_{D\|E}$

	Verdaderamente no E	
Clasificados	**D**	**no D**
D´	$Se_{D\|no\ E}$	$1\text{-}Sp_{D\|no\ E}$
no D´	$1\text{-}Se_{D\|no\ E}$	$Sp_{D\|no\ E}$

Probabilidades de la clasificación errónea de la exposición (dada la condición de enfermedad)

	Verdaderamente D	
Clasificados	**E**	**no E**
E´	$Se_{E\|D}$	$1\text{-}Sp_{E\|D}$
no E´	$1\text{-}Se_{E\|D}$	$Sp_{E\|D}$

	Verdaderamente no D	
Clasificados	**E**	**no E**
E´	$Se_{E\|no\ D}$	$1\text{-}Sp_{E\|no\ D}$
no E´	$1\text{-}Se_{E\|no\ D}$	$Sp_{D\|no\ D}$

Mediante la notación dada en estos cuadros, la **clasificación errónea diferencial** ocurre si es válida alguna de las siguientes cuatro desigualdades:

Clasificación errónea de la enfermedad:

$\mathbf{Se_{D|E} \neq Se_{D|no\ E}}$ $\mathbf{Sp_{D|E} \neq Sp_{D|no\ E}}$

Clasificación errónea de la exposición:

$\mathbf{Se_{E|D} \neq Se_{E|no\ D}}$ $\mathbf{Sp_{E|D} \neq Sp_{E|no\ D}}$

Clasificación errónea independiente tanto de la exposición como de la enfermedad

Dos eventos A y B son independientes si la probabilidad de que ambos ocurran es igual al producto de sus probabilidades individuales.

Dos eventos A y B son independientes si:

$$P(A \text{ y } B) = P(A) * P(B)$$

Cuando consideramos la clasificación errónea de la exposición y la enfermedad, los dos eventos que nos interesan son **A**, la forma en que clasificamos la condición de enfermedad, y **B**, la forma en que clasificamos la condición de exposición. Estos dos eventos son independientes si la probabilidad de clasificar la condición de exposición Y de enfermedad es igual al producto de las probabilidades de clasificar la condición de enfermedad y

la condición de exposición en forma separada.

Clasificación errónea independiente de E y D

$A = (\text{condición de D clasificado})\ B = (\text{condición de E clasificado})$

$\Pr(\text{ de clasificar la condición de exposición Y de enfermedad }) =$

$\Pr(\text{ de clasificar D }) \times \Pr(\text{de clasificar E })$

En muchos estudios epidemiológicos, tiene sentido suponer que las determinaciones de la condición de exposición y enfermedad son más o menos independientes. Normalmente, la condición de enfermedad está determinada por el examen físico o los registros médicos; la condición de exposición está determinada por un proceso separado, a menudo en un momento diferente como, por ejemplo, una entrevista, antecedentes ocupacionales o técnica de medición. Por ejemplo, la forma en que se diagnostica cierto cáncer normalmente no influye ni depende de la manera en que se determinan los antecedentes relativos a la ocupación, al tabaquismo o a la dieta.

Por otro lado, si tanto la condición de enfermedad como de exposición se determinan al mismo tiempo y por informes personales de los individuos o por recuerdos, entonces podrían no ser independientes. Dicha situación surgió en el Estudio de usuarios de las playas de Sydney de 1989, donde se les pidió a los individuos que recordaran una exposición previa por bañarse en el mar y la condición de enfermedad una semana después de que fueron entrevistados en la playa.

Entonces, ¿por qué es importante el supuesto de clasificación errónea independiente? Necesitamos este supuesto para evaluar o corregir convenientemente el sesgo de clasificación errónea cuando, *tanto* la enfermedad *como* la exposición, están clasificadas erróneamente en el mismo estudio. Si se clasifica erróneamente sólo la exposición o sólo la enfermedad, no se necesita el supuesto de independencia.

Preguntas de estudio (P9.7)

1. ¿El supuesto de **clasificación errónea independiente** de la enfermedad y la exposición es equivalente a suponer la **clasificación errónea no diferencial** de la condición de enfermedad o exposición?

Para las siguientes preguntas, presuma que la enfermedad y la exposición están clasificadas erróneamente en forma independiente. Suponga también:

(sensibilidad D|E) = 0,8; (sensibilidad D|no E) = 0,75
(sensibilidad E|D) = 0,9; (especificidad E|D) = 0,95

2. ¿La clasificación errónea es no diferencial o diferencial?
3. ¿Faltan sensibilidades y especificidades en la información anterior?
4. ¿De qué manera puede expresar Pr(D´E´|DE) como el producto de los parámetros de sensibilidad y/o especificidad?
5. Calcule Pr(D´E´|DE) utilizando la información anterior.
6. ¿De qué manera puede expresar Pr(D´E´|no D no E) como el producto de los parámetros de sensibilidad y/o especificidad?
7. Calcule Pr(D´E´|D no E).

Resumen

- Otra forma importante de caracterizar la clasificación errónea implica determinar si hay **clasificación errónea independiente** o no, tanto de la exposición como de la enfermedad.
- La clasificación errónea de la exposición y la enfermedad se define como independiente si:
  Pr(clasificación de D y E|verdaderamente D y E) =
  Pr(clasificación de D|verdaderamente D y E) x Pr(clasificación de E|verdaderamente D y E)
- El supuesto de clasificación errónea independiente permite la evaluación del sesgo de información a partir de la clasificación errónea de la exposición y la enfermedad en forma simultánea.

Más información sobre la clasificación errónea independiente

Este recuadro brinda una descripción matemática de cómo las frecuencias de celda en el cuadro de 2 x 2 clasificado erróneamente pueden expresarse en términos de las frecuencias de celda verdaderas bajo el supuesto de que hay clasificación errónea independiente de la enfermedad y de la exposición. Es posible que no desee leer lo que sigue si no está interesado en los fundamentos matemáticos del sesgo de clasificación errónea.

Hemos definido la **clasificación errónea independiente de la enfermedad y la exposición** de la siguiente manera:

Pr(clasificación de D y E|verdaderamente D y E) =
Pr(clasificación de D|verdaderamente D y E) x Pr(clasificación de E|verdaderamente D y E)

Si suponemos una clasificación errónea independiente como se definió anteriormente, entonces podemos expresar las frecuencias de celda $\mathbf{A^o}$, $\mathbf{B^o}$, $\mathbf{C^o}$ y $\mathbf{D^o}$ en el cuadro de 2 x 2 para la población clasificada erróneamente en términos de las verdaderas frecuencias de celda **A**, **B**, **C** y **D** en el correspondiente cuadro de 2 x 2 para la población clasificada correctamente (es decir, verdadera).

Éstas son las ecuaciones:

$$\mathbf{A^o} = Se_{D|E}Se_{E|D}\mathbf{A} + Se_{D|no\ E}(1-Sp_{E|D})\mathbf{B} + (1-Sp_{D|E})Se_{E|no\ D}\mathbf{C} + (1-Sp_{D|no\ E})(1-Sp_{E|no\ D})\mathbf{D}$$

$$\mathbf{B^o} = Se_{D|E}(1-Se_{E|D})\mathbf{A} + Se_{D|no\ E}Sp_{E|D}\mathbf{B} + (1-Sp_{D|E})(1-Se_{E|no\ D}\mathbf{C} + (1-Sp_{D|no\ E})Sp_{E|no\ D}\mathbf{D}$$

$$\mathbf{C^o} = (1-Se_{D|E})Se_{E|D}\mathbf{A} + (1-Se_{D|no\ E})\ (1-Sp_{E|D})\mathbf{B} + Sp_{D|E}Se_{E|no\ D}\mathbf{C} + Sp_{D|no\ E}(1-Sp_{E|no\ D})\mathbf{D}$$

$$\mathbf{D^o} = (1-Se_{D|E})\ (1-Se_{E|D})\mathbf{A} + (1-Se_{D|no\ E})1\text{-}Sp_{E|D}\mathbf{B} + Sp_{D|E}\ (1-Se_{E|no\ D})\mathbf{C} + Sp_{D|no\ E}Sp_{E|no\ D}\mathbf{D}$$

La primera de estas ecuaciones se deriva mediante la aplicación de la definición de independencia dada anteriormente a los términos de probabilidad en la siguiente expresión para $\mathbf{A^o}$:

$$A^o = A^o_{11} + A^o_{12} + A^o_{21} + A^o_{22}$$
$$= \mathbf{A}\ Pr(D'E'|DE) + \mathbf{B}\ Pr(D'E'|D, \text{no } E) + \mathbf{C}\ Pr(D'E'|\text{no } D, E) + \mathbf{D}\ Pr(D'E'|\text{no } D, \text{no } E)$$

Las otras tres ecuaciones (es decir, para $\mathbf{B^o}$, $\mathbf{C^o}$ y $\mathbf{D^o}$) se derivan en forma similar. Estas ecuaciones representan un sistema de cuatro ecuaciones para las frecuencias de celda clasificadas erróneamente $\mathbf{A^o}$, $\mathbf{B^o}$, $\mathbf{C^o}$ y $\mathbf{D^o}$ en términos de cuatro frecuencias de celda verdaderas desconocidas **A**, **B**, **C** y **D**. Resolviendo este sistema de ecuaciones para obtener las expresiones de **A**, **B**, **C** y **D** en términos de $\mathbf{A^o}$, $\mathbf{B^o}$, $\mathbf{C^o}$ y $\mathbf{D^o}$, podemos derivar fórmulas para ajustar el cálculo observado de una medida de efecto (por ejemplo, un *odds ratio*) para corregir la clasificación errónea. Dichas fórmulas de corrección se describen en explicaciones posteriores sobre la clasificación errónea.



Auto-evaluación (P9.8)

Coloque **verdadero** o **falso** en las siguientes afirmaciones.

1. Si hay clasificación errónea de la condición de enfermedad, pero no de la condición de exposición en un estudio de seguimiento, y si la probabilidad de sensibilidad es igual para el grupo expuesto y el grupo no expuesto, entonces todo sesgo existente (debido a la clasificación errónea) debe ser hacia el valor nulo. **???**
2. Si hay clasificación errónea diferencial de la condición de enfermedad, pero no de la condición de exposición en un estudio de casos y controles, entonces puede haber sesgo en el cálculo del *odds ratio*, que puede ser hacia el valor nulo o alejarse de éste. **???**
3. Suponga que hay clasificación errónea independiente, tanto de la condición de exposición como de la condición de enfermedad. Entonces:

   $P(D', \text{no } E' | D, \text{no } E) = (Se_{D|\text{no } E})*(Sp_{E|D})$. **???**

9-3 Sesgo de información (continuación)

Evaluación cuantitativa del sesgo de clasificación errónea

Supongamos que sabemos que existe una probable clasificación errónea de la exposición o de la enfermedad que podría sesgar los resultados de nuestro estudio. ¿Cómo podemos corregir cuantitativamente dicho sesgo para obtener una medida de efecto ajustada que ya no esté sesgada?

Para cuantificar el sesgo, debemos ajustar las frecuencias de celda para los datos observados a fin de obtener un cuadro de doble entrada de las frecuencias de celda corregidas de las cuales podamos calcular una medida corregida de efecto. Luego podemos comparar el cálculo observado y posiblemente sesgado con nuestro cálculo corregido para determinar el monto del posible sesgo y su dirección.

Datos observados (clasificados erróneamente)

	E'	No E'
D'	a	b
No D'	c	d

$$\hat{OR} = \frac{ad}{bc} \qquad \hat{RR} = \frac{a/(a+c)}{b/(b+d)}$$

Datos corregidos (es decir, ajustados)

	E	No E
D	A	B
No D	C	D

$$\hat{OR}_{(ajustado)} = \frac{AD}{BC} \qquad \hat{RR}_{(ajustado)} = \frac{A/(A+C)}{B/(B+D)}$$

Preguntas de estudio (P9.9)

Utilice la información del siguiente cuadro sobre las medidas de efecto observadas y corregidas para determinar si el sesgo observado es hacia el valor nulo o se aleja de éste.

Número de pregunta	Efecto observado	Efecto corregido	¿Hacia el valor nulo?	¿Se aleja del valor nulo?
a.	2,2	1,7		
b.	2,5	3,8		
c.	4,0	6,1		
d.	4,1	1,2		
e.	0,5	0,9		
f.	0,8	0,9		
g.	0,3	0,2		
h.	0,7	0,1		

Suponga que hemos determinado que cualquier sesgo existente es el resultado de una clasificación errónea no diferencial de la exposición o de la enfermedad.

1. ¿Necesitamos obtener una medida de efecto ajustada que corrija dicho sesgo?
2. ¿Cómo podemos determinar si la clasificación errónea es no diferencial?

Suponga que hemos determinado que hay clasificación errónea diferencial de la exposición o de la enfermedad.

3. ¿Necesitamos obtener una medida de efecto ajustada que corrija el posible sesgo de clasificación errónea?

En las siguientes presentaciones, brindamos fórmulas para obtener cálculos corregidos. Estas fórmulas requieren cálculos confiables de sensibilidad y especificidad. ¿Cómo podemos determinar sensibilidad y especificidad si todo lo que tenemos son los datos observados? Una opción es tomar una pequeña muestra de los datos observados para los que se determina la verdadera condición de enfermedad y exposición, de forma que se puedan calcular las sensibilidades y las especificidades (una desventaja de este enfoque es que la muestra podría ser demasiado pequeña como para brindar valores confiables de sensibilidad y especificidad). Otra opción es determinar los parámetros de sensibilidad y especificidad a partir de datos separados que se deriven de un estudio previo que haya incluido las mismas variables. Una tercera opción es simplemente hacer una estimación razonable de los parámetros de sensibilidad y especificidad, en base al conocimiento clínico o de otra naturaleza sobre las variables del estudio. Una cuarta opción es realizar lo que a menudo se denomina análisis de sensibilidad con varias estimaciones razonables para determinar un rango de posibles sesgos.

Opciones para calcular
la sensibilidad y la especificidad:

1. Una submuestra de los datos del estudio
2. Una muestra separada de otro estudio
3. Un cálculo aproximado que se base en la teoría/experiencia clínica o de otra naturaleza
4. Un análisis de sensibilidad que considere varios cálculos aproximados

Preguntas de estudio (P9.9) (continuación)

Considere los siguientes resultados de un análisis de sensibilidad para corregir la clasificación errónea que se supone que es no diferencial.

OR observado	¿Sensibilidad?	¿Especificidad?	OR corregido
1,5	80%	80%	3,5
1,5	80%	90%	2,5
1,5	90%	80%	2,5
1,5	90%	90%	1,8

4. ¿Por qué los resultados observados comparados con los resultados corregidos correspondientes ilustran un sesgo hacia el valor nulo?
5. ¿Qué valores de sensibilidad y especificidad están asociados con un sesgo mayor?
6. ¿Qué valores de sensibilidad y especificidad están asociados con un sesgo menor?
7. Sobre la base del análisis de sensibilidad descrito anteriormente, ¿cómo podría decidir qué OR corregido es "mejor"?

Resumen

- Podemos corregir el sesgo de clasificación errónea calculando una medida de efecto ajustada a partir de un cuadro de doble entrada, donde ya se haya corregido la clasificación errónea encontrada en las frecuencias de celda observadas.

- La corrección requiere un cálculo preciso de los parámetros de sensibilidad y especificidad.
- Opciones para calcular la sensibilidad y la especificidad:
  - una submuestra de los datos del estudio
  - una muestra separada de otro estudio
  - un cálculo aproximado que se base en la teoría/experiencia clínica o de otra naturaleza
  - un análisis de sensibilidad que considere varios cálculos aproximados

Corrección de la clasificación errónea no diferencial de la enfermedad

Este cuadro contiene los datos observados de un estudio hipotético de cohorte que analiza la relación entre el sexo y la enfermedad de úlcera péptica.

Datos observados (clasificados erróneamente)

	Hombre	Mujer	
UPE´	380	240	UPE = úlcera péptica
no UPE´	620	760	
Total	1000	1000	

$$\hat{RR} = \frac{(380/1000)}{(240/1000)} = 1.58$$

Suponga que la variable de exposición, el sexo del individuo, no ha sido clasificada erróneamente, pero que el diagnóstico de la úlcera péptica tiene cierto grado de clasificación errónea. ¿Cómo podemos corregir la posible clasificación errónea de la enfermedad para obtener un cálculo ajustado de la razón de riesgos que brinde un cálculo preciso de la relación entre la exposición y la enfermedad?

Primero debemos obtener cálculos confiables de las sensibilidades y especificidades de la clasificación errónea de la enfermedad. Suponga que realizamos un examen físico más exhaustivo de una submuestra de 200 individuos para obtener una determinación del "estándar de oro" de la condición de úlcera péptica. La clasificación de estos 200 individuos se muestra a continuación. Suponga que se obtienen cuadros separados de clasificación errónea de hombres y mujeres.

Clasificación errónea de la enfermedad (submuestra)

Hombres:

	Estándar de oro: UPE	Estándar de oro: no UPE	Total
UPE´	32	6	38
no UPE´	8	54	62
Total	40	60	100

Mujeres:

	Estándar de oro: UPE	Estándar de oro: no UPE	Total
UPE´	16	8	24
no UPE´	4	72	76
Total	20	80	100

Preguntas de estudio (P9.10)

Las siguientes preguntas se refieren al cuadro titulado "Clasificación errónea de la enfermedad (submuestra)".

1. ¿Cuál es el valor de la sensibilidad y especificidad para los hombres?
2. ¿Cuál es el valor de la sensibilidad y especificidad para las mujeres?

3. ¿La clasificación errónea es no diferencial o diferencial?

Como en nuestro ejemplo, cuando la clasificación errónea es no diferencial y tenemos cálculos confiables de los parámetros de sensibilidad y especificidad, podemos transformar las frecuencias de celda clasificadas erróneamente de nuestros datos observados en frecuencias de celda "corregidas", de las cuales se puede calcular una medida ajustada y corregida de efecto.

Podemos escribir una fórmula relativamente simple de las frecuencias de celda corregidas si, como en nuestro ejemplo, sólo la variable de enfermedad está clasificada erróneamente.

Datos observados (clasificados)

	Hombre	Mujer
UPE	a= 380	b= 240
no UPE	c= 620	d= 760
Total	1000	1000

Datos corregidos

	Hombre	Mujer
UPE	A	B
no UPE	C	D
Total	1000	1000

$$A = [a\ SpD - c\ (1 - SpD)] / q$$
$$B = [b\ SpD - d\ (1 - SpD)] / q$$
$$C = [c\ SeD - a\ (1 - SeD)] / q$$
$$D = [d\ SeD - b\ (1 - SeD)] / q$$
$$q = SeD + SpD - 1$$

Para calcular las frecuencias de celda corregidas, sustituimos las frecuencias de celda observadas en la fórmula, tal como se muestra a continuación. También sustituimos nuestros valores calculados de la sensibilidad y especificidad en esta fórmula, de la siguiente manera:

$$A = [380\,0.90 - 620(1 - 0.90)] / 0.70$$
$$B = [240\,0.90 - 760(1 - 0.90)] / 0.70$$
$$C = [620\,0.80 - 380(1 - 0.80)] / 0.70$$
$$D = [760\,0.80 - 240(1 - 0.80)] / 0.70$$
$$q = 0.70$$
$$SeD = 0.80 \quad SpD = 0.90$$

Aquí se muestran las resultantes frecuencias de celda corregidas.

Datos Corregidos

	Hombre	Mujer
UPE	400	200
no UPE	600	800
Total	1000	1000

Preguntas de estudio (P9.10) (continuación)

4. Utilizando las frecuencias de celda corregidas, calcule la razón de riesgos ajustada que corrige la clasificación errónea de la enfermedad.
5. ¿Cuál es la dirección del sesgo?
6. Aunque no haya problema al transformar las frecuencias de celda en este ejemplo, ¿qué debería llamarle la atención acerca del denominador q = sensibilidad(D) + especificidad(D) – 1 en las fórmulas para A, B, C y D?

Resumen

- La clasificación errónea no diferencial de la enfermedad sin clasificación errónea de la exposición puede corregirse cuando están disponibles los cálculos de sensibilidad y especificidad.
- Este procedimiento implica transformar las frecuencias de celda clasificadas erróneamente en frecuencias de celda "corregidas" de las cuales puede calcularse una medida corregida de efecto que ya no esté sesgada.
- El procedimiento requiere entradas de cálculos de las sensibilidades y especificidades de la clasificación errónea de la enfermedad, así como también frecuencias de celda en el cuadro de 2 x 2 observado que relaciona la exposición y la enfermedad.
- Hay una fórmula algebraica explícita para el procedimiento de corrección, aunque estos cálculos se realizan con mayor eficacia mediante el uso de un programa de computación.

Corrección por computadora de la clasificación errónea no diferencial de la enfermedad

Existen programas de computación disponibles que permiten la clasificación errónea diferencial tanto de la exposición como de la enfermedad, de modo que la clasificación errónea no diferencial de la enfermedad es un caso especial del programa en el que las sensibilidades y especificidades para la exposición, dado el estatus de enfermedad, (1) son todas especificadas como igual a 1; (2) las sensibilidades correspondientes a la enfermedad, dado el estatus de exposición, son especificadas como iguales; y (3) las especificidades correspondientes a la enfermedad, dado el estatus de exposición, también son especificadas como iguales. Es decir, la entrada de datos en el programa sería:

1. Frecuencias de celda observadas **a**, **b**, **c** y **d**.
2. Sensibilidades y especificidades de la exposición dada la enfermedad:

   $Se_{E|D} = Se_{E|no\ D}\ (=Se_E) = 1$
   $Sp_{E|D} = Sp_{E|no\ D}\ (=Sp_E) = 1$

3. Sensibilidades iguales de la enfermedad dada la exposición:

   $Se_{D|E} = Se_{D|no\ E}\ (= Se_D)$

4. Especificidades iguales de la enfermedad dada la exposición:

   $Sp_{D|E} = Sp_{D|no\ E}\ (= Sp_D)$

Problemas con la corrección de la clasificación errónea no diferencial de la enfermedad

Las fórmulas para corregir la clasificación errónea no diferencial de la enfermedad cuando no hay clasificación errónea de la exposición pueden dar resultados inadecuados para las frecuencias de celda corregidas **A**, **B**, **C** y **D** bajo las siguientes circunstancias:

1. **Se + Sp = 1** (es decir, **q** = Se + Sp - 1 = 0, donde **q** es el denominador de la fórmula de corrección para cada una de las cuatro celdas y **Se** y **Sp** son la sensibilidad y la especificidad de la clasificación errónea de la **enfermedad**).
2. **Sp < c / (a + c)**, en cuyo caso el valor corregido para **A** es negativo.
3. **Sp < d / (b + d)**, en cuyo caso el valor corregido para **B** es negativo.
4. **Se < a / (a + c)**, en cuyo caso el valor corregido para **C** es negativo.
5. **Se < b / (b + d)**, en cuyo caso el valor corregido para **D** es negativo.

Corrección de la clasificación errónea no diferencial de la exposición

El siguiente cuadro describe los datos observados de un estudio hipotético de casos y controles que analizó la relación entre los antecedentes informados de úlcera péptica, la variable de exposición y el cáncer de estómago. Los controles eran personas con cáncer de colon, obtenidos de la misma población de hospitales que los casos de cáncer de estómago. Suponga que la condición de caso y control ha sido diagnosticada correctamente, pero que los antecedentes de úlcera péptica informados por los individuos involucra algún tipo de clasificación errónea.

Datos observados (clasificados)

	UPE´	no UPE´	**Total**
Casos	75	225	300
Controles	30	270	300

$$\hat{OR} = \frac{(75 \times 270)}{(225 \times 30)} = 3.0$$

UPE = enfermedad de úlcera péptica

Como en la actividad anterior, sólo una de las variables de enfermedad y exposición está clasificada erróneamente. Esta vez se trata de la variable de exposición. Una vez más, preguntamos, "¿cómo podemos corregir la posible clasificación errónea en esta situación?" Podemos obtener una respuesta corregida siempre que tengamos cálculos confiables de la sensibilidad y la especificidad para la clasificación errónea de la exposición.

Suponga que un estudio separado que analizó los factores de riesgo para la enfermedad de úlcera péptica, realizado con anterioridad, había evaluado la medida de la clasificación errónea de la enfermedad informada, obteniendo una determinación del estándar de oro de la condición de úlcera péptica en una muestra de 200 individuos. Suponga, además, que el cuadro de clasificación errónea es el que se muestra a continuación.

Cuadro de clasificación errónea de la exposición: (**estudio separado**)

	Estándar de oro		
	UPE	no UPE	Total
UPE'	48	7	55
no UPE'	12	133	145
Total	60	140	200

Preguntas de estudio (P9.11)

1. ¿Cuál es el valor de la sensibilidad y especificidad de la clasificación errónea de la condición de UPE?
2. ¿Cómo puede determinar si la clasificación errónea es no diferencial o diferencial?

Es razonable suponer que la clasificación errónea de la exposición es no diferencial en este ejemplo. Ahora podemos transformar las frecuencias de celda clasificadas erróneamente y que contienen nuestros datos observados, en frecuencias de celda corregidas de las que se puede calcular un *odds ratio* corregido. A continuación se presentan las expresiones para las frecuencias de celda corregidas si sólo la variable de exposición está clasificada erróneamente.

$$A = [a\ SpE - b(1 - SpE)]/q$$
$$B = [b\ SeE - a(1 - SeE)]/q$$
$$C = [c\ SpE - d(1 - SpE)]/q$$
$$D = [d\ SeE - c(1 - SeE)]/q$$
$$q = SeE + SpE - 1 \quad SeE = 0.80 = \text{Sensibilidad E}$$
$$SpE = 0.95 = \text{Especificidad E}$$

Para calcular las frecuencias de celda corregidas, sustituimos las frecuencias de celda observadas en la fórmula, tal como se muestra a continuación.

También sustituimos nuestros valores calculados de sensibilidad y especificidad en esta fórmula, de la siguiente manera:

$$A = [75\ 0.95 - 225(1 - 0.95)]/0.75 = 80$$
$$B = [225\ 0.80 - 75(1 - 0.80)]/0.75 = 220$$
$$C = [30\ 0.95 - 270(1 - 0.95)]/0.75 = 20$$
$$D = [270\ 0.80 - 30(1 - 0.80)]/0.75 = 280$$
$$q = 0.75 \qquad SeE = 0.80 = \text{Sensibilidad E}$$
$$SpE = 0.95 = \text{Especificidad E}$$

Aquí se muestran las frecuencias resultantes de celda corregidas.

Datos Corregidos

	UPE′	no UPE′	Total
Casos	80	220	300
Controles	20	280	300

Preguntas de estudio (P9.11) (continuación)

(Nota: no hay pregunta 3 en las preguntas de estudio P9.11).

4. Mediante las frecuencias de celda corregidas, calcule el *odds ratio* ajustado que corrige la clasificación errónea de la enfermedad.

5. ¿Cuál es la dirección del sesgo?

Resumen

- La clasificación errónea no diferencial de la exposición sin clasificación errónea de la enfermedad puede corregirse cuando están disponibles los cálculos de sensibilidad y especificidad.
- Este procedimiento implica transformar las frecuencias de celda clasificadas erróneamente en frecuencias de celda "corregidas" de las cuales puede calcularse una medida corregida de efecto.
- El procedimiento requiere entradas de cálculos de sensibilidades y especificidades de la clasificación errónea de la exposición, así como también frecuencias de celda en el cuadro de 2 x 2 observado que relaciona la exposición y la enfermedad.
- Hay una fórmula algebraica explícita para el procedimiento de corrección, aunque estos cálculos se realizan con mayor eficacia mediante el uso de un programa de computación.

Corrección por computadora de la clasificación errónea no diferencial de la exposición

Existen programas de computación disponibles que permiten la clasificación errónea diferencial tanto de la exposición como de la enfermedad. La clasificación errónea no diferencial de la exposición es un caso especial del programa en el que las sensibilidades y especificidades de la enfermedad, dado el estatus de exposición, (1) son especificadas como igual a 1; (2) las sensibilidades correspondientes a la exposición, dado el estatus de enfermedad, son especificadas como iguales; y (3) las especificidades correspondientes a la exposición, dado el estatus de enfermedad, también son especificadas como iguales. Es decir, la entrada de datos en el programa sería:

1. Frecuencias de celda observadas **a**, **b**, **c** y **d**.
2. Sensibilidades y especificidades de la enfermedad dada la exposición:

   $Se_{D|E} = Se_{D|no\ E}\ (= Se_D) = 1$
   $Sp_{D|E} = Sp_{D|no\ E}\ (= Sp_D) = 1$

3. Sensibilidades iguales de la exposición dada la enfermedad:

   $Se_{E|D} = Se_{E|no\ D}\ (= Se_E)$

4. Especificidades iguales de la exposición dada la enfermedad:

   $Sp_{E|D} = Sp_{E|no\ D}\ (= Sp_E)$

Problemas con la corrección de la clasificación errónea no diferencial de la exposición

Las fórmulas para corregir la clasificación errónea no diferencial de la exposición, cuando no hay clasificación errónea de la enfermedad, pueden dar resultados inadecuados de frecuencias de celda corregidas **A**, **B**, **C** y **D** bajo las siguientes circunstancias:

1. **Se + Sp = 1** (es decir, **q** = Se + Sp - 1 = 0, donde **q** es el denominador de la fórmula de corrección para cada una de las cuatro celdas y **Se** y **Sp** son la sensibilidad y la especificidad de la clasificación errónea de la **exposición**).
2. **Sp < c / (a + c)**, en cuyo caso el valor corregido para **A** es negativo.

3. **Sp < d / (b + d)**, en cuyo caso el valor corregido para **B** es negativo.
4. **Se < a / (a + c)**, en cuyo caso el valor corregido para **C** es negativo.
5. **Se < b / (b + d)**, en cuyo caso el valor corregido para **D** es negativo.

Corrección de la clasificación errónea no diferencial de la exposición y la enfermedad

El siguiente cuadro considera datos hipotéticos de cohorte de individuos encuestados en las playas de cierta isla del Caribe. Este estudio se asemeja mucho al estudio de usuarios de las playas de Sydney descrito en la Lección 2.

Datos observados (recuerdos)

	Nivel de contaminación: Alto'	Bajo'	**Total**
Enfermo'	**400**	**375**	**775**
No enfermo'	**600**	**1125**	**1725**
Total	**1000**	**1500**	**2500**

$$\hat{RR} = \frac{(400/1000)}{(375/1500)} = 1.6$$

El cuadro describe los datos observados de 2500 individuos que se bañaron en las playas el día de la entrevista. El nivel de contaminación para cada individuo se basó en conteos promedio de la calidad del agua tomados de tres muestras especificadas previamente el día en que el individuo fue entrevistado en la playa. Una semana después, se les pidió por teléfono que recordaran si habían desarrollado tos, gripe o enfermedad gastrointestinal desde su entrevista en la playa.

Ya que la información sobre la enfermedad se obtuvo por el recuerdo de los individuos, es razonable esperar que posiblemente algunos individuos hayan informado incorrectamente la condición de enfermedad. Asimismo, como el nivel de contaminación por individuo podría no corresponder al momento del día que se bañó realmente dicha persona, y puesto que la técnica de medición estaba sujeta a error, también es probable que existiera clasificación errónea del nivel de contaminación.

Suponga que se investigó exhaustivamente una submuestra de individuos del estudio para determinar su verdadera condición de enfermedad, y se descubrió que la sensibilidad era de 80% y la especificidad de 90% de la clasificación errónea de la enfermedad. Suponga además que un estudio anterior sobre las mediciones de la calidad del agua de la playa indicó que la sensibilidad y la especificidad de la clasificación errónea del nivel de contaminación eran del 90% y 75%, respectivamente.

Información Enfermedad Clasificación errónea	Nivel de Contaminación Clasificación errónea
estudio submuestra indicó:	estudio anterior indicó:
Sensibilidad = 80%	Sensibilidad = 90%
Especificidad = 90%	Especificidad = 70%

Preguntas de estudio (P9.12)

1. Si hubiera información adecuada para este estudio, ¿cómo determinaría si la clasificación errónea de la enfermedad o la exposición fue no diferencial o diferencial?
2. Sobre la base de la información brindada para este ejemplo, ¿es posible determinar explícitamente si la clasificación errónea de E o D es no diferencial o diferencial?
3. Sobre la base de la información brindada, ¿es necesario suponer una clasificación errónea no diferencial o diferencial para corregir el sesgo de clasificación errónea?
4. ¿Cree que hubo una clasificación independiente del nivel de contaminación y de la condición de enfermedad?

Las respuestas a las preguntas de estudio indican que es razonable suponer que la clasificación errónea, tanto de la exposición como de la enfermedad, es no diferencial y que las probabilidades de clasificación son independientes. Según estos supuestos, las frecuencias de celda observadas pueden transformarse en frecuencias de celda corregidas mediante las siguientes expresiones:

$$A = \{SpD[(a+b)SpE - b] - (1-SpD)[(c+d)SpE - d]\}/q^*$$
$$B = \{SpD[(a+b)SeE - a] - (1-SpD)[(c+d)SeE - c]\}/q^*$$
$$C = \{SeD[(c+d)SpE - d] - (1-SeD)[(a+b)SpE - b]\}/q^*$$
$$D = \{SeD[(c+d)SeE - c] - (1-SeD)[(a+b)SeE - a]\}/q^*$$
$$q^* = (SeD + SpD - 1)(SeE + SpE - 1)$$

Estas fórmulas son más complicadas que aquellas en las que sólo la exposición o sólo la enfermedad estaban clasificadas erróneamente. Para calcular las frecuencias de celda corregidas, sustituimos las frecuencias de celda observadas en las fórmulas, tal como se muestra a continuación. También sustituimos nuestros valores calculados de la sensibilidad y la especificidad en estas fórmulas, de la siguiente manera:

$$A = \{0.90[(775)0.70 - 375] - (1-0.90)[(1725)0.70 - 1125]\}/0.42$$
$$B = \{0.90[(775)0.90 - 400] - (1-0.90)[(1725)0.90 - 600]\}/0.42$$
$$C = \{0.80[(1725)0.70 - 1125] - (1-0.80)[(775)0.70 - 375]\}/0.42$$
$$D = \{0.80[(1725)0.90 - 600] - (1-0.80)[(775)0.90 - 400]\}/0.42$$
$$q^* = 0.42$$

Aquí se muestran las frecuencias resultantes de celda corregidas.

Celda corregida

	Alto	Bajo
enfermo	339.3	410.8
no enfermo	77.4	1672.6
Total	416.7	2083.4

Preguntas de estudio (P9.12) (continuación)

5. Mediante las frecuencias de celda corregidas, calcule la razón de riesgos ajustada que corrige la clasificación errónea de la enfermedad y la exposición.
6. ¿Cuál es la dirección del sesgo?
7. ¿Qué observa sobre el alcance del sesgo cuando tanto la exposición como la enfermedad están clasificadas erróneamente?
8. Aunque no hay problema con q* en este ejemplo, ¿qué debería llamar su atención sobre el uso de q* en la fórmula de cálculo?

Resumen

- La clasificación errónea no diferencial tanto de la exposición como de la enfermedad puede corregirse cuando están disponibles los cálculos de sensibilidad y especificidad.
- Las fórmulas utilizadas para los cálculos corregidos suponen que la clasificación errónea de la exposición y la enfermedad son independientes.
- El procedimiento requiere entradas de cálculos de sensibilidades y especificidades para la clasificación errónea de la exposición y la enfermedad, así como también frecuencias de celda en el cuadro de 2 x 2 observado que relaciona la exposición y la enfermedad.
- Hay una fórmula algebraica explícita para el procedimiento de corrección, aunque estos cálculos se realizan con mayor eficacia mediante el uso de un programa de computación.

Problemas con la corrección de la clasificación errónea no diferencial tanto de la enfermedad como de la exposición

Las fórmulas para corregir la clasificación errónea no diferencial tanto de la enfermedad como de la exposición son las siguientes:

$$\mathbf{A = \{SpD[(a + b)\ SpE - b] - (1 - SpD)[(c + d)\ SpE - d]\} / q^*}$$
$$\mathbf{B = \{SpD[(a + b)\ SeE - a] - (1 - SpD)[(c + d)\ SeE - c]\} / q^*}$$
$$\mathbf{C = \{SeD[(c + d)\ SpE - d] - (1 - SeD)[(a + b)\ SpE - b]\} / q^*}$$
$$\mathbf{D = \{SeD[(c + d)\ SeE - c] - (1 - SeD)[(a + b)\ SeE - a]\} / q^*}$$

donde $\mathbf{q^* = (SeD + SpD - 1)\ (\ SeE + SpE - 1)}$

Estas fórmulas pueden dar resultados inadecuados de frecuencias de celda corregidas **A**, **B**, **C** y **D** en las siguientes circunstancias:

1. SeD + SpD = 1 o SeE + SpE = 1 (es decir, q* = 0)
2. (a+b+c+d)SpDSpE - (b+d)SpD - (c+d)SpE + d < 0, en cuyo caso el valor corregido para **A** es negativo.
3. (a+b+c+d)SpDSeE - (a+c)SpD - (c+d)SeE + c < 0, en cuyo caso el valor corregido para **B** es negativo.
4. (a+b+c+d)SeDSpE - (b+d)SeD - (a+b)SpE + b < 0, en cuyo caso el valor corregido para **C** es negativo.
5. (a+b+c+d)SeDSeE - (a+c)SeD - (a+b)SeE + a < 0, en cuyo caso el valor corregido para **D** es negativo.

Ahora se brinda un ejemplo con valores negativos de frecuencias de celda corregidas. Suponga que el cuadro observado de 2 x 2 de frecuencias de celda es el siguiente:

Datos observados (es decir, clasificados)

	Se bañó	**No se bañó**
Enfermo'	**a** = 500	**b** = 150
No enfermo'	**c** = 1500	**d** = 750
Total	2000	900

Suponga, además, que se presume la clasificación errónea no diferencial tanto de la exposición como de la enfermedad y que los valores calculados para sensibilidades y especificidades son los siguientes:

SeD = 0,90; SpD = 0,80; SeE = 0,85; SpE = 0,95.

Ahora utilizamos las fórmulas de corrección brindadas anteriormente para obtener las frecuencias de celda corregidas para el siguiente cuadro de 2 x 2:

Cuadro corregido	**Se bañó**	**No se bañó**
Enfermo	**A**	**B**
No enfermo	**C**	**D**

Ahora verificamos las cinco condiciones (1-5) descritas anteriormente que determinan si alguna de las frecuencias de celda corregidas no estará definida o será negativa:

Continúa en la página siguiente.

1. SeD + SpD = 0,90 + 0,80 = 1,70 > 1, y SeE + SpE = 0,85 + 0,95 > 1, de forma que q* = (SeD + SpD - 1) x (SeE + SpE - 1) = 0,70 x 0,80 = 0,56 > 0, por lo tanto, no hay problema asociado al denominador, q*

.

2. (a+b+c+d)SpDSpE - (b+d)SpD - (c+d)SpE + d = 2900 x 0,80 x 0,95 – 900 x 0,80 – 2250 x 0,95 + 750 = 2204 - 2857 + 750 = 96,5 > 0, en cuyo caso el valor corregido para **A** debería ser positivo.
3. (a+b+c+d)SpDSeE - (a+c)SpD - (c+d)SeE + c = 2900 x 0,80 x 0,85 – 2000 x 0,80 – 2250 x 0,85 + 1500 = 1972 – 3512,5 + 1500 = -40,5 < 0, en cuyo caso el valor corregido para **B** debería ser negativo.
4. (a+b+c+d)SeDSpE - (b+d)SeD - (a+b)SpE + b = 2900 x 0,90 x 0,95 – 900 x 0,90 – 650 x 0,95 + 150 = 2479,5 – 1427,5 + 150 = 1202 > 0, en cuyo caso el valor corregido para **C** debería ser positivo.
5. (a+b+c+d)SeDSeE - (a+c)SeD - (a+b)SeE + a = 2900 x 0,90 x 0,85 – 2000 x 0,90 – 650 x 0,85 + 500 = 2218,5 – 2352,5 + 500 = -366 > 0, en cuyo caso el valor corregido para **D** debería ser positivo.

Problemas con la corrección de la clasificación errónea no diferencial tanto de la enfermedad como de la exposición (continuación)

Ahora utilizamos la fórmula de corrección real para demostrar que las verificaciones anteriores realmente generaron los valores positivos de A, C y D, y un valor negativo de B:

A = { SpD [(a+b)SpE - b] - (I-SpD) [(c+d)SpE - d] } / q*
= { 0,80 [(650)0,95 - 150] - (1-0,80) [(2250)0,95 - 750] }/0,56
= 96,5/0,56 = 172,3
B = { SpD [(a+b)SeE - a] - {1-SpD) [(c+d)SeE - c] } / q*
= { 0,80 [(650)0,85 - 500] - {1-0,80) [(2250)0,85 - 1500 }/0,56
= -40,5/0,56 = -72,3
C = { SeD [(c+d)SpE - d] - (I-SeD) [(a+b)SpE - b] } / q*
= { 0,90 [(2250)0,95 - 750] - (1-0,90) [(650)0,95 - 150] }/0,56
= 1202,0/0,56 = 2146,4
D = { SeD [(c+d)SeE - c] - (I-SeD) [(a+b)SeE - a] } / q*
= { 0,90 [(2250)0,85 - 1500] - (1-0,90) [(650)0,85 - 500] }/0,56
= 366/0,56 = 653,6

Por lo tanto, el cuadro corregido es incorrecto porque la celda B es negativa:

Cuadro corregido (incorrectamente)	**Se bañó**	**No se bañó**
Enfermo	172,3	-72,3
No enfermo	2146,4	653,6
Total	2318,7	581,3

Auto-evaluación (P9.13)

Utilice los datos “observados” en el siguiente cuadro para responder las preguntas que se consignan a continuación. Suponga que hay clasificación errónea no diferencial de la enfermedad, SeD = 0,8; SpD = 0,7 y que no hay clasificación errónea de la exposición.

1. Calcule el *odds ratio* que corrige la clasificación errónea o mala clasificación **???**
2. El sesgo de clasificación errónea **???** valor nulo.

Opciones
1,4 2,3 5,1 7,4 se aleja del es hacia el

	E	no E	
D'	70	30	100
no D'	50	50	100
Total	120	80	200

Coloque **verdadero** o **falso** en las siguientes afirmaciones.

3. Es posible obtener resultados indeterminados cuando se trata de corregir la clasificación errónea no diferencial, incluso si se encuentran disponibles buenos cálculos de las probabilidades de sensibilidad y especificidad. . . . . **???**
4. Si la clasificación errónea es no diferencial e independiente, con SpD=SeD=SeE=SpE=0,5; entonces cada persona en la población objetivo tiene la misma probabilidad de ser clasificada erróneamente en cualquiera de las cuatro celdas del cuadro de 2 x 2. . . . . **???**
5. Para la situación descrita anteriormente, es posible obtener un cálculo ajustado del *odds ratio* que corregirá la clasificación errónea o mala clasificacion. . . . . **???**

9-4 Sesgo de información (continuación)

Corrección de la clasificación errónea diferencial de la exposición y/o la enfermedad

Previamente, ejemplificamos la clasificación errónea diferencial de la exposición mediante datos de casos y controles hipotéticos sobre la relación entre la dieta y la enfermedad coronaria. Supusimos que conocíamos el cuadro de 2 x 2 que describe la **verdadera** condición de exposición y enfermedad; luego determinamos el cuadro de 2 x 2 para los datos observados sobre la base de una clasificación errónea diferencial de la exposición, pero no de la enfermedad.

Datos observados (clasificados)

	Bajo'	Alto'	Total
EC	600	400	1000
no EC	275	725	1000

$\widehat{OR} = \frac{600 \times 725}{400 \times 275} = 2.6$

Aquí, en forma más realista, comenzamos suponiendo que tenemos los datos observados, como se muestra a continuación. Deseamos corregir la clasificación errónea diferencial, transformando las frecuencias de celda observadas en frecuencias de celda corregidas, de las cuales podemos obtener un *odds ratio* que corrige la mala clasificación diferencial. Suponga, como antes, que un caso de EC, al que le preocupan los motivos de su enfermedad, no tiene tantas probabilidades de sobrestimar su consumo de frutas y verduras como un control.

Suponga que de una submuestra o de un estudio separado, las sensibilidades y especificidades de la clasificación errónea de la exposición son conocidas.

Datos observados (clasificados)

	Bajo'	Alto'	Total
EC	a= 600	b= 400	1000
no EC	c= 275	d= 725	1000

$$\hat{OR} = \frac{(600 \times 725)}{(400 \times 275)} = 3.95$$

Frecuencias de celda corregidas

	Bajo	Alto	Total
EC	A	B	1000
no EC	C	D	1000

$$\hat{OR}_{ajustado} = \frac{(A \times D)}{(B \times C)}$$

Suponga:

Un caso no tiene tantas probabilidades de sobrestimar su consumo de frutas y verduras como un control

Casos: Sensibilidad: = 96.7%, Especificidad = 95%

Controles: Sensibilidad: = 80%, Especificidad = 95%

Preguntas de estudio (P9.14)

Las siguientes preguntas están basadas en las sensibilidades y especificidades descritas anteriormente.

1. ¿Hay clasificación errónea no diferencial o diferencial de la exposición?
2. ¿Qué indica la información de sensibilidad sobre la probabilidad de que un caso sobrestime su consumo de frutas y verduras, en comparación con un control?

Cuando la clasificación errónea es diferencial, pueden derivarse fórmulas generales para calcular las frecuencias de celda corregidas, **A**, **B**, **C**, y **D** mayúsculas, en términos de las frecuencias de celda observadas, **a**, **b**, **c**, y **d** minúsculas. (Nota: en el recuadro que se encuentra al final de esta actividad están las fórmulas generales que corrigen la clasificación errónea diferencial, tanto de la exposición como de la enfermedad). Aquí se muestran las resultantes frecuencias de celda corregidas.

Datos observados (clasificados)

	Bajo'	Alto'	Total
EC	a= 600	b= 400	1000
no EC	c= 275	d= 725	1000

$$\hat{OR} = \frac{(600 \times 725)}{(400 \times 275)} = 3.95$$

Frecuencias de celda corregidas

	Bajo	Alto	Total
EC	599.7	400.2	1000
no EC	300.0	700	1000

$$\hat{OR}_{ajustado} = \frac{(A \times D)}{(B \times C)}$$

Preguntas de estudio (P9.14) (continuación)

Las siguientes preguntas están basadas en los cuadros anteriores.

3. Mediante las frecuencias de celda corregidas, calcule el *odds ratio* que corrige la mala clasificación de la enfermedad.
4. ¿Cuál es la dirección del sesgo?

Resumen

- Como en el caso no diferencial, la clasificación errónea diferencial puede corregirse cuando están disponibles los cálculos de sensibilidad y especificidad.
- El procedimiento implica transformar las frecuencias de celda clasificadas erróneamente en frecuencias de celda "corregidas"; además, requiere suponer que la clasificación errónea de la exposición y la enfermedad son independientes.
- El procedimiento requiere entradas de cálculos de sensibilidades y especificidades de la clasificación errónea de la exposición y la enfermedad, así como también frecuencias de celda en el cuadro de 2 x 2 observado que relaciona la exposición y la enfermedad.
- No hay una fórmula algebraica explícita para el procedimiento de corrección, aunque los cálculos se pueden realizar mediante un programa de computación.

Corrección de la clasificación errónea diferencial y los problemas asociados

Las fórmulas generales para corregir la clasificación errónea diferencial son las siguientes:

$A = \{Sp(D|E)[(a + b)Sp(E|D) - b] - [1 - Sp(D|E)][(c + d)Sp(E|D) - d]\} / q_A$

$B = \{Sp(D|no\ E)[(a + b)Se(E|D) - a] - [1 - Sp(D|no\ E)][(c + d)Se(E|D) - c]\} / q_B$

$C = \{Se(D|E)[(c + d)Sp(E|no\ D) - d] - [1 - Se(D|E)][(a + b)Sp(E|no\ D) - b]\} / q_C$

$D = \{Se(D|no\ E)[(c + d)Se(E|no\ D) - c] - [1 - Se(D|no\ E)][(a + b)Se(E|no\ D) - a]\} / q_D$

Corrección de la clasificación errónea diferencial y los problemas asociados (continuación)

donde

$q_A = [Se(D|E) + Sp(D|E) - 1] \times [Se(E|D) + Sp(E|D) - 1]$

$q_B = [Se(D|no\ E) + Sp(D|no\ E) - 1] \times [Se(E|D) + Sp(E|D) - 1]$

$q_C = [Se(D|E) + Sp(D|E) - 1] \times [Se(E|no\ D) + Sp(E|no\ D) - 1]$

$q_D = [Se(D|no\ E) + Sp(D|no\ E) - 1] \times [Se(E|no\ D) + Sp(E|no\ D) - 1]$

$Se(D|E)$ = sensibilidad **D** dado **E**

$Se(D|no\ E)$ = sensibilidad **D** dado no **E**

$Sp(D|E)$ = especificidad **D** dado **E**

$Sp(D|no\ E)$ = especificidad **D** dado no **E**

$Se(E|D)$ = sensibilidad **E** dado **D**

$Se(E|no\ D)$ = sensibilidad **E** dado no **D**

$Sp(E|D)$ = especificidad **E** dado **D**

$Sp(E|no\ D)$ = especificidad **E** dado no **D**

Estas fórmulas pueden dar resultados inadecuados para las frecuencias de celda corregidas **A**, **B**, **C** y **D** en las siguientes circunstancias:

1. $q_A = 0$, $q_B = 0$, $q_C = 0$ ó $q_D = 0$
2. El valor corregido para **A** es negativo.
3. El valor corregido para **B** es negativo.
4. El valor corregido para **C** es negativo.
5. El valor corregido para **D** es negativo.

Auto-evaluación (P9.15)

Utilice los datos del siguiente cuadro para responder las preguntas que se consignan a continuación. Suponga que hay clasificación errónea diferencial de exposición, Se(E|D) = 0,9; Se(E|no D) = 0,6; Sp(E|D) = 0,7; Sp(E|no D) = 0,9 y que no hay clasificación errónea de enfermedad.

1. Calcule el *odds ratio* que corrige la clasificación errónea o mala clasificación **???**
2. El sesgo de clasificación errónea **???** valor nulo.

Opciones
3,5 4,0 5,1 5,8 se aleja del es hacia el

	E	No E	Total
D	660	340	1000
No D	250	750	1000
Total	910	1090	2000

La prueba de diagnóstico y su relación con la clasificación errónea

Ejemplo de diagnóstico clínico

En una clínica médica se examina a un paciente de 43 años cuya pierna presenta una hinchazón aguda. Al médico le preocupa que el paciente pueda tener trombosis venosa profunda (TVP), un trombo o coágulo de sangre en el sistema venoso profundo de la pierna. ¿Cómo diagnostica el médico a este paciente? Clásicamente, la prueba de diagnóstico de referencia, o **estándar de oro,** para la TVP ha sido una flebografía o venografía ascendente con contraste. Sin embargo, el contraste radiográfico en sí mismo conlleva riesgos y puede inducir flebotrombosis local. Además, reacciones alérgicas por el medio del contraste yodado son comunes y pueden impedir que el paciente sea sometido a dicho procedimiento. Asimismo, una flebografía es invasiva e incómoda, por lo cual se preferiría una prueba menos dolorosa en el sitio de punción, si estuviese disponible y tuviese suficiente precisión. La verdadera condición de TVP del paciente podría aún evaluarse mediante un seguimiento prolongado, ya que la TVP debería declararse al causar síntomas persistentes. Sin embargo, el médico clínico prefiere hacer un diagnóstico rápido sin seguimiento prolongado, no sólo para evitar los aspectos negativos de realizar una flebografía sino, principalmente, debido al riesgo inmediato de embolia pulmonar (cuando un coágulo en una

vena profunda de la pierna migra dirigiéndose a los pulmones).

Una prueba de diagnóstico alternativa para la TVP es la ecografía, que puede programarse rápidamente y no es invasiva para el paciente. Sin embargo, la ecografía puede ser menos precisa para diagnosticar la TVP que la flebografía. Por lo tanto, ¿cómo elegimos entre el procedimiento considerado el de referencia o estándar de oro con sus complicaciones y el procedimiento más simple con el potencial de clasificación errónea? Los estudios de pruebas de diagnóstico pueden brindar algunas respuestas. Dichos estudios se describen en la próxima actividad.

Resumen

- El procedimiento de examen considerado estándar de oro para el diagnóstico de una condición de enfermedad específica puede presentar algunos problemas (por ejemplo, ser invasivo, riesgoso, costoso) asociados con su uso.
- Un ejemplo de dicho procedimiento es el uso de un venograma con contraste para diagnosticar la trombosis venosa profunda (TVP).
- Una prueba de diagnóstico alternativa para la TVP es la ecografía, que puede no ser tan precisa como un venograma.
- Un estudio de prueba de diagnóstico puede ayudar a un médico clínico a evaluar el desempeño de un procedimiento de diagnóstico alternativo en comparación con el uso del procedimiento considerado estándar de oro.

Estudios de pruebas de diagnóstico

En la medicina clínica, los estudios referidos a la clasificación errónea de la enfermedad generalmente se llaman **estudios de pruebas de diagnóstico.** El objetivo principal de dicho estudio es evaluar el desempeño de un examen para diagnosticar una condición de enfermedad de interés. Suponga, por ejemplo, que la condición de enfermedad es la trombosis venosa profunda, o TVP. En un estudio de diagnóstico de la TVP, el médico clínico sólo se enfoca en pacientes con un síntoma específico, por ejemplo, "hinchazón aguda de la pierna" o trombosis local; luego, a estos pacientes se les realiza tanto la prueba de diagnóstico (normalmente una ecografía) como el procedimiento considerado estándar de oro (en general, una flebografía). Aquí están los resultados de dicha prueba de diagnóstico, presentados como un cuadro de clasificación errónea:

Cuadro de clasificación errónea - Pruebas de diagnóstico

		Estándar de Oro		
		+	-	Total
Resultados de la Prueba	+	48	14	62
	-	12	126	138
	Total	60	140	200

Mediante la prueba de diagnóstico, la condición de clasificación de la enfermedad que se determina para un paciente dado se llama **resultado de la prueba**, que aparece como positivo (+) o negativo (-) en las filas del cuadro. El procedimiento utilizado para definir la verdadera enfermedad se denomina **estándar de oro**, aunque puede ser imperfecto. En el cuadro de clasificación errónea, los resultados generados al utilizar el estándar de oro aparecen en las columnas del cuadro. Normalmente, el estándar de oro es una prueba más detallada, cara o riesgosa que la prueba de diagnóstico utilizada por el médico. El estándar de oro podría incluso requerir un seguimiento prolongado del paciente si se espera que la enfermedad finalmente se declare, un examen *post mortem* o una medida que combine más de una estrategia, a veces, con formas complicadas pensadas exclusivamente para la enfermedad específica.

Mediante la información del cuadro de clasificación errónea, el desempeño de la prueba de diagnóstico

puede evaluarse empleando varias medidas importantes, lo que incluye **sensibilidad**, **especificidad** y **prevalencia**. Recuerde que la **sensibilidad** describe el desempeño de la prueba en los pacientes que verdaderamente padecen la enfermedad, y se define como la probabilidad condicional de un resultado positivo en la prueba dada la verdadera enfermedad, P(Prueba + | D+), donde "D+" indica estatus positivo de enfermedad.

Preguntas de estudio (P9.16)

1. ¿Cuál es la sensibilidad de la prueba en el cuadro anterior?
2. Si la sensibilidad hubiese sido 0,99, ¿qué conclusión puede sacar sobre un paciente verdaderamente enfermo que tuvo un resultado negativo en la prueba?

La **especificidad** describe el desempeño de las pruebas entre los pacientes que verdaderamente no padecen la enfermedad. Se define como la probabilidad condicional de un resultado negativo en la prueba dada la ausencia de enfermedad, P(Prueba - | D-), donde "D-" indica estatus negativo de enfermedad.

Preguntas de estudio (P9.16) (continuación)

3. ¿Cuál es la especificidad de la prueba?
4. Si la sensibilidad hubiese sido 0,99, ¿qué conclusión puede sacar sobre un paciente verdaderamente sano que tuvo un resultado negativo en la prueba?

La **prevalencia** se calcula como la proporción de pacientes en la muestra del estudio que verdaderamente padecen la enfermedad, P(D+). Si se sabe poco sobre un paciente, la prevalencia de la enfermedad en un estudio de prueba de diagnóstico es el mejor cálculo de probabilidad previa a la prueba de que el paciente padezca la enfermedad.

Preguntas de estudio (P9.16) (continuación)

5. ¿Cuál es la prevalencia de la verdadera enfermedad a partir de estos datos?
6. Basándose en los cálculos de sensibilidad, especificidad y prevalencia anteriores, ¿cree que la prueba es una buena herramienta de diagnóstico para la TVP? Desarrolle su respuesta brevemente.

Aunque las tres medidas (sensibilidad, especificidad y prevalencia) brindan importante información de resumen sobre el desempeño de la prueba de diagnóstico, la medida más útil del desempeño general se llama **valor predictivo**. Éste se describe en la próxima actividad.

Resumen

- En la medicina clínica, los estudios referidos a la clasificación errónea de la enfermedad generalmente se llaman **estudios de pruebas de diagnóstico**.
- El propósito de un estudio de prueba de diagnóstico es evaluar el desempeño de la prueba, en lugar de hacer el ajuste del sesgo de información.
- El procedimiento utilizado para definir la verdadera enfermedad se denomina **estándar de oro**.

- En un estudio de diagnóstico, el médico clínico se enfoca en pacientes con un síntoma específico y luego realiza tanto la prueba de diagnóstico como el procedimiento considerado el estándar de oro en estos pacientes.
- El desempeño de una prueba de diagnóstico puede evaluarse mediante varias medidas importantes, lo que incluye **sensibilidad**, **especificidad** y **prevalencia**.
- Una medida más útil del desempeño de la prueba de diagnóstico la brinda el **valor predictivo**.

Pruebas de detección

Un segundo tipo de estudio clínico relacionado con la clasificación errónea se denomina **prueba de detección**. A diferencia de la prueba de diagnóstico, la prueba de detección se enfoca en una población amplia de individuos asintomáticos para identificar a quiénes pueden requerir una evaluación de diagnóstico más detallada. Los individuos en una prueba de detección no han acudido a un médico por una molestia específica.

Normalmente, se invita a los miembros del público en general a someterse a pruebas de detección de varios tipos para separarlos en aquellos con mayores y menores probabilidades de padecer la enfermedad. Luego, se insta a aquellos con mayores probabilidades a solicitar atención médica para un diagnóstico definitivo. Los que poseen menores probabilidades no reciben ningún beneficio directo de salud porque no tienen la condición de enfermedad , objeto del programa de detección sistemática o tamizaje . Asimismo, según la invasividad de la prueba de detección y/o la condición de enfermedad bajo análisis, los participantes del programa de detección sistemática pueden correr riesgos, así como también tener inconvenientes, ansiedad, algún costo personal y, a veces, molestias como, por ejemplo, las que se presentan a raíz del uso de una colonoscopía para detectar cáncer intestinal.

El valor predictivo de la prueba de diagnóstico

La probabilidad de una verdadera condición de enfermedad para un paciente individual, dado el resultado de la prueba de diagnóstico, se llama **valor predictivo** de la prueba. Este valor predictivo es especialmente útil para fines del diagnóstico médico del paciente, porque calcula directamente la probabilidad de que el paciente verdaderamente padezca o no la enfermedad, según los resultados de la prueba de diagnóstico.

Preguntas de estudio (P9.17)

Suponga que T+ denota que un paciente verdaderamente tiene una condición de enfermedad de interés, mientras que D+ denota el evento de un resultado positivo de la prueba de diagnóstico en el mismo paciente.

1. ¿Cuál de las dos siguientes afirmaciones de probabilidad describe la sensibilidad y cuál describe el valor predictivo?
   A. P(T+|D+) B. P(D+|T+)

El valor predictivo puede obtenerse directamente del cuadro de clasificación errónea generado por un estudio de prueba de diagnóstico. Debido a que hay dos resultados posibles para una prueba, hay dos valores predictivos diferentes. La probabilidad de padecer verdaderamente la enfermedad cuando la prueba es positiva se llama **valor predictivo positivo** y se denota como **VP+**. La probabilidad de no sufrir verdaderamente la enfermedad si la prueba es negativa se llama **valor predictivo negativo** y se denota como **VP-**. Tanto **VP+** como **VP-** son proporciones. Cuanto más cercanas a 1 sean estas proporciones, mejor será el desempeño predictivo de la prueba.

Cuadro de clasificación errónea: estudio prueba de diagnóstico

		Padecer verdaderamente la enfermedad (T) +	-	
Resultado de prueba de diagnóstico (D)	+	48	14	62
	-	12	126	138
		60	140	200

VP+: $P(T+ \mid D+) =$ Valor predictivo positivo

VP-: $P(T- \mid D-) =$ Valor predictivo negativo

El valor predictivo positivo se calcula como el número de resultados positivos verdaderos, dividido por todos los resultados positivos.

$$\mathbf{VP+} = \frac{\text{\# resultados positivos verdaderos}}{\text{todos los resultados positivos}}$$

Preguntas de estudio (P9.17) (continuación)

2. ¿Cuál es el VP+ para el cuadro anterior?

El valor predictivo positivo generalmente se denomina **probabilidad posprueba** de padecer la enfermedad. Es distinta de la prevalencia, que brinda la **probabilidad** promedio **previa a la prueba** que tiene el paciente de padecer la enfermedad.

Preguntas de estudio (P9.17) (continuación)

3. Sobre la base de los datos del cuadro de clasificación errónea, ¿cuál es el valor de la probabilidad promedio que tiene el paciente de padecer TVP antes de realizarse una ecografía?
4. ¿El uso de la ecografía mejoró el diagnóstico de la enfermedad para las personas con resultados positivos en la ecografía? Desarrolle su respuesta brevemente.

El **valor predictivo negativo** es el número de resultados negativos verdaderos, dividido por el número total de individuos con resultados negativos en las pruebas.

$$\mathbf{VP-} = \frac{\text{\# resultados negativos verdaderos}}{\text{\# de individuos con resultados negativos}}$$

Preguntas de estudio (P9.17) (continuación)

5. ¿Cuál es el VP- para estos datos?
6. Sobre la base de los datos del cuadro de clasificación errónea, ¿cuál es el valor de la probabilidad promedio que tiene el paciente de no padecer TVP antes de realizarse una ecografía?
7. ¿El uso de la ecografía mejoró el diagnóstico de la enfermedad para las personas con resultados negativos en la ecografía? Desarrolle su respuesta brevemente.

La prevalencia de la enfermedad verdadera en un estudio de prueba de diagnóstico puede influir mucho en el tamaño de los valores predictivos obtenidos. Para ejemplificar esto, consideremos ahora un segundo cuadro de clasificación errónea para un grupo diferente de pacientes que se han presentado, con dolor pero sin hinchazón en la pierna, ante su médico clínico.

Cuadro de clasificación errónea: estudio prueba de diagnóstico

		Venograma de TVP +	Venograma de TVP -	
Ecografía	+	16	18	34
	-	4	162	166
		20	180	200

Preguntas de estudio (P9.17) (continuación)

8. ¿Cuál es la sensibilidad, especificidad y prevalencia en el cuadro?
9. En el cuadro de clasificación errónea considerado anteriormente, la sensibilidad, especificidad y prevalencia eran iguales a 0,80; 0,90 y 0,30 respectivamente. ¿Cómo se comparan estos valores con aquellos que se calcularon en la pregunta anterior?
10. ¿Cuáles son los valores de VP+ y VP- en el cuadro anterior?
11. En el cuadro de clasificación errónea considerado anteriormente, VP+ y VP- eran iguales a 0,77 y 0,91, respectivamente, y la prevalencia fue 0,30. ¿Cómo se comparan estos valores con los valores predichos que se calcularon en la pregunta anterior?
12. ¿Cuál es la lección de esta historia que relaciona el valor predictivo con la prevalencia de la enfermedad?

Resumen

- El valor predictivo (o probabilidad posterior a la prueba) es la probabilidad de la verdadera condición de enfermedad dado el resultado de la prueba de diagnóstico.
- El valor predictivo puede obtenerse directamente del cuadro de clasificación errónea generado por un estudio de prueba de diagnóstico.
- La probabilidad de enfermedad, cuando la prueba da positivo, se llama valor predictivo positivo y se denota como VP+.
- La probabilidad de enfermedad, cuando la prueba da negativo, se llama valor predictivo negativo y se denota como VP-.
- VP+ = número de positivos verdaderos / todos los positivos
- VP- = número de negativos verdaderos / todos los negativos
- Cuanto más cercanos VP+ y VP- sean a 1, mejor será la prueba.

- La prevalencia de la verdadera enfermedad en un estudio de prueba de diagnóstico puede influir mucho en el tamaño del valor predictivo.

Expresión del valor predictivo en términos de sensibilidad, especificidad y prevalencia

La prevalencia de la verdadera enfermedad en un estudio de prueba de diagnóstico puede influir mucho sobre el tamaño del valor predictivo. Primero ilustramos esta afirmación mediante un ejemplo. Luego describimos matemáticamente cómo calculamos los valores numéricos en nuestro ejemplo mediante fórmulas que expresan los valores predictivos positivos y negativos, en términos de los parámetros de sensibilidad, especificidad y prevalencia.

Suponga que fijamos los valores de sensibilidad y especificidad en 0,80 y 0,90; respectivamente, y consideramos cuáles serían los valores predictivos para diferentes prevalencias, por ejemplo, 0,10; 0,30; 0,50; 0,70 y 0,90. Se obtendrán los siguientes resultados:.

Prevalencia de verdadera enfermedad	Sensibilidad D	Especificidad D	VP+	VP-
0,01	0,80	0,90	0,08	0,998
0,10	0,80	0,90	0,47	0,98
0,30	0,80	0,90	0,77	0,91
0,50	0,80	0,90	0,89	0,82
0,70	0,80	0,90	0,95	0,66
0,90	0,80	0,90	0,99	0,33

Expresión del valor predictivo en términos de sensibilidad, especificidad y prevalencia (*continuación*)

Del cuadro anterior, podemos ver que para los valores fijos de sensibilidad y especificidad, el valor predicho positivo (**VP +**) aumenta de 0,08 a 0,99 a medida que la prevalencia aumenta de 0,01 a 0,90; asimismo, aunque tanto las sensibilidades como las especificidades son relativamente altas, el valor predictivo positivo puede ser muy bajo (por ejemplo, 8%) si la prevalencia de la enfermedad es baja (por ejemplo, 1%). Una tendencia inversa se observa para el valor predicho negativo (**VP-**), que disminuye a medida que aumenta la prevalencia de la enfermedad.

Ahora describimos cómo las fórmulas para **VP+** y **VP-** pueden expresarse en términos de los parámetros de sensibilidad, especificidad y prevalencia. Para hacerlo, necesitamos utilizar un teorema famoso sobre probabilidades condicionales, llamado **teorema de Bayes**, que nos permite expresar probabilidades condicionales de la forma:

P(A|B) y P(A|no B)

en términos de probabilidades condicionales de la forma:

P(B|A) y P(B|no A)

en donde **A** y **B** han sido cambiadas al lado opuesto del signo dado. En los estudios de pruebas de diagnóstico, el evento **A** se refiere a la condición positiva de verdadera enfermedad de un individuo, que denotamos:

A = T+

El evento **B** se refiere a los resultados positivos de la prueba de diagnóstico de un individuo, que denotamos:

B = D+

Por lo consiguiente

no A = T- y
no B = D-

denotan ser un verdadero negativo y tener un resultado negativo en la prueba de diagnóstico, respectivamente. Por lo tanto:

P(A|B) = P(T+|D+)

define el **valor predictivo positivo** (**VP+**) de la prueba de diagnóstico,

P(no A|no B) = P(T -|D-)

define el **valor predictivo negativo** (**VP-**) de la prueba de diagnóstico,

P(B|A) = P(D+|T+)

denota la **sensibilidad** de la prueba de diagnóstico,

(no B|no A) = P(D-|T-)

denota la **especificidad** de la prueba de diagnóstico, y

P(A) = P(T+)

denota la **prevalencia** de la verdadera enfermedad en la población bajo estudio.

Mediante el teorema de Bayes, podemos expresar **P(A|B)** de la siguiente forma:

Expresión del valor predictivo en términos de sensibilidad, especificidad y prevalencia (continuación)

$$P(A \mid B) = \frac{P(A)P(B \mid A)}{P(A)P(B \mid A) + [1 - P(A)]P(B \mid \text{no } A)}$$

Sustituyendo **A**, **no A**, **B** en esta fórmula por **T+**, **T-**, **D+** , respectivamente, obtenemos el siguiente resultado:

$$P(T+ \mid D+) = \frac{P(T+)P(D+ \mid T+)}{P(T+)P(D+ \mid T+) + [1 - P(T+)]P(D+ \mid T-)}$$

que puede reescribirse como:

$$VP+ = \frac{PrevD \square SensD}{[PrevD \square SensD] + [(1 - PrevD) \square (1 - EspecD)]}$$

.

En forma similar, podemos utilizar el **teorema de Bayes** para describir **VP-** en términos de sensibilidad, especificidad y prevalencia de la enfermedad:

$$VP- = \frac{(1 - PrevD) \square EspecD}{[PrevD \square (1 - SensD)] + [(1 - PrevD) \square EspecD]}$$

Auto-evaluación (P9.18)

Para el cuadro de clasificación que se muestra a la derecha, determine lo siguiente:

		Verdaderamente		
		D	no D	
Clasificado	D	90	20	110
	no D	10	180	190
		100	200	300

1. ¿Cuál es la sensibilidad? . . **???**
2. ¿Cuál es la especificidad? . . **???**
3. ¿Cuál es la prevalencia de la enfermedad? **???**
4. ¿Cuál es el valor predictivo positivo? **???**

Opciones
10% **33,3%** **36,7%** **81,8%** **90%**

La sensibilidad y la especificidad para el cuadro de clasificación que se muestra a la derecha siguen siendo de 90% como en las preguntas anteriores. Para este cuadro, responda lo siguiente:

		Verdaderamente		
		D	no D	
Clasificado	D	90	50	140
	no D	10	450	460
		100	900	600

5. ¿Cuál es la prevalencia de la enfermedad? **???**
6. ¿Cuál es el valor predictivo positivo? **???**

7. La prevalencia en este cuadro es menor que en el cuadro anterior; por lo tanto, el valor predictivo positivo es **???** que en el cuadro anterior.

Opciones
16,7% **64,3%** **90%** **mayor** **menor**

Nuevamente, la sensibilidad y la especificidad para el cuadro de clasificación que se muestra a la derecha son de 90%. Para este cuadro, responda lo siguiente:

		Verdaderamente		
		D	no D	
Clasificado	D	90	90	180
	no D	10	810	820
		100	900	1000

8. ¿Cuál es la prevalencia de la enfermedad? **???**
9. ¿Cuál es el valor predictivo positivo? **???**
10. La prevalencia en este cuadro es menor que en los dos cuadros anteriores; por lo tanto, el valor predictivo positivo es **???** que en los dos cuadros anteriores.
11. Estos resultados ilustran el hecho de que si la prevalencia es baja, el valor predictivo puede ser bastante **???** incluso si los parámetros de sensibilidad y especificidad son bastante **???**

Opciones

10% **36,7%** **50%** **alto/s** **mayor/es** **bajo/s** **menor/es**

Nomenclatura

Cuadros de clasificación errónea para la enfermedad y la exposición

Cuadro de clasificación errónea para la enfermedad:

		Verdaderamente		
		D	no D	
Clasificados	D´			
	no D´			

Cuadro de clasificación errónea para la exposición:

		Verdaderamente		
		E	no E	
Clasificados	E´			
	no E´			

Cuadros de datos observados (clasificados erróneamente) y corregidos

Datos observados (es decir, clasificados erróneamente)

	E´	no E´	
D´	a	b	
no D´	c	d	

Datos corregidos (es decir, ajustados)

	E	no E	
D	A	B	
no D	C	D	

Fórmulas para corregir la clasificación errónea no diferencial de la enfermedad (sin clasificación errónea de la exposición)

$A = [a\ SpD – c(1 – SpD)] / q$
$B = [b\ SpD – d(1 – SpD)] / q$
$C = [c\ SeD – a(1 – SeD)] / q$
$D = [d\ SeD – b(1 – SeD)] / q$

Donde $q = SeD + SpD - 1$

Fórmulas para corregir la clasificación errónea no diferencial de la exposición (sin clasificación errónea de la enfermedad)

$A = [a\ SpE – b(1 – SpE)] / q$
$B = [b\ SeE – a(1 – SeE)] / q$
$C = [c\ SpE – d(1 – SpE)] / q$
$D = [d\ SeE – c(1 – SeE)] / q$

Donde q = SeE + SpE - 1

Fórmulas para corregir la clasificación errónea no diferencial de la exposición y la enfermedad

A = {SpD[(a + b) SpE – b] – (1 – SpD)[(c + d) SpE – d]} / q*
B = {SpD[(a + b) SeE – a] – (1 – SpD)[(c + d) SeE – c]} / q*
C = {SeD[(c + d) SpE – d] – (1 – SeD)[(a + b) SpE – b]} / q*
D = {SeD[(c + d) SeE – c] – (1 – SeD)[(a + b) SeE – a]} / q*

Donde q* = (SeD + SpD – 1) (SeE + SpE – 1)

D	Verdaderamente padece la enfermedad
D´	Clasificado como que padece la enfermedad
E	Verdaderamente expuesto
E´	Clasificado como expuesto
no D	Verdaderamente no padece la enfermedad
no D´	Clasificado como que no padece la enfermedad
no E	Verdaderamente no expuesto
no E´	Clasificado como no expuesto
$\hat{OR}$	*Odds ratio* basado en datos *observados* = ad/bc
$\hat{OR}_{ajustado}$	*Odds ratio* basado en datos *corregidos* o *ajustados* = AD/BC
$\hat{RR}$	Razón de riesgos basada en datos *observados* = [a/(a+c)]/[b/(b+d)]
$\hat{RR}_{ajustada}$	Razón de riesgos basada en los datos *corregidos* o *ajustados* = [A/(A+C)]/[B/(B+D)]
Sensibilidad	De quienes verdaderamente poseen la característica, la proporción que se clasificará correctamente como que posee la característica; para la enfermedad, Pr(D´|D+); para la exposición, Pr(E´|E+).
SeD	Sensibilidad de la clasificación errónea de la enfermedad
SeE	Sensibilidad de la clasificación errónea de la exposición
SpD	Especificidad de la clasificación errónea de la enfermedad
SpE	Especificidad de la clasificación errónea de la exposición
Especificidad	De quienes verdaderamente no poseen la característica, la proporción que se clasificará correctamente como que no posee la característica; para la enfermedad, P(no D´|no D); para la exposición, P(no E´|no E).

Referencias sobre sesgo de información/clasificación errónea

Descripciones generales

Greenberg RS, Daniels SR, Flanders WD, Eley JW, Boring JR. *Medical Epidemiology* (3era. Ed). Lange Medical Books, Nueva York, 2001.

Hill H, Kleinbaum DG. *Bias in Observational Studies*. En Encyclopedia of Biostatistics, páginas 323-329, Oxford University Press, 1999.

Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic Research: Principles and Quantitative Methods*. John Wiley and Sons Publishers, Nueva York, 1982.

Temas especiales

Barron, BA. *The effects of misclassification on the estimation of relative risk*. Biometrics 1977; 33(2):414-8.

Copeland KT, Checkoway H, McMichael AJ and Holbrook RH. *Bias due to misclassification in the estimation of*

relative risk. Am J Epidemiol 1977; 105(5):488-95.
Dosemeci M, Wacholder S, and Lubin JH. *Does nondifferential misclassification of exposure always bias a true effect toward the null value?* Am J Epidemiol 1990:132(4):746-8.
Espeland MA, Hui SL. *A general approach to analyzing epidemiologic data that contain misclassification errors*. Biometrics 1987; 43(4):1001-12.
Greenland S. *The effect of misclassification in the presence of covariates*. Am J Epidemiol 1980; 112(4):554-69.
Greenland S, Kleinbaum DG. *Correcting for misclassification in two-way tables and matched-pair studies*. Int J Epidemiol 1983; 12(1):93-7.
Reade-Christopher SJ, Kupper LL. *Effects of exposure misclassification on regression analyses of epidemiologic follow-up study data*. Biometrics 1991; 47(2):535-48.
Satten GA and Kupper LL. *Inferences about exposure-disease associations using probability-of-exposure information*. J Am Stat Assoc 1993; 88:200-8.
Wynder EL. *Investigator bias and interviewer bias: the problem of reporting systematic error in epidemiology*. J Clin Epidemiol 1994; 47(8):825-7.

Ejercicios

EPI-1. Exposición a la radiación vs. tumores gastrointestinales

Un investigador tiene interés en estudiar la relación entre la exposición a la radiación y el desarrollo de tumores gastrointestinales. Reúne una cohorte de individuos que no padecen cáncer, los divide en grupos de exposición a la radiación "alta" y "baja", y los sigue durante 10 años en busca de pruebas de la aparición de tumores gastrointestinales.

a. Suponga que el investigador utiliza una prueba de diagnóstico que no logra identificar al 20% de <u>todos</u> los tumores gastrointestinales, pero nunca registra un falso resultado positivo (es decir, no se ha diagnosticado erróneamente como que sufre de cáncer a nadie que no lo padece). Describa esta situación en términos de sensibilidades y especificidades para los individuos expuestos y no expuestos:

   Sensibilidad (D|E) = Sensibilidad (D|no E) =

   Especificidad (D|E) = Especificidad (D|no E) =

   A continuación se resumen los datos de la población de ORIGEN (sin clasificación errónea):

	Alta radiación	**Baja radiación**	**Total**
Tumor gastrointestinal	600	200	800
Sin tumor	1400	1300	2700
Total	2000	1500	3500

b. ¿Cuál es el cálculo no sesgado de la razón de riesgos (RR)?

c. Utilice las sensibilidades y las especificidades del ítem a) para mostrar los datos que habría observado el investigador:

	Alta radiación	**Baja radiación**	**Total**
Tumor gastrointestinal			
Sin tumor			
Total			

d. Calcule la RR observada. ¿Hay ausencia de sesgo, sesgo hacia el valor nulo, sesgo que se aleja del valor nulo o sesgo de cambio?

e. El siguiente cuadro resume los datos OBSERVADOS de un estudio <u>diferente</u> sobre radiación y tumores gastrointestinales:

	Alta radiación	**Baja radiación**	**Total**
Tumor gastrointestinal	476	376	852
Sin tumor	524	624	1148
Total	1000	1000	2000

Suponiendo el escenario descrito en el ítem "a" anterior, ¿cuál es el cálculo correcto (no sesgado) de la RR correspondiente a la relación entre la exposición a la radiación y los tumores gastrointestinales para este estudio? Muestre sus cálculos. ¿Cuál es la naturaleza del sesgo, si lo hay?

f. Calcule el *odds ratio* (en lugar de la RR) tanto para los datos observados como para los datos corregidos en el ítem e). ¿Su conclusión sobre el sesgo es la misma?

EPI-2. EC y comportamiento

Los siguientes datos representan la población de ORIGEN en un estudio de casos y controles que analiza la enfermedad coronaria (EC) y comportamiento de tipo A.

	Tipo A	**No tipo A**	**Total**
EC	80	25	105
no EC	50	55	105
Total	130	80	210

a. Calcule el valor no sesgado del *odds ratio* de exposición (EOR, por sus siglas en inglés).
b. Suponga que la exposición se clasificó erróneamente con las siguientes sensibilidades y especificidades:

Casos:	sensibilidad = 0,8	especificidad = 1
Controles:	sensibilidad = 0,9	especificidad = 0,6

Indique los datos que se habrían observado, dada la clasificación errónea:

	Tipo A	**No tipo A**	**Total**
EC			
no EC			
Total			

c. Calcule el EOR a partir de los datos observados. ¿Hay ausencia de sesgo, sesgo hacia el valor nulo, sesgo que se aleja del valor nulo o sesgo de cambio?

EPI-3. Alteración del sueño vs. depresión clínica

El estudio sobre precursores de Johns Hopkins, un estudio de cohorte prospectivo a largo plazo, se utilizó para evaluar la relación entre la alteración del sueño informada por los individuos y la consiguiente depresión clínica. Un total de 1053 hombres en varias universidades de Estados Unidos brindaron información sobre sus hábitos de sueño durante los años lectivos de la carrera de Medicina; luego se les hizo un seguimiento durante 20 años para investigar el desarrollo de depresión. Los individuos se sometieron a pruebas psicológicas exhaustivas y largas entrevistas estructuradas a cargo de psiquiatras capacitados en el diagnóstico de la depresión. Los resultados del estudio se resumen a continuación (puede suponer que estos datos están clasificados correctamente):

	Alteración del sueño	
	Sí	**No**
Depresión	168	93
Sin depresión	258	534
	426	627

Ahora suponga que los investigadores tenían fondos limitados y no pudieron utilizar las herramientas de diagnóstico sofisticadas descritas anteriormente. Tenían dos opciones disponibles para el diagnóstico de la depresión entre los individuos del estudio: (1) una entrevista o (2) un cuestionario auto-administrado.

La entrevista pudo clasificar la condición de depresión con las siguientes sensibilidades y especificidades:

Expuestos:	sensibilidad = 0,84 y especificidad = 0,90
No expuestos:	sensibilidad = 0,84 y especificidad = 0,90

El cuestionario pudo clasificar la condición de depresión con las siguientes sensibilidades y especificidades:

Expuestos:	sensibilidad = 0,80 y especificidad = 0,88
No expuestos:	sensibilidad = 0,90 y especificidad = 0,95

a. ¿Cuál es la verdadera razón de riesgos (RR) que describe la asociación entre la alteración del sueño y la depresión?
b. ¿Cuál es la RR observada cuando se utiliza la entrevista para clasificar la depresión?
c. ¿Cuál es la RR observada cuando se utiliza el cuestionario para clasificar la depresión?

d. Considerando sólo los problemas de validez, ¿qué herramienta se prefiere para diagnosticar la depresión: la entrevista o el cuestionario? Justifique su respuesta.

EPI-4. Uso de aspirina y hemorragia gastrointestinal

Un epidemiólogo tenía interés en determinar si el uso de un nuevo analgésico que contenía aspirina estaba asociado a un mayor riesgo de hemorragia gastrointestinal. Identificó a 600 pacientes que tomaban la droga en forma regular y 600 individuos no expuestos. Se siguió a los individuos durante un año para detectar la presencia de hemorragia gastrointestinal. Debido a la publicidad sobre los peligros potenciales de la nueva droga, los médicos que participaron en el estudio siguieron a los individuos expuestos más de cerca que a los individuos no expuestos y, por lo tanto, tenían una mayor probabilidad de diagnosticar hemorragias gastrointestinales cuando ocurrían en este grupo de individuos del estudio.

a. ¿Cuál o cuáles de las siguientes afirmaciones describen la situación? [Puede elegir más de una opción].

   i. Clasificación errónea no diferencial de la exposición
   ii. Clasificación errónea diferencial de la exposición
   iii. Clasificación errónea no diferencial de la enfermedad
   iv. Clasificación errónea diferencial de la enfermedad
   v. Sesgo de recuerdo
   vi. Sesgo de detección
   vii. Sesgo de Berkson

Suponga que el siguiente cuadro representa a la población de ORIGEN (es decir, correctamente clasificada) para el estudio sobre la nueva droga y la hemorragia gastrointestinal.

	Uso de la nueva droga	
	Sí	**No**
Hemorragia gastrointestinal	400	300
Sin hemorragia gastrointestinal	200	300

b. Suponga que cuando se llevó a cabo el estudio, las hemorragias gastrointestinales se clasificaron con sensibilidad = 0,9 y especificidad = 0,75 para los individuos que estaban utilizando la droga. Para los individuos que no estaban utilizando la droga, las hemorragias gastrointestinales se clasificaron con sensibilidad = 0,6 y especificidad = 0,75. Utilice esta información para completar el siguiente cuadro.

	Uso de la nueva droga	
	Sí	**No**
Hemorragia gastrointestinal		
Sin hemorragia gastrointestinal		

c. Calcule la medida de asociación adecuada en forma de razón para los datos OBSERVADOS (clasificados erróneamente). Indique si hay sesgo y, de ser así, cuál es su dirección.
d. Suponga que el estudio se repitiera, con hemorragias gastrointestinales detectadas por los médicos que no conocían la condición de exposición de los individuos. Si la sensibilidad para los individuos no expuestos aumentase a 0,9, mientras que todas las otras sensibilidades y especificidades continuaran igual, ¿qué impacto tendría ésto sobre el sesgo?

i. El sesgo se eliminaría.
ii. La magnitud del sesgo disminuiría, pero la dirección continuaría siendo igual.
iii. La magnitud del sesgo aumentaría, pero la dirección continuaría siendo igual.
iv. La dirección del sesgo cambiaría.
v. Ninguna de las anteriores.

EPI-5. Medicación antidepresiva y riesgo de cáncer de mama

Un estudio titulado *Antidepressant Medication and Breast Cancer Risk* (Medicación antidepresiva y riesgo de cáncer de mama) fue publicado en un número reciente del *American Journal of Epidemiology*. Según la sección del artículo dedicada a los métodos, "Los casos se componían de una muestra aleatoria estratificada por edad (< 50 y > = 50 años) de mujeres de 25 a 74 años a las que se les diagnosticó cáncer primario de mama durante 1995 y 1996 (informe de patología confirmado) y que se encontraban inscritas en el Registro de cáncer de Ontario basado en la población. Como la supervivencia por 1 año al cáncer de mama es de 90%, no se utilizaron encuestados sustitutos. Se tomó una muestra aleatoria de controles de la población de 25 a 74 años de las listas de evaluación de propiedades del Ministerio de Finanzas de Ontario; esta base de datos incluye todos los propietarios de viviendas y a los inquilinos, con indicación de edad, sexo y domicilio".

a. Analice el abordaje de los autores con respecto a la identificación de los casos en términos del potencial sesgo de clasificación errónea.
b. Analice el enfoque de los autores con respecto a la identificación de los controles en términos del potencial de: (a) sesgo de selección y (b) sesgo de clasificación errónea.
c. Analice las ventajas y desventajas de la decisión de no permitir encuestados sustitutos.

La sección del artículo sobre los métodos continúa diciendo: "Se recabaron datos a través de cuestionarios enviados por correo, contestados en forma personal y estructurados que incluían información sobre: (1) datos sociodemográficos; (2) duración, dosis, momento de administración y tipo de medicamentos antidepresivos utilizados; y (3) factores potenciales de confusión. Se les preguntó a los individuos: ¿Alguna vez tomó antidepresivos durante, al menos, 2 semanas en algún momento de su vida?". (Se brindó una lista de 11 antidepresivos a modo de ejemplo).

d. Analice las ventajas y desventajas del enfoque de los autores con respecto a la evaluación de la exposición. Enumere y describa cómo podría surgir algún sesgo de información a partir de este abordaje.

EPI-6. Estudio de validación

Ha realizado:
i) un estudio de cohorte de una enfermedad y exposición dicotómica; y
ii) un estudio de validación separado para evalua sensibilidad y especificidad de la medida utilizada en el estudio de cohorte, ya que la condición de enfermedad fue clasificada erróneamente en forma diferencial con respecto a la exposición.

Resultados del estudio de validación:

Expuestos:

	Enfermedad verdadera	
Enfermedad medida	+	-
+	95	10
-	5	90

No expuestos:

	Enfermedad verdadera	
Enfermedad medida	+	-
+	97	5
-	3	95

Resultados del estudio de cohorte (datos observados):

	Condición de exposición	
Condición de enfermedad	+	-
+	540	284
-	1460	1716

Utilizando la información anterior, calcule la verdadera razón de riesgos ajustada por la clasificación errónea.

EPI-7. Prueba de diagnóstico: fiebre del baseball

Un investigador de la *Javier Lopez University* desarrolló una nueva prueba para detectar la fiebre del baseball. Dicha prueba se evaluó en una población de 10.000 personas, 21% de las cuales padecían con certeza fiebre del baseball. La cantidad de pruebas negativas fue 8162. Se determinó que el valor predictivo positivo era del 91,4%.

a. Utilice la información brindada anteriormente para completar el siguiente cuadro de 2 x 2.

	Fiebre del baseball	
	presente	ausente
Resultado de la prueba		
Positivo	_____	_____
Negativo	_____	_____

b. ¿Cuál es la sensibilidad y especificidad de la prueba? Brinde una interpretación de la sensibilidad mediante el valor que acaba de calcular.
c. Si aplicara esta prueba a un paciente y el resultado fuese negativo, ¿que le diría a esa persona sobre sus posibilidades de padecer la fiebre del baseball?

EPI-8. Valor predictivo: diabetes

En cierta comunidad, el 8% de todos los adultos de más de 50 años tienen diabetes. Si un servicio de salud en esta comunidad diagnostica correctamente al 95% de todas las personas con diabetes como que padecen la enfermedad y diagnostica incorrectamente al 10% de todas las personas sin diabetes como que sufren la enfermedad, encuentre las probabilidades de que:

a. el servicio de salud diagnostique a un adulto de más de 50 años como que padece diabetes.
b. una persona de más de 50 años de edad, a quien el servicio de salud le haya diagnosticado diabetes, en verdad tenga la enfermedad.

EPI-9. Prueba de diagnóstico: cáncer de próstata

Se evaluó a un grupo de 50.000 hombres de más de 60 años para saber si padecían cáncer de próstata mediante la prueba del antígeno prostático específico (PSA, por sus siglas en inglés) con los siguientes resultados:

	Cáncer de próstata	
Resultado de la prueba del PSA	**Sí**	**No**
Positivo	2.900	20.000
Negativo	100	27.000
Total	3.000	47.000

a. ¿Cuál es la sensibilidad y especificidad de esta prueba?
b. ¿Cuál es el valor predictivo positivo de esta prueba?
c. Si una persona obtiene un resultado negativo en la prueba del PSA, ¿debería preocuparse de tener cáncer de próstata? Desarrolle su respuesta.
d. Si una persona obtiene un resultado positivo en la prueba del PSA, ¿debería preocuparse de tener cáncer de próstata? Desarrolle su respuesta.

e. Suponga que se desarrolló un nuevo examen que tenía la misma sensibilidad que la prueba del PSA, pero tenía una especificidad de 95%. Suponiendo el mismo número de personas con y sin cáncer de próstata (verdadero), describa el cuadro de clasificación errónea que se obtendría mediante la prueba nueva.

	Cáncer de próstata	
Resultado de la prueba nueva	**Sí**	**No**
Positivo		
Negativo		
Total	3.000	47.000

f. ¿Cuál es el valor predictivo positivo de la nueva prueba?
g. Si una persona obtiene un resultado negativo en la prueba nueva, ¿debería preocuparse de tener cáncer de próstata? Desarrolle su respuesta.
h. Si una persona obtiene un resultado positivo en la prueba nueva, ¿debería preocuparse de tener cáncer de próstata? Desarrolle su respuesta.

EPI-10. Prueba de detección del cáncer de próstata

Suponga que se ha estado desarrollando una nueva prueba de detección para el cáncer de próstata y que está casi lista para su lanzamiento en el mercado. Se requiere que los individuos sometidos a la prueba de detección den una pequeña muestra de sangre, que luego se analiza para determinar el nivel de cierto factor (Factor P). Cuanto mayor es el nivel del Factor P, mayor es la probabilidad de padecer cáncer de próstata. Sin embargo, el Factor P puede ser elevado como resultado de otras condiciones no cancerosas que involucran la próstata. A continuación, se presentan situaciones hipotéticas y preguntas para discusión.

a. Los científicos en Europa y Estados Unidos no han llegado a un acuerdo sobre el punto de corte adecuado para determinar si la prueba de detección se considera positiva. Discuta las implicaciones de la elección del punto de corte para esta prueba. Piense en términos de evaluación del desempeño de la prueba, así como también en las ramificaciones para el cuidado del paciente.
b. Diseñe el estudio epidemiológico que sería apropiado para evaluar la efectividad de la nueva prueba de detección. Asegúrese de hacer un comentario sobre el diseño del estudio, los individuos del estudio, la(s) exposición(es) de interés, el plan analítico, los sesgos potenciales y cualquier otro aspecto importante de su estudio.
c. Tan pronto la nueva prueba de detección está disponible para uso, ésta recibe mucha atención de los medios. Rápidamente es avalada por la Asociación Médica Estadounidense, la Asociación de Urología de Estados Unidos y la Universidad Estadounidense de Cirujanos. Un grupo comunitario de apoyo a la salud, fundado por pacientes de cáncer de próstata y sus familias, comienza un llamado difundido para programas de detección sistemática usando la nueva prueba. Su objetivo es lograr que todos los hombres adultos en los Estados Unidos realicen una prueba anual de detección de cáncer de próstata. Analice las ventajas y desventajas de dicho plan de someter al público a un tamizaje general para la detección del cáncer de próstata.

Respuestas a las preguntas de estudio y auto-evaluaciones

P9.1

1. 2
2. 7
3. 100 x (6/8) = 75%
4. 100 x (7/10) = 79%

P9.2

1. Falsa. Normalmente se utilizan varios momentos y ubicaciones en la misma residencia, y a menudo se calcula un promedio ponderado por tiempo (PPT).
2. Falsa. Una buena instrumentación para medir el promedio ponderado por tiempo ha estado disponible desde hace un tiempo.
3. Falsa. Un sistema de códigos de cables para medir la distancia y la configuración se utiliza sistemáticamente desde 1979 para calificar a los hogares en forma rudimentaria según la intensidad de los CEM. Sin embargo, la utilidad de este sistema para predecir la exposición pasada continúa siendo una pregunta abierta.
4. Verdadera
5. Verdadera
6. Verdadera

(Nota: no hay preguntas numeradas del 7 al 9).

10. Sesgo del entrevistador. Al considerar los antecedentes del uso de anticonceptivos orales, los individuos que con certeza experimentaron una trombosis venosa podrían ser investigados más exhaustivamente que los controles.
11. Se aleja del valor nulo. La proporción de expuestos entre los controles sería menor que la que debería haber habido si tanto los casos como los controles fueran investigados en la misma medida. En consecuencia, el *odds ratio* en los datos clasificados erróneamente sería más alto que lo que debería ser.
12. Manteniendo a los entrevistadores en desconocimiento de la condición de caso y control del individuo del estudio.
13. 5
14. 70
15. 100 x (95/100) = 95%
16. 100 x (70/80) = 87,5%

P9.3

1. El valor de la razón de riesgos para los datos observados es (380/1000)/(240/1000) = 1,58.
2. Porque la razón de riesgos observada de 1,58 es distinta de la verdadera (es decir, correcta) razón de riesgos de 3,14.
3. Hacia el valor nulo, ya que el valor sesgado de 1,58 está más cerca del valor nulo que el valor correcto.
4. No hay forma de saberlo a partir de un solo ejemplo, pero la respuesta es no, el sesgo podría ser hacia el valor nulo o alejarse de él.
5. El OR observado de 2,6, que surge de clasificar erróneamente la exposición, es distinta del verdadero *odds ratio* de 3,5.
6. La dirección del sesgo es hacia el valor nulo. El valor sesgado del OR de 2,6 está más cerca del valor nulo de 1 que el OR correcto.

P9.4

1. Sensibilidad = 720 / 900 = 0,8 u 80%.
2. Especificidad = 910 / 1100 = 0,83 u 83%.
3. Sí. Tanto la sensibilidad como la especificidad son menores que 1. Sin embargo, sin corregir el sesgo, no queda claro que el grado del sesgo será grande.

P9.5

1. Sensibilidad = 480/600 = 0,80 u 80% y especificidad = 380/400 = 0,95 ó 95%.
2. Sensibilidad = 240/300 = 0,80 u 80% y especificidad = 665/700 = 0,95 ó 95%.
3. Las sensibilidades para los casos de EC y los no casos son iguales. Asimismo, las especificidades para los casos de EC y los no casos son iguales. La información de sensibilidad indica que el 20% de los casos y los no casos con bajo consumo de frutas y verduras tienden a sobrestimar su consumo. La información de especificidad indica que sólo el 5% de los casos y los no casos con alto consumo tienden a subestimar su consumo.
4. En el cuadro de 2 x 2 correctamente clasificado, a = 600, b = 400, c = 300 y d = 700, por lo tanto, el valor del *odds ratio* es ad/bc = (600 x 700) / (400 x 300) = 3,5.

5. El OR observado de 2,6, que surge de clasificar erróneamente la exposición, es distinto del verdadero *odds ratio* de 3,5.
6. La dirección del sesgo es hacia el valor nulo. El OR sesgado de 2,6 está más cerca del valor nulo de 1 que el OR correcto.

P9.6

1. Sensibilidad = 580/600 = 0,97 ó 97% y especificidad = 380/400 = 0,95 ó 95%.
2. Sensibilidad = 240/300 = 0,80 u 80% y especificidad = 665/700 = 0,95 ó 95%.
3. No. Aunque las especificidades para los casos y los no casos son iguales (es decir, 95%), la sensibilidad para los casos (97%) es bastante distinta de la sensibilidad para los no casos (80%). Esta diferencia entre sensibilidades indica que los casos con bajo consumo de frutas y verduras tienen menos probabilidades de sobrestimar su consumo que los no casos.
4. No mucho. El OR observado de 3,95, que surge de clasificar erróneamente la exposición, es levemente más alto que el verdadero *odds ratio* de 3,50.
5. El sesgo se aleja apenas del valor nulo. El OR sesgado de 3,95 se aleja más del valor nulo de 1 que el OR correcto de 3,5.

P9.7

1. No. Suponer una **clasificación errónea independiente** no equivale a suponer una **clasificación errónea no diferencial**. La última supone que la forma en que un individuo clasifica la exposición no variará con su condición de verdadera enfermedad, es decir, Pr(clasificación de la condición de enfermedad|verdaderamente E) = Pr(clasificación de la condición de enfermedad|verdaderamente no E) o Pr(clasificación de la condición de exposición|verdaderamente D) = Pr(clasificación de la condición de exposición|verdaderamente no D).
2. Diferencial porque (Sensibilidad D|E) no es igual a (Sensibilidad D|no E) = 0,75.
3. Sí. Las siguientes faltan:
   (Especificidad D|E)
   (Especificidad D|no E)
   (Especificidad E|no D)
4. Pr(D´E´ | D E) = Pr(D´|D E) x Pr(E´|D E) = (Sensibilidad D|E) x (Sensibilidad E|D)
5. Pr(D´ E´|D E) = 0,8 x 0,9 = 0,72.
6. Pr(D´ E´|D no E) = Pr(D´|D no E) x Pr(E´|D no E) = (Sensibilidad D|no E) x (1 – (Especificidad E|D)).
7. Pr(D´ E´|D no E) = 0,75 x (1 – 0,95) = 0,0375.

P9.8

1. Falso. Además de las sensibilidades, si las especificidades tanto para los expuestos como para los no expuestos son iguales, entonces el sesgo debe ser hacia el valor nulo.
2. Verdadero
3. Verdadero

P9.9

a) Se aleja
b) Hacia
c) Hacia
d) Se aleja
e) Se aleja
f) Se aleja
g) Hacia
h) Hacia

1. Depende. Sabemos que el sesgo debe ser hacia el valor nulo. Si la dirección del sesgo es todo lo que nos interesa, entonces no necesitamos corregir el sesgo. Sin embargo, si queremos determinar el alcance del sesgo y obtener una medida cuantitativa de efecto verdadero, entonces debemos corregir el sesgo.
2. Podemos inferir que la clasificación errónea es no diferencial por nuestro conocimiento o experiencia con las variables de exposición y enfermedad de nuestro estudio, o podemos basar nuestra decisión sobre cálculos confiables de los parámetros de sensibilidad y especificidad.
3. Depende. El sesgo puede ser hacia el valor nulo o puede alejarse de éste. Podríamos determinar la dirección del sesgo mediante el razonamiento lógico sobre las características del estudio. De lo contrario, la única forma en que podemos determinar el alcance o la

dirección del sesgo es mediante la comparación de un cálculo corregido con un cálculo observado.

4. El *odds ratio* observado y sesgado (es decir, clasificado erróneamente) se acerca más al valor nulo que el *odds ratio* corregido.
5. El mayor alcance de sesgo se ve con el OR observado de 1,5, en comparación con el OR corregido de 3,5, que surge cuando tanto la sensibilidad como la especificidad son del 80%.
6. El sesgo es más bajo cuando el OR corregido es 1,8, lo que sucede cuando tanto la sensibilidad como la especificidad equivalen al 90%.
7. Una forma de decidir es elegir el OR corregido correspondiente al conjunto de valores de la sensibilidad y la especificidad más realista. Otra forma es elegir el OR corregido (aquí, 3,5) que se aleja más del OR observado. Una tercera alternativa es optar por el OR corregido que cambia menos (aquí, 1,8) con respecto al OR observado.

P9.10

1. Para los hombres, SeD = 32/40 = 80% y SpD = 54/60 = 90%.
2. Para las mujeres, SeD = 16/20 = 80% y SpD = 72/80 = 90%.
3. No diferencial: la SeD para hombres y mujeres es igual (80%), y la SpD para hombres y mujeres es igual (90%).
4. RR(ajustada) = (A/1000)/(B/1000) = (400/1000)/(200/1000) = 2,0.
5. El sesgo es hacia el valor nulo, ya que el valor sesgado de la razón de riesgos de 1,58 se acerca más al valor nulo que la razón de riesgos corregida.
6. q será cero si tanto SeD como SpD suman 1. Por ejemplo, si tanto SeD como SpD son iguales a 0,5, entonces q = 0. En este caso, habría iguales posibilidades de clasificación errónea en cualquiera de las cuatro celdas de los cuadros de 2 x 2. No tendría sentido calcular los valores del efecto corregido para dicha situación, ya que la clasificación errónea habría invalidado completamente los resultados del estudio.

P9.11

1. SeE = 48/60 = 80% y SpE = 133/140 = 95%.
2. Necesitaría estratificar el cuadro de clasificación errónea en dos cuadros, uno para los casos y el otro para los controles, y luego determinar si las sensibilidades y las especificidades correspondientes a los casos y los controles eran iguales. Sin dicha información, podría inferir que ya que tanto los casos como los controles padecen cáncer gastrointestinal, pueden ser propensos a tener tendencias de información similares sobre los antecedentes de UPE. Tal argumento sugiere que es probable que la clasificación errónea sea no diferencial.
3. Nota: no hay pregunta 3 en esta sección.
4. OR(ajustado) = (A x D)/(B x C) = (80 x 280)/(220 x 20) = 5,1.
5. El sesgo es hacia el valor nulo, ya que el valor sesgado de la razón de riesgos de 3,0 está más cerca del valor nulo que la razón de riesgos corregida de 5,1.

P9.12

1. Estratificando la información de clasificación relativa al nivel de contaminación según la condición de verdadera enfermedad, y haciendo una estratificación de la información de clasificación relativa a la enfermedad en función del nivel de contaminación verdadero.
2. No realmente, porque no se brinda la información real de clasificación errónea estratificada que involucra la condición de enfermedad y el nivel de contaminación verdaderos.
3. Se supone una clasificación errónea no diferencial, ya que no se brindan valores de sensibilidad y especificidad particulares para cada estrato en la submuestra ni en el estudio previo sobre contaminación que se aplicará aquí. El supuesto de clasificación errónea no diferencial parece razonable. No es probable que la enfermedad se haya informado erróneamente una semana después según el nivel de contaminación, o que la calidad del agua se haya medido o informado incorrectamente según el nivel de enfermedad.
4. Es razonable suponer una clasificación independiente, ya que las variables de

exposición y enfermedad fueron medidas en momentos diferentes y probablemente por investigadores distintos.

5. RR(corregida) = (339,3/416,7)/(410,8/2083,4) = 4,1
6. El sesgo es hacia el valor nulo porque el valor sesgado de la razón de riesgos de 1,6 se acerca más al valor nulo que la razón de riesgos corregida de 4,1.
7. El valor sesgado de 1,6 sube bastante a 4,1 cuando se le corrige. Un salto tan grande del valor sesgado al corregido ocurre a menudo cuando, tanto la enfermedad como la exposición, están clasificadas erróneamente, incluso cuando los parámetros de sensibilidad y especificidad son cercanos al 100%.
8. Si la sensibilidad y especificidad para la enfermedad suman 1 o si la sensibilidad y especificidad para la exposición suman 1, entonces las frecuencias de celda corregidas no están definidas porque q* = 0.

P9.13

1. 7,4. Para responder a esta pregunta deberá utilizar la fórmula adecuada para corregir la clasificación errónea no diferencial de la enfermedad; el cuadro corregido es:

	E	**no E**	
D	68	12	80
no D	52	68	120
Total	120	80	200

2. Es hacia el
3. Verdadero; si la suma de la especificidad y la sensibilidad es igual a 1, obtendrá resultados indeterminados.
4. Verdadero
5. Falso; ya que la suma de cualquiera de las especificidades y sensibilidades es igual a 1, los resultados serán indeterminados.

P9.14

1. Diferencial porque las sensibilidades de 96,7% y 80% son distintas, aunque las especificidades sean iguales.
2. Un caso de EC, al que le podría preocupar las razones de su enfermedad, no tiene tantas probabilidades de sobrestimar su consumo de frutas y verduras como un control.
3. OR(corregido) = (599,7 x 700) / (400,2 x 300) = 3,5. Éste es el mismo valor que obtuvimos previamente para el verdadero *odds ratio* en nuestra presentación anterior sobre la clasificación errónea diferencial que mostraba cómo obtener frecuencias de celda observadas cuando se comienza con las frecuencias de celda verdaderas.
4. El sesgo se aleja del valor nulo, ya que el valor sesgado del *odds ratio* de 3,95 se aleja más del valor nulo que el *odds ratio* corregido de 3,5.

P9.15

1. 3,5. Véanse las fórmulas adecuadas que se necesitan para calcular este valor. A = 600, B = 400, C = 300, D = 700.
2. Se aleja del; como el OR observado se aleja más del valor nulo que el valor ajustado, el valor observado debe alejarse del valor nulo.

P9.16

1. Sensibilidad = 48 / 60 = 0,80.
2. Es muy poco probable que el paciente padezca la enfermedad, ya que la probabilidad de obtener un resultado negativo en la prueba para un paciente con la enfermedad es 0,01, que es muy baja.
3. Especificidad = 126 / 140 = 0,90.
4. Es muy probable que el paciente padezca la enfermedad, ya que la probabilidad de obtener un resultado positivo para un paciente sin la enfermedad es 0,01, que es muy baja.
5. Prevalencia de la verdadera enfermedad = 60 / 200 = 0,30.
6. No se puede responder completamente esta pregunta. Tanto la sensibilidad como la especificidad son relativamente altas (0,80 y 0,90), pero la prevalencia es sólo del 30%. Lo que se necesita es la proporción de resultados positivos totales de la ecografía de quienes verdaderamente padecen TVP, que en este estudio es 48 / 62 = 0,77, lo que es alto, pero no supera 0,90 ó 0,95.

P9.17

1. La opción A es el valor predictivo y la opción B es la sensibilidad.
2. VP+ = 48 / 62 = 0,77
3. Sobre la base del cuadro, la probabilidad anterior de desarrollar TVP es 60 / 200 = 0,30, que es la prevalencia de la enfermedad entre los pacientes estudiados.

4. Sí. La probabilidad anterior fue 0,30, mientras que la probabilidad (posterior a la prueba) mediante una ecografía aumentó a 0,77 dado un resultado positivo en la prueba.
5. VP- = 126 / 138 = 0,91
6. Sobre la base del cuadro, la probabilidad anterior de **no** desarrollar TVP es 140 / 200 = 0,70, que es 1 menos la prevalencia de la enfermedad entre los pacientes estudiados.
7. Sí. La probabilidad anterior de no desarrollar TVP fue 0,70, mientras que la probabilidad (posterior a la prueba) de no desarrollar TVP mediante una ecografía aumentó a 0,91 dado un resultado negativo en la prueba.
8. Sensibilidad = 16 / 20 = 0,80, especificidad = 162 / 180 = 0,90, prevalencia = 20 / 200 = 0,10.
9. Los correspondientes valores de la sensibilidad y especificidad son idénticos en ambos cuadros, pero la prevalencia calculada para estos datos es mucho menor (0,10) que la calculada para el cuadro anterior (0,30).
10. VP+ = 16 / 34 = 0,47 y VP- = 162 / 166 = 0,98.
11. El VP+ ha disminuido de 0,77 a 0,47 y el VP- ha aumentado de 0,91 a 0,98, mientras que la prevalencia ha caído de 0,30 a 0,10, en tanto que la sensibilidad y la especificidad se han mantenido iguales y altas.
12. Si la prevalencia disminuye, el valor predictivo positivo disminuirá y puede ser bastante bajo, incluso si la sensibilidad y la especificidad son altas. De forma similar, el valor predictivo negativo aumentará y puede ser muy alto, incluso si la sensibilidad y la especificidad no son muy altas.

P9.18

1. 90%
2. 90%
3. 33,3%
4. 81,8%
5. 16,7%
6. 64,3%
7. menor
8. 10%
9. 50%
10. menor
11. bajo, altos

LECCIÓN 10

SESGO DE CONFUSIÓN

*La **confusión** es una forma de sesgo que se refiere a la manera en que una medida de asociación puede cambiar de valor si en el análisis se controlan variables distintas de la variable de exposición.*

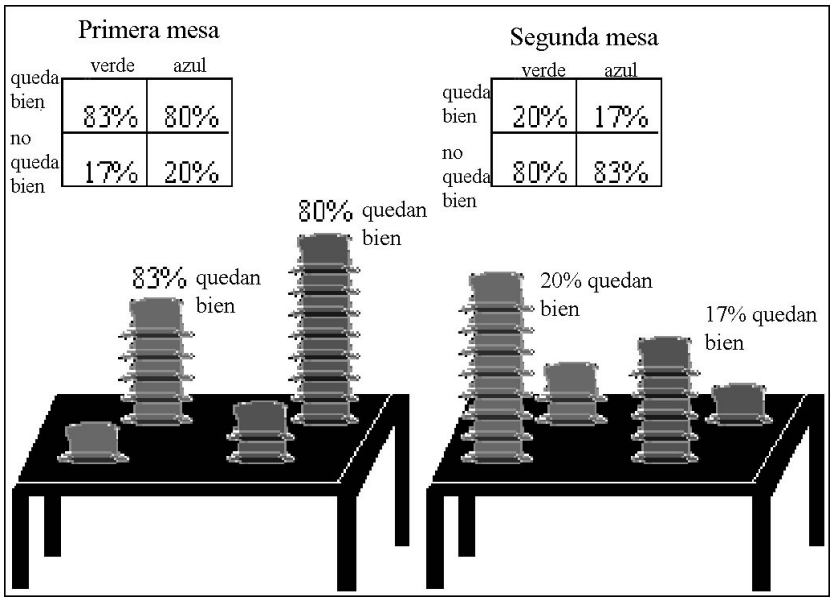
10-1 Concepto y definición

La paradoja de Simpson

La paradoja de Simpson se ejemplifica mediante una historia sobre la compra de un sombrero, en la que se llega a una conclusión diferente sobre qué color queda mejor cuando la misma colección de sombreros se mueve de mesas separadas a una única mesa.

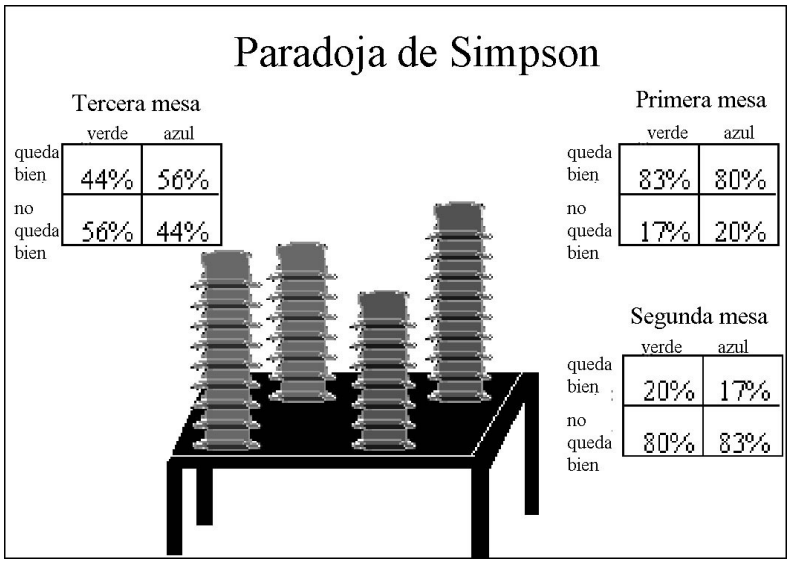
Un hombre entra a un establecimiento comercial para comprar un sombrero. Ve dos mesas en las que hay sombreros de sólo dos colores, verde y azul. La primera mesa tiene 6 sombreros verdes y 10 azules. Se da cuenta de que 5 de los 6 sombreros verdes y 8 de los 10 azules le quedan bien. Por lo tanto, 83% de los sombreros verdes y 80% de los sombreros azules de la primera mesa le quedan bien.

La otra mesa tiene 10 sombreros verdes y 6 azules. Sólo 2 de los 10 sombreros verdes y 1 de los 6 sombreros azules le quedan bien. Por lo tanto, 20% de los sombreros verdes y 17% de los sombreros azules de la segunda mesa le quedan bien. Sin embargo, el hombre vuelve a su casa sin haber elegido ningún sombrero.



(Nota: los sombreros verdes están en el lado izquierdo de las mesas, los azules en el lado derecho).

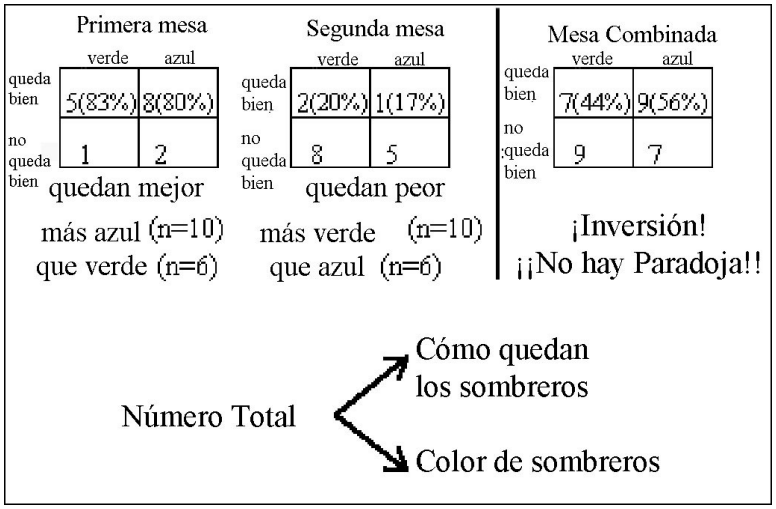
Al día siguiente, el hombre vuelve el establecimiento comercial justo cuando abre y recuerda que el día anterior, un mayor porcentaje de sombreros verdes que de azules de cada mesa le quedaba bien. Pero, ahora, los mismos 32 sombreros están mezclados en una sóla mesa grande. Como él es un comprador muy meticuloso, vuelve a probarse todos los sombreros. Se da cuenta que 7 de los 16 sombreros verdes y 9 de los 16 sombreros azules le quedan bien. Por lo tanto, 44% de los sombreros verdes y 56% de los sombreros azules le quendan bien. ¡Ahora los sombreros azules parecen quedarle mejor que los verdes! ¿Cómo se puede explicar este resultado invertido? Esta situación ejemplifica un resultado sorprendente llamado paradoja de Simpson.



(Nota: los sombreros verdes están en el lado izquierdo de la mesa, los azules en el lado derecho).

Explicación de la paradoja de Simpson

En nuestra historia sobre la compra de un sombrero, vemos que, aunque los sombreros **verdes** quedaban mejor que los azules cuando los sombreros estaban colocados en **mesas separadas**, los sombreros **azules** quedaban mejor **en general** cuando todos los sombreros estaban colocados en **una sola mesa**. De esta manera, la combinación de todos los sombreros en una sola mesa muestra un resultado invertido. La aparente paradoja surge aquí porque hay dos relaciones importantes subyacentes que involucran a las tres variables consideradas en esta historia. Estas variables son: el número de mesa; el resultado de cómo queda el sombrero; y el color del sombrero. El número de mesa se relaciona tanto a cómo queda el sombrero así como al color del sombrero. La primera asociación que vincula el número de mesa a cómo queda el sombrero significa que, para cada respectivo color de sombrero, los sombreros en la mesa 1 quedan mejor que los de la mesa 2. Sin embargo, la segunda asociación que vincula el número de mesa al color del sombrero indica que hay más sombreros azules que verdes en la mesa 1, donde los sombreros quedan mejor, y hay más sombreros verdes que azules en la mesa 2, donde los sombreros quedan peor. La inversión observada no es realmente una paradoja porque dicha inversión puede suceder si en verdad, como en nuestra historia, el número de mesa está asociado tanto a cómo quedan los sombreros así como al color del sombrero. En epidemiología, llamamos a este principio **confusión basada en los datos**. Describiremos ésto con más detalle en las lecciones siguientes.



(Nota: el "número de mesa" es el factor de confusión (C); "cómo queda el sombrero" es el resultado (D) y el "color del sombrero" es la exposición (E)).

Resumen

- La paradoja de Simpson ejemplifica un principio general sobre la relación entre dos variables, E y D, cuando también se considera una tercera variable, C.
- El principio: si la variable C está fuertemente relacionada tanto con E como con D, entonces la relación entre E y D para los datos combinados, que ignora a C, puede ser diferente de la relación entre E y D cuando estratificamos por categorías de C.
- El principio es el fundamento del **criterio basado en los datos** para la confusión, el cual se describe en las actividades siguientes.

El concepto de confusión en la epidemiología

La paradoja de Simpson es un ejemplo de la equivocación que puede provocarse debido a la **confusión**. La confusión es un problema importante para los investigadores de salud públic y medicina cuando llevan a cabo estudios para evaluar una relación entre una exposición, **E**, y algún resultado de salud o enfermedad de interés, **D**. La confusión es un tipo de sesgo que puede ocurrir cuando no tomamos en cuenta otras variables, como, por ejemplo, edad, sexo o condición de tabaquismo, al intentar evaluar una relación del tipo E→D.

Para ejemplificar la confusión, considere los resultados de un estudio hipotético de cohorte retrospectivo para determinar el efecto de la exposición a un químico tóxico sospechado en el desarrollo de cáncer de pulmón en trabajadores de una industria química. Llamaremos al químico "TCX". Se calcula que el riesgo en diez años de desarrollar cáncer de pulmón es de 0,36 para los que estuvieron expuestos al TCX y de 0,17 para los que no estuvieron expuestos a este químico. El valor de la razón de riesgos es de 2,1, lo que indica que aquellos expuestos al TCX tienen un riesgo dos veces mayor de desarrollar cáncer de pulmón que los no expuestos. Hasta ahora, hemos considerado dos variables: la exposición al TCX y la condición de cáncer de pulmón (el resultado de salud).

Trabajadores de una industria química	TCX	no TCX	Total
CP	27	14	41
No CP	48	67	115
Total	75	81	156

Riesgo en diez años de desarrollar CP

TCX: $27/75 = 0.36$

no TCX: $14/81 = 0.17$

$$\hat{RR} = 0.36/0.17 = 2.1$$

Todavía no hemos considerado ninguna otra variable que también podría haberse medido u observado en los pacientes de este estudio. Por ejemplo, podríamos preguntarnos si había relativamente más fumadores entre los que estuvieron expuestos al TCX que entre los no expuestos. De ser así, eso podría explicar por qué se descubrió que los trabajadores expuestos al TCX tenían un riesgo mayor de 2,1, en comparación con los trabajadores no expuestos. Simplemente, quizás los que estuvieron expuestos al TCX fumaron más y, por lo tanto, tuvieron más probabilidad de desarrollar cáncer de pulmón que los no expuestos al TCX, independientemente de la exposición a este químico. Quizás, la exposición al TCX es un determinante de alguna otra forma de cáncer u otra enfermedad, pero no necesariamente del cáncer de pulmón.

Suponga que clasificamos nuestros datos del estudio en dos categorías relacionadas con los antecedentes de tabaquismo: no fumadores y fumadores. Para estos cuadros, el cálculo de la razón de riesgos da 1,0 para los no fumadores y 1,3 para los fumadores. Observe que estas razones de riesgos específicas para dos estratos sugieren que no existe ninguna asociación entre la exposición al TCX y el desarrollo de cáncer de pulmón.

Trabajadores de una industria química	TCX	no TCX	Total
CP	27	14	41
No CP	48	67	115
Total	75	81	156

$$\hat{RR} = 2.1$$

no fumadores

	TCX	no TCX	Total
CP	1	2	3
No CP	24	48	72
Total	25	50	75

fumadores

	TCX	no TCX	Total
CP	26	12	38
No CP	24	19	43
Total	50	31	81

$\hat{RR} = 1.0$ no existe ninguna asociación $\hat{RR} = 1.3$

Preguntas de estudio (P10.1)

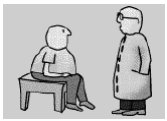
1. Concluir que no hay asociación después de controlar el tabaquismo a partir de estos datos es debatible. ¿Puede explicar por qué?

Cuando formamos estratos estableciendo categorías en todo el conjunto de datos según una o más variables, como antecedentes de tabaquismo en este ejemplo, decimos que estamos **controlando** estas variables, a las que a menudo denominamos **variables de control**. Por lo tanto, lo que parece un aumento al doble en el riesgo cuando ignoramos los antecedentes de tabaquismo, cambia a una ausencia de asociación cuando controlamos dichos antecedentes. Esto sugiere que la razón por la cual el riesgo de los trabajadores expuestos al TCX aumentó al doble, en comparación con los trabajadores no expuestos, podría explicarse simplemente observando que hay relativamente más fumadores entre los expuestos al TCX. Éste es un ejemplo de lo que llamamos **confusión**, y decimos que los antecedentes de tabaquismo constituyen un factor de confusión de la relación entre la condición de exposición al TCX y el riesgo en diez años de padecer cáncer de pulmón. En general, la confusión puede describirse como una distorsión en una medida de asociación, como ser una razón de riesgos, que puede surgir debido a que no controlamos otras variables, por ejemplo, los antecedentes de tabaquismo, que podrían ser factores de riesgo para el resultado de salud que se está estudiando. Si no controlamos el factor de confusión, obtendremos un cálculo incorrecto, o sesgado, de la medida *de efecto*.

Resumen

- La confusión es una distorsión en una medida *de efecto* como, por ejemplo, la RR, que puede surgir debido a que no controlamos otras variables, por ejemplo, los antecedentes de tabaquismo, que son factores de riesgo previamente conocidos para el resultado de salud que se está estudiando.
- Si ignoramos el efecto de un factor de confusión, obtendremos un cálculo incorrecto, o sesgado, de la medida *de efecto*.

Auto-evaluación (P10.2)



1. La confusión es una **???** en una **???** que puede surgir debido a que no **???** otras variables que son **???** previamente conocidos para el resultado de salud que se está estudiando.

Opciones
estudio de casos y controles **controlamos** **distorsión** **modificadores de efecto** **eliminamos** **medida** de efecto **factores de riesgo**

Un estudio halla que el consumo de alcohol está asociado al cáncer de pulmón, OR crudo = 3,5. Mediante los siguientes datos, determine si el tabaquismo podría estar confundiendo esta relación.

2. ¿Cuál es el OR entre los fumadores? . . . . . **???**
3. ¿Cuál es el OR entre los no fumadores? . . . . **???**
4. ¿Cambia el OR cuando controlamos la condición de tabaquismo? . **???**
5. ¿Hay evidencia en estos datos de que el tabaquismo es un factor de confusión de la relación entre el consumo de alcohol y el cáncer de pulmón? . . . **???**

Opciones
0,01 **1,0** **3,5** **5,4** **no** **sí**

	Fumadores			No fumadores	
	Alcohol	No Alc		Alcohol	No Alc
Casos	560	140	Casos	40	160
Controles	240	60	Controles	160	640

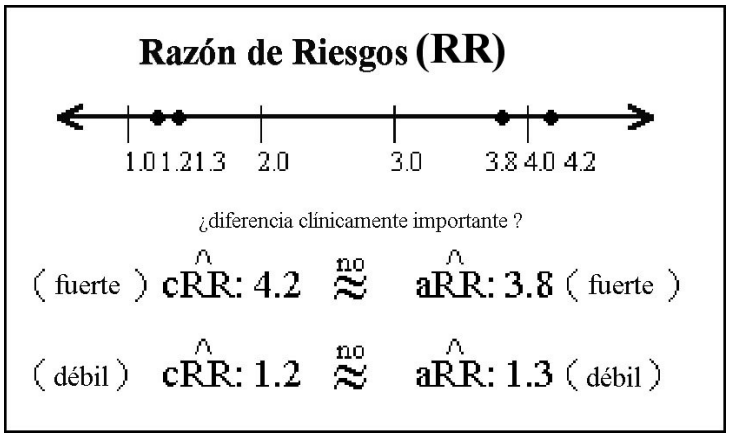
Cálculos brutos versus ajustados

En los estudios epidemiológicos, la confusión se evalúa comparando el **cálculo crudo , o bruto, de efecto** (por ejemplo, $c\hat{R}R$), en donde no se controla ninguna variable, con el **cálculo ajustado de efecto** (por ejemplo, $a\hat{R}R$), en donde se controlan una o más variables. El cálculo ajustado normalmente se calcula mediante la combinación de los cálculos específicos para cada estrato en un único número.

Por ejemplo, para evaluar la confusión por antecedentes de tabaquismo en el estudio de cohorte retrospectivo sobre los efectos de la exposición al químico TCX en el desarrollo de cáncer de pulmón descrito anteriormente, podemos comparar la razón de riesgos cruda o bruta de 2,1 con una razón de riesgos ajustada que combina las razones de riesgos de 1,0 y 1,3 para las dos categorías de antecedentes de tabaquismo. El método para combinar estos dos cálculos en una única medida de resumen es el tema de la siguiente actividad. Una vez que hemos combinado estos cálculos específicos para cada estrato, ¿cómo decidimos si hay confusión? El **criterio basado en los datos** para establecer si hay evidencia de confusión requiere que el cálculo crudo de efecto sea distinto del cálculo ajustado de efecto. ¿Qué tan distintos deben ser estos dos cálculos para llegar a la conclusión de que hay confusión? Para responder a esta pregunta, el investigador debe decidir si existe o no una diferencia clínicamente relevante.

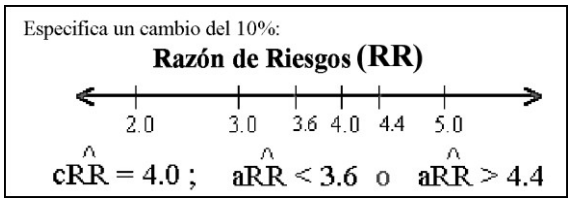
En nuestro estudio de cohorte retrospectivo, el cálculo ajustado sería un número entre 1,0 y 1,3, lo que sugiere una relación mucho más débil que la indicada por el cálculo crudo de 2,1. La mayoría de los investigadores considerarían que ésta es una diferencia clínicamente importante. Suponga que el cálculo crudo hubiera dado 4,2, en lugar de 2,1: la diferencia entre el valor del cálculo crudo y el ajustado indicaría una confusión incluso mucho más fuerte.

Podemos comparar otros cálculos crudos y ajustados. Suponga, por ejemplo, que una razón de riesgos cruda dio 4,2 y que la razón de riesgos ajustada dio 3,8. Ambos valores indican una asociación de casi igual magnitud; por lo tanto, no hay una diferencia clínicamente importante entre los valores del cálculo crudo y ajustado. En forma similar, si la razón de riesgos cruda es 1,2 y la razón de riesgos ajustada es 1,3, estos dos valores indican que existe aproximadamente la misma relación muy débil o ninguna asociación. Por lo tanto, aquí, nuevamente, no hay una diferencia clínicamente importante entre el valor crudo y ajustado.

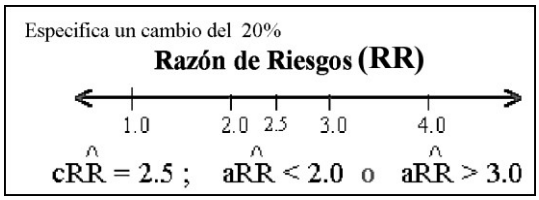


Claramente, la determinación de lo que es clínicamente importante requiere una decisión subjetiva por parte de los investigadores. Un investigador puede llegar a la conclusión, por ejemplo, de que la diferencia entre 1,2 y 1,3 es clínicamente importante, mientras que otro investigador puede llegar a la conclusión contraria. Este problema nos puede llevar a querer utilizar una prueba de significancia estadística para decidir si hay una diferencia entre el cálculo crudo y el ajustado. Sin embargo, **debido a que la confusión es un problema de validez, no debería evaluarse mediante una prueba estadística**, sino, más bien, buscando una diferencia considerada relevante, aunque sea imprecisa.

Un abordaje comúnmente utilizado para evaluar la confusión es especificar *a priori,* o antes de observar los datos, qué margen de diferencia entre un cálculo bruto y uno ajustado se requiere en nuestro estudio para considerar dicha diferencia "relevante". Normalmente, se especifica un cambio o diferencia de 10%, de forma tal que si la razón de riesgos cruda es, por ejemplo, 4, entonces el margen de cambio determinado *a priori* es de $\pm$ 10%, para una razón de riesgos ajustada de 3,6 a 4,4. Por lo tanto, si la razón de riesgos ajustada fuese menor que 3,6 ó mayor que 4,4, diríamos que hay evidencia de sesgo de confusión basada en los datos con, al menos, un margen de cambio de 10% en la asociación calculada.



Para dar otro ejemplo, si se especifica un cambio *a priori* de 20% y la razón de riesgos cruda es, por ejemplo, 2,5, la razón de riesgos ajustada tendría que ser menor que 2 o mayor que 3 para llegar a la conclusión de que hay confusión.



Resumen

- La confusión se evalúa comparando el cálculo crudo de efecto, en donde no se controla ninguna variable, con el cálculo ajustado de efecto, en donde se controlan una o más variables.
- Hay confusión si llegamos a la conclusión de que hay una diferencia clínicamente importante o considerada relevante entre los valores del cálculo crudo y ajustado.
- No utilizamos pruebas estadísticas para evaluar la confusión.
- Un abordaje comúnmente utilizado para evaluar la confusión es especificar, antes de observar los datos, qué margen de diferencia entre un cálculo bruto y uno ajustado se requiere en nuestro estudio para considerar dicha diferencia "relevante".

Auto-evaluación (P10.3)

Se realizó un estudio de casos y controles para evaluar la relación entre el uso de anticonceptivos orales y el cáncer de ovario. El OR crudo dio 0,77. Ya que la edad se consideró un factor potencial de confusión en este estudio, los datos se estratificaron en 3 grupos etarios, como se muestra a continuación.

1. El OR para el grupo de 20-39 años es **???**
2. El OR para el grupo de 40-49 años es **???**

3. El OR para el grupo de 50-54 años es **???**
4. ¿Cree que estos datos brindan alguna evidencia de que la edad es un factor de confusión de la relación entre el uso de anticonceptivos orales y el cáncer de ovario? **???**

Opciones
0,58 0,61 0,65 0,69 0,77 1,45 no sí

	Edades 20-39		Edades 40-49		Edades 50-54	
	COs	No COs	COs	No COs	COs	No COs
Caso	46	12	30	30	17	44
Control	285	51	463	301	211	331

En el estudio descrito en la pregunta anterior, el OR crudo fue 0,77 y el OR ajustado que controla la edad fue 0,64.

5. Si el investigador especifica un cambio *a priori* de 15% entre el OR crudo y el ajustado como una diferencia relevante, un OR ajustado menor que **???** o mayor que **???** brinda evidencia de que hay confusión.
6. ¿Hay evidencia de confusión? **???**
7. Suponga que los investigadores determinaron que un margen de cambio de 20% era una diferencia relevante. ¿Hay evidencia de confusión? **???**

Opciones
0,15 0,63 0,65 0,85 0,89 0,92 no sí

10-2 Cálculos ajustados

Características de los cálculos ajustados

Considere nuevamente los resultados de un estudio de cohorte retrospectivo seguido durante 10 años cuyo propósito es determinar si la exposición al químico TCX es un determinante del desarrollo de cáncer de pulmón para los trabajadores de la industria química. La razón de riesgos en 10 años es 2,1. Este valor es nuestro cálculo crudo de efecto porque su cómputo ignora el control de otras variables, como los antecedentes de tabaquismo.

Estudio de cohorte retrospectivo

	TCX	No TCX	Total
CP	27	14	41
No CP	48	67	115
Total	75	81	156

$$\hat{cRR} = 2.1$$

Hacemos un cálculo ajustado combinando los cálculos específicos para cada estrato a fin de obtener una única medida de resumen. La típica medida de resumen utilizada como cálculo ajustado es un promedio ponderado elegido adecuadamente de los cálculos específicos para cada estrato.

En nuestro ejemplo del estudio de cohorte retrospectivo, donde las razones de riesgos para dos categorías de tabaquismo fueron 1,0 y 1,3, una manera simple de obtener esta razón de riesgos ajustada como un promedio

ponderado es utilizar esta fórmula:

Cálculo ajustado $(\hat{aRR})$: promedio ponderado

Razones de riesgos calculadas: 1.0 y 1.3

$$\hat{aRR} = (w_1 \text{ x } 1.0 + w_2 \text{ x } 1.3) / (w_1 + w_2)$$

Promedio aritmético

$$w_1 = w_2$$

$$\hat{aRR} = (1.0 + 1.3) / 2 = 1.15$$

Los términos w_1 y w_2 representan las ponderaciones que se multiplican, respectivamente, por las razones de riesgos específicas para cada estrato y luego se combinan para formar un promedio ponderado. (Nota: una versión más complicada del promedio ponderado que se usa normalmente para los cálculos de las razones de riesgos ajustadas se describe en la sección 14-3 de la lección 14, donde se analizan de nuevo las fórmulas de promedio ponderado que se utilizan en el análisis estratificado). ¿Qué ponderaciones deberíamos utilizar? En un promedio aritmético simple, w_1 es igual a w_2, y el cálculo ajustado se simplifica a la suma de las dos razones de riesgos dividido por 2. El valor 1,15 es un número razonable para utilizar como nuestro cálculo ajustado, porque está en el medio entre las dos razones de riesgos calculadas. Sin embargo, la mayoría de los epidemiólogos prefieren asignar pesos distintos a cada cálculo específico para cada estrato, especialmente si los tamaños de las muestras son muy diferentes entre estratos distintos. Si observamos nuevamente los cuadros para las dos categorías de antecedentes de tabaquismo en nuestro ejemplo, los tamaños de las muestras para cada grupo son 75 y 81, que, aunque no son exactamente iguales, no son muy diferentes. Por lo tanto, quizás un promedio aritmético no es tan malo para este ejemplo.

Estudio de cohorte retrospectivo

Estrato 1: No Fumadores

	TCX	No TCX	Total
CP	1	2	3
No CP	24	48	72
Total	25	50	75

Estrato 2: Fumadores

	TCX	No TCX	Total
CP	26	12	38
No CP	24	19	43
Total	50	31	81

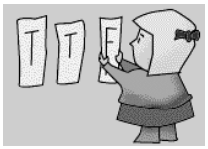
$$\hat{RR}_1 = 1.0 \qquad \hat{RR}_2 = 1.3$$

Por otro lado, si hubiese 10 veces más no fumadores que fumadores, podríamos querer asignar un mayor peso a los no fumadores, de los cuales tenemos más información. De hecho, podríamos considerar asignar 10 veces más peso al primer estrato que al segundo, dando, por consiguiente, el siguiente promedio ponderado, que resulta ser igual a 1,03. Observe que este valor es mucho más cercano a la razón de riesgos de 1,0 para el estrato 1, porque le estamos asignando más peso a dicho estrato.

$$\hat{aRR} = (10 \text{ x } 1.0 + 1 \text{ x } 1.3) / (10 + 1) = 1.03$$

Resumen

- El típico cálculo ajustado es un promedio ponderado elegido adecuadamente de los cálculos específicos para cada estrato de la medida de efecto.
- Los epidemiólogos prefieren utilizar distintas ponderaciones, especialmente si los tamaños de las muestras son diferentes en los distintos estratos.
- Cuando se eligen las ponderaciones según el tamaño de la muestra, los estratos más grandes reciben ponderaciones mayores.
- Una versión más complicada de un promedio ponderado que se usa normalmente para los cálculos ajustados de la RR se describe en la sección 14-3 de la lección 14 sobre análisis estratificado. Esta versión más complicada se llama **RR ajustada basada en la precisión**.



Auto-evaluación (P10.4)

Determine si cada una de las siguientes afirmaciones es **verdadera** o **falsa**:

1. El cálculo ajustado es un promedio ponderado elegido adecuadamente de los cálculos específicos para cada estrato. . . . . **???**
2. El cálculo ajustado siempre es menor que el cálculo crudo correspondiente. . . **???**
3. La mayoría de los epidemiólogos prefieren asignar el mismo peso a cada cálculo específico para cada estrato en los estudios de casos y controles. . . . . **???**
4. La confusión es un problema de validez y, por lo tanto, requiere el uso de una prueba estadística para determinar su significancia. . . . . **???**

Utilice la fórmula que se presenta a continuación para calcular la RR ajustada en los siguientes ejemplos, para los cuales se dan los cálculos específicos para cada estrato. (Nota: aunque la siguiente fórmula brinda un promedio ponderado, la fórmula usual para la aRR es un promedio ponderado más complicado "basado en la precisión", descrito en la lección 14).

$$a\hat{R}R=(w_1 \times \hat{RR}_1 + w_2 \times \hat{RR}_2)/(w_1 + w_2)$$

5. Estrato 1: RR = 1,13; w = 13,1; estrato 2: RR = 1,00; w = 7,7. La RR ajustada es: **???**
6. Estrato 1: RR = 2,25; w = 31,3; estrato 2: RR = 1,75; w = 5,6. La RR ajustada es: **???**

Opciones

1,06 1,07 1,08 1,09 1,98 2,08 2,17

Criterios para el sesgo de confusión

Además del criterio basado en los datos para la confusión, debemos evaluar varios **criterios *a priori***. Éstos son condiciones que se deben tener en cuenta en la etapa de diseño del estudio, antes de recabar los datos, a fin de identificar las variables que se deben medir para un posible control en el análisis de datos.

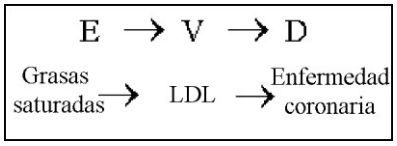
El primer criterio *a priori* es que un **factor de confusión debe ser un factor de riesgo asociado con el resultado de salud**. Este criterio asegura que una asociación cruda entre la exposición y la enfermedad no puede ser explicada convincentemente por otras variables que con certeza predicen la enfermedad. Dichas variables se llaman **factores de riesgo**. Por ejemplo, suponga que estamos estudiando la asociación entre la exposición a un químico tóxico y el desarrollo de cáncer de pulmón en la industria química. Basado en nuestra revisión de la literatura epidemiológica sobre los determinantes del cáncer de pulmón, queremos controlar la edad y la condición de tabaquismo, dos factores de riesgo conocidos. Nuestro objetivo es determinar si la exposición al químico aporta algo más a los efectos de la edad y el tabaquismo en el desarrollo de cáncer de pulmón.

El segundo criterio es que un factor de confusión no puede ser una **variable interviniente** entre la exposición y la enfermedad. La variable interviniente pura (**V**) es cualquier variable cuya relación con la exposición y la enfermedad se encuentra totalmente dentro de la cadena causal entre la exposición y la enfermedad.

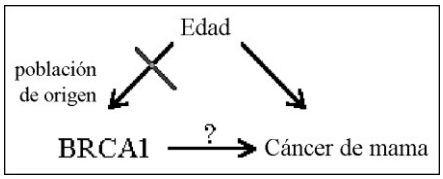
Preguntas de estudio (P10.5)

1. Dado un escenario hipotético donde se miden los niveles de grasas saturadas para determinar sus efectos en la EC, ¿quisieramos controlar los niveles de LDL?

Si controlamos el nivel de LDL, básicamente controlaríamos el nivel de grasas saturadas, y probablemente encontraríamos una razón de riesgos u *odds ratio* ajustado que relacione la grasa saturada con la enfermedad coronaria cerca del valor nulo. La variable interviniente aquí es el nivel de LDL, y no deberíamos controlarla.

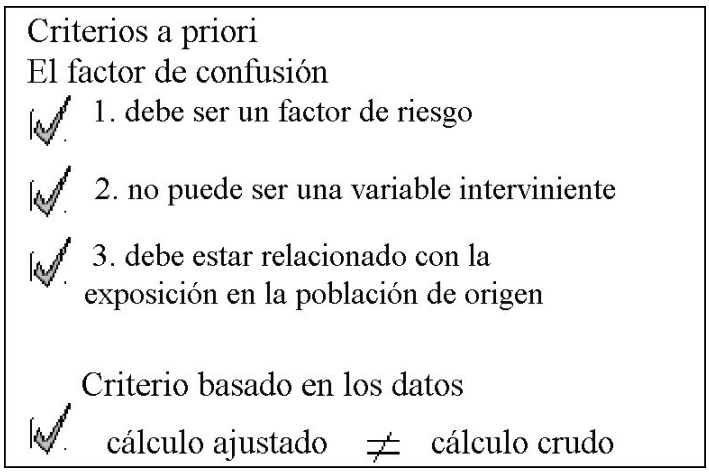


El tercer criterio es que un **factor de confusión debe estar asociado a la exposición en la población de origen que se está estudiando**. Llamamos población de origen a la cohorte de población subyacente de donde surgen los casos que se utilizan en el estudio. Considere un estudio para evaluar si un factor genético específico, BRCAl, es un determinante del cáncer de mama. La edad es un factor de riesgo conocido del cáncer de mama, pero claramente no está asociada a la presencia o ausencia del gen, independientemente de la población de origen que se esté estudiando.



Por lo tanto, no se satisface el tercer criterio *a priori*. En consecuencia, aun cuando por casualidad la edad resultara asociada al gen en los datos del estudio, no la controlaríamos ni siquiera cuando hubiera confusión basada en los datos, porque la edad no satisface todos los criterios *a priori*.

Para que una variable sea considerada un verdadero factor de confusión, se requieren los tres criterios *a priori* en adición al criterio basado en los datos.



Resumen

Un factor de confusión debe satisfacer los 3 criterios *a priori*, además del criterio basado en los datos para la confusión. Éstos se enuncian a continuación:

- El factor de confusión debe ser un **factor de riesgo** asociado con el resultado de salud.
- El factor de confusión **no puede ser una variable interviniente** en la cadena causal entre la exposición y la enfermedad.
- El factor de confusión debe estar **relacionado con la exposición en la población de origen** de la cual se derivan los casos.

Algunas "sutilezas" sobre los factores de riesgo

La decisión sobre qué variables incluir en la lista de factores de riesgo pocas veces es, en la práctica, un asunto bien definido. Este es el caso cuando hay poca literatura disponible sobre un tema de estudio dado. Por otro lado, gran parte de la bibliografía puede ser controvertida en términos de qué variables estudiadas anteriormente son verdaderamente predictivas de la enfermedad.

Asimismo, después de recabar los datos, pero antes del análisis principal de datos, es posible que sea necesario reevaluar la lista de factores de riesgo para que se puedan agregar o eliminar variables ya medidas, pero no consideradas explícitamente para control. Las variables medidas para otros fines, por ejemplo, en un estudio amplio para evaluar varias cuestiones etiológicas, se pueden agregar a la lista de factores de riesgo si anteriormente se habían pasado por alto.

Asimismo, es posible que deba usarse un **sustituto** de un factor de riesgo cuando este último sea difícil de medir. Por ejemplo, a menudo se utiliza el número de años trabajados en un empleo determinado dentro de una industria específica, como una medida sustituta para la cantidad real de tiempo de exposición a una sustancia tóxica que se sospecha que es un agente cancerígeno relacionado con la ocupación.

Algunas "sutilezas" sobre el tercer criterio *a priori*

El tercer criterio *a priori* para la confusión interesa especialmente en los estudios de casos y controles, donde los controles generalmente se seleccionan para el estudio después de que ya hayan sucedido los casos. En dichos estudios, es posible que los datos no sean representativos de la población de origen con respecto a la exposición,

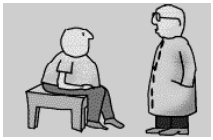
ni tampoco de otras variables.

Por lo tanto, una variable, por ejemplo, **C**, que puede no estar asociada a la exposición en la población de origen puede, sin embargo, estar asociada a la exposición en los datos reales del estudio. En dicho caso, el tercer criterio establece que la variable **C** no puede ser considerada un factor de confusión, y no debería ser controlada, incluso si hay confusión basada en los datos.

En los estudios de cohorte, la condición de exposición se determina antes de que se haya presentado la condición de enfermedad, de forma que la población de origen es la cohorte de estudio. En dichos estudios, una variable, **C**, que no está asociada a la exposición en los datos del estudio, no satisface el tercer criterio y, por lo tanto, no debería ser considerada un factor de confusión.

La principal dificultad al evaluar el tercer criterio *a priori* tiene que ver con la manera de determinar la asociación del factor de confusión sospechado, **C**, con la exposición, **E**, en la población de origen. Esto requiere cierto conocimiento de la literatura epidemiológica referida a la relación entre **C** y **E** y la población de origen que se está estudiando.

Auto-evaluación (P10.6)



Se realizó un estudio para evaluar la relación entre el tipo sanguíneo (O positivo y O negativo) y una enfermedad específica. Como la edad muchas veces está relacionada con el resultado de la enfermedad, se consideró ser un factor potencial de confusión. Determine si se satisface cada uno de los siguientes criterios *a priori*:

1. El factor de confusión debe ser un factor de riesgo para el resultado de salud. . . **???**
2. El factor de confusión no puede ser una variable interviniente entre la exposición y la enfermedad. . . . . **???**
3. El factor de confusión debe estar asociado a la exposición en la población de origen que se está estudiando. . . . . **???**
4. ¿Puede la edad ser un factor de confusión en este estudio? . . . . **???**

Opciones

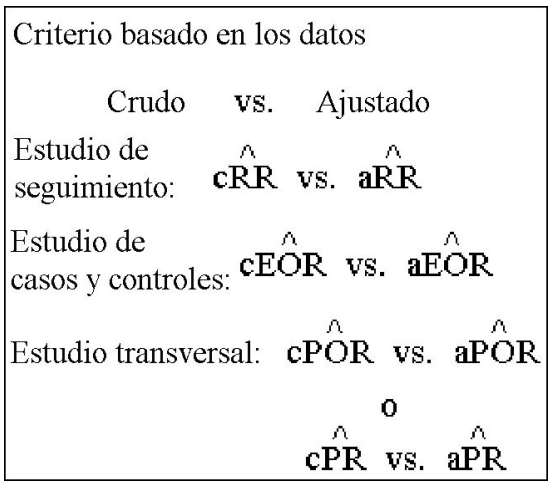
no **no se satisface** **se satisface** **sí**

10-3 Criterios *a priori*/diseños de estudios

Confusión en diferentes diseños de estudios

Hemos visto que la evaluación de la confusión demanda criterios basados en los datos y criterios *a priori*. El criterio basado en los datos requiere que el cálculo crudo de efecto sea relevantemente distinto del cálculo ajustado de efecto. El cálculo ajustado se computa como un promedio ponderado de los cálculos específicos para cada estrato, obtenidos de diferentes categorías del factor potencial de confusión. La medida de efecto utilizada para esta comparación cambia con el diseño del estudio, pero siempre compara un cálculo crudo con uno ajustado.

Hasta este momento, sólo hemos considerado estudios de seguimiento donde la medida de efecto de interés es la razón de riesgos. Para los estudios de casos y controles, comparamos los valores del cálculo crudo y ajustado del *odds ratio* de exposición. En los estudios transversales, comparamos los valores del cálculo crudo y ajustado del *odds ratio* de prevalencia o la razón de prevalencia.



Resumen

- La medida de asociación utilizada para evaluar la confusión dependerá del diseño del estudio.
- En un estudio de seguimiento, normalmente comparamos una razón de riesgos cruda con una razón de riesgos ajustada.
- En un estudio de casos y controles, normalmente comparamos un *odds ratio* de exposición crudo con un *odds ratio* de exposición ajustado.
- En un estudio transversal, normalmente comparamos un *odds ratio* de prevalencia crudo con un *odds ratio* de prevalencia ajustado, o podríamos utilizar una razón de prevalencia en lugar del *odds ratio* de prevalencia.
- Independientemente del diseño del estudio, la confusión basada en los datos se evalúa mediante la comparación del cálculo crudo de efecto con el cálculo ajustado de efecto adecuado.

Evaluación de la confusión en los estudios de casos y controles

Estos cuadros muestran los resultados de un estudio de casos y controles que evalúa la relación entre el consumo de alcohol y el cáncer bucal. Los cuadros describen los datos crudos cuando se ignora la edad, y los datos estratificados describen los datos cuando se ha categorizado la edad en tres grupos. Los investigadores querían evaluar si posiblemente existía confusión debido a la edad.

Estudio de casos y controles
La relación entre el consumo de alcohol y el cáncer bucal

Crudos:

	Alc	no Alc
CB	27	47
no CB	90	443

n = 607

Estratificados:

	Edades 40-49		Edades 50-59		Edades 60+	
	Alc	no Alc	Alc	no Alc	Alc	no Alc
CB	4	25	12	10	11	12
no CB	22	309	37	67	31	67
	n = 360		n = 126		n = 131	

Preguntas de estudio (P10.7)

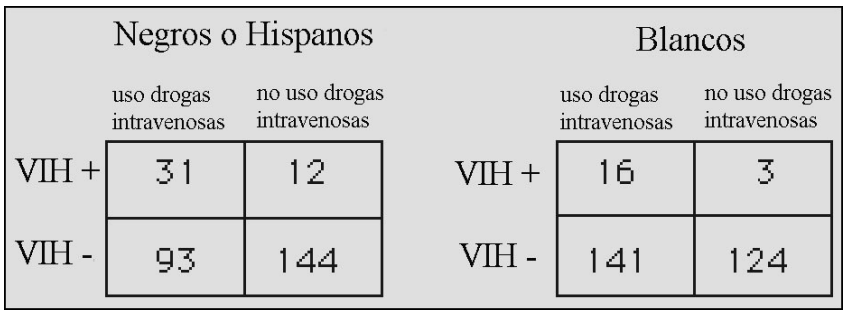
1. ¿Qué expresión debería utilizarse para calcular el *odds ratio* crudo que relaciona el consumo de alcohol con los cánceres bucales?
   A. (27*90)/(47*443)
   B. (27*443)/(47*90)
   C. (27*443)/(47*90)
2. ¿Cuáles son los *odds ratios* específicos para cada estrato?
   D. 2,2; 2,2; 2,0
   E. 0,45; 0,45; 0,5
   F. 0,6; 0,2; 0,2
2. ¿Hay confusión basada en los datos debido a la edad? (Sí o no)
3. Suponiendo que la edad satisfizo las tres condiciones *a priori* para la confusión, ¿por qué la edad es un confusor de la relación entre el consumo de alcohol y el cáncer bucal?

Resumen

- Como ejemplo para evaluar la confusión que involucra el *odds ratio* de exposición, consideramos un estudio de casos y controles que analiza la relación entre el consumo de alcohol y el cáncer bucal.
- La edad, un posible factor de confusión, ha sido categorizada en tres grupos.
- El cálculo crudo de 2,8 indica tres veces el exceso de riesgo de cáncer bucal, mientras que el cálculo ajustado de 2,1 indica dos veces el exceso de riesgo, para los que consumen alcohol con respecto a los que no lo hacen.
- Los resultados indican que hay confusión debido a la edad.

Auto-evaluación (P10.8)

Los siguientes datos provienen de una encuesta transversal de seroprevalencia del VIH entre prostitutas con relación al uso de drogas intravenosas. El *odds ratio* crudo de prevalencia es 3,59.



	Negros o Hispanos		Blancos	
	uso drogas intravenosas	no uso drogas intravenosas	uso drogas intravenosas	no uso drogas intravenosas
VIH +	31	12	16	3
VIH -	93	144	141	124

1. ¿Cuál es el valor del POR para el grupo de negros o hispanos? . . . **???**
2. ¿Cuál es el valor del POR para los blancos? . . . . . **???**
3. ¿Qué cuadro cree que debería recibir más peso cuando se calcula un *odds ratio* ajustado? **???**

Opciones
3,25 **3,59** **4,00** **4,31** **4,69** **negros o hispanos** **blancos**

En el estudio descrito en las preguntas anteriores, el POR calculado para el grupo de negros o hispanos fue 4,00 y el POR calculado para los blancos fue 4,69. El POR ajustado "basado en la precisión" para este estudio es 4,16. Recuerde que el POR crudo fue 3,59.

4. ¿Hay confusión? . . . . . **???**

Opciones
1,00 **4,35** **4,97** **quizás** **no** **sí**

10-4 Confusión, interacción y modificación de efecto

Confusión versus interacción

Otro motivo para controlar las variables en un estudio epidemiológico es evaluar la interacción. Para hacerlo, necesitamos determinar si el cálculo de la medida de efecto difiere a niveles diferentes de la variable de control. Considere los resultados de un estudio de casos y controles que evalúa la relación potencial entre el consumo de alcohol y el cáncer de vejiga. Estos datos están estratificados en tres categorías por raza.

Estudio casos y controles

Alcohol ? **Cáncer de vejiga**

	Blancos		**Negros**		**Asiáticos**		**Combinados**	
	Alc	no Alc	Alc	no Alc	Alc	no Alc	Alc	no Alc
Caso	72	41	93	54	68	33	233	128
Control	106	105	113	113	78	142	297	360
	$\hat{OR}$ = 1.74		$\hat{OR}$ = 1.72		$\hat{OR}$ = 3.75		$\hat{cOR}$ =	**2.21**

Los *odds ratios* calculados para estos tres estratos de raza y los estratos combinados dan 1,74; 1,72; 3,75 y 1,73, respectivamente. Hay evidencia clara de que existe interacción aquí, ya que el efecto es fuerte en los asiáticos, pero menor en los blancos y negros. No queda claro si hay confusión, ya que el valor del cálculo ajustado podría variar entre 1,72 y 3,75, según las ponderaciones asignadas a los estratos. El *odds ratio* ajustado basado en la precisión da 2,16, que no es muy distinto del *odds ratio* crudo de 1,73; esto sugiere que hay poca evidencia de confusión.

Confusión e interacción son conceptos distintos. La confusión compara los efectos calculados antes y después del control, mientras que la interacción compara los efectos calculados después del control. Cuando se evalúa la confusión y la interacción en el mismo estudio, es posible hallar una con o sin la otra.

Considere el siguiente cuadro que brinda las razones de riesgos crudas y específicas para cada estrato de varios grupos de datos hipotéticos en los cuales se está considerando una variable dicotómica de control. Para cada grupo de datos en este cuadro, ¿cree que hay interacción o confusión? Piénselo por unos minutos y luego observe las respuestas a continuación.

Confusión versus interacción

Grupo de datos	Estrato 1 RR	Estrato 2 RR	Crudo RR	Interacción?	Confusión?
1	1.02	3.50	6.00		
2	1.02	3.50	2.00		
3	0.03	3.50	1.70		
4	1.00	1.00	4.10		
5	4.00	4.10	4.20		

Veamos los conjuntos de datos, uno por vez. Para el grupo de datos 1, claramente hay interacción, porque el cálculo para el estrato 1 indica que no hay asociación, pero el cálculo para el estrato 2 indica una asociación razonablemente fuerte. Claramente hay confusión, porque cualquier promedio ponderado de los valores 1,02 y 3,50 será relevantemente distinto del cálculo crudo de 6,0.

Nuevamente, para el conjunto de datos 2 claramente hay interacción, como en el conjunto de datos 1. Sin embargo, no queda claro si hay confusión o no. El valor del cálculo ajustado dependerá de las ponderaciones asignadas a cada estrato. Si se da todo el peso al estrato 1 o al estrato 2, entonces el cálculo crudo de 2,0 diferirá considerablemente del cálculo ajustado, pero si se da igual peso a cada estrato, el cálculo ajustado será mucho más cercano al cálculo crudo. Sin embargo, el uso del cálculo ajustado aquí no es tan importante como la conclusión de que la asociación E→D es diferente para estratos distintos.

El conjunto de datos 3 también muestra interacción, aunque esta vez la naturaleza de la interacción es diferente de la que observamos en los conjuntos de datos 1 y 2. Aquí, los dos cálculos específicos para cada estrato

están en lados opuestos del valor nulo de 1 de la razón de riesgos. Parece que hay un efecto protector de la exposición sobre la enfermedad en el estrato 1, pero un efecto dañino de la exposición sobre la enfermedad en el estrato 2. En esta situación, la evaluación de la confusión es cuestionable y, potencialmente, muy engañosa, ya que el resultado importante aquí es el efecto de interacción, especialmente si esta fuerte interacción persiste después de realizar una prueba estadística para verificarla.

En el conjunto de datos 4, los dos cálculos específicos para cada estrato son exactamente iguales a 1, mostrando que no hay interacción. Sin embargo, hay evidencia clara de confusión, ya que el cálculo crudo de 4,0 es relevantemente distinto de ambos cálculos específicos para cada estrato.

En el conjunto de datos 5, no hay interacción porque los cálculos específicos para cada estrato son ambos iguales a 4. Tampoco hay confusión porque el cálculo crudo básicamente es igual a los dos cálculos específicos para cada estrato.

Grupo de datos	Estrato 1 RR	Estrato 2 RR	Crudo RR	Interacción?	Confusión?
1	1.02	3.50	6.00	sí	sí
2	1.02	3.50	2.00	sí	?
3	0.03	3.50	1.70	sí	??
4	1.00	1.00	4.10	no	sí
5	4.00	4.10	4.20	no	no

Resumen

- Confusión e interacción son conceptos distintos.
- La interacción considera lo que sucede después de que controlamos otra variable.
- Hay interacción si el cálculo de la medida de asociación es distinto a diferentes niveles de una variable que se está controlando.
- Cuando se evalúa la confusión y la interacción en el mismo estudio, es posible hallar una con o sin la otra.
- Cuando hay interacción fuerte, la evaluación de la confusión puede ser irrelevante o errónea.

Dos tipos de interacción: aditiva y multiplicativa

Considere el siguiente cuadro que brinda las frecuencias de celda para la categorización de tres variables dicotómicas **e**, **d** y **f**. Suponga que estos datos se derivan de un estudio de cohorte que calcula el riesgo.

	e = 1, f = 1	e = 1, f = 0	e = 0, f = 1	e = 0, f = 0
d = 1	a1	a2	b1	b2
d = 0	c1	c2	d1	d2

Observe que este cuadro representa, en forma equivalente, dos cuadros de 2 x 2 y describe la relación entre **e** y **d** después de estratificar **f**.

A partir de este cuadro, ahora definimos tres medidas de riesgo que utilizaremos para brindar fórmulas que distinguen la interacción aditiva de la multiplicativa.

R_{11} = riesgo para (e = l, f = 1) = al / (al + c1)

R_{10} = riesgo para (e = l, f = 0) = a2 / (a2 + c2)

R_{01} = riesgo para (e = 0, f = 1) = bl / (bl + dl)

R_{00} = riesgo para (e = 0, f = 0) = b2 / (b2 + d2) (riesgo basal)

Luego definimos la **ausencia de interacción a escala aditiva** si: **$R_{11} - R_{10} - R_{01} + R_{00} = 0$**
También definimos la **ausencia de interacción a escala multiplicativa** si: **$RR_{11} = RR_{10} \times RR_{01}$**

donde **$RR_{11} = R_{11} / R_{00}$**, **$RR_{10} = R_{10} / R_{00}$** y **$RR_{01} = R_{01} / R_{00}$**

Para ejemplificar estas dos definiciones, considere los siguientes datos de riesgo que tienen en cuenta la posible modificación de efecto de la composición genética (**f**) sobre la relación entre la presencia de aterosclerosis (**e**) y el mal de Alzheimer (**d**) en una población anciana en Rotterdam, Países Bajos.

R_{11} = riesgo para (e = l, f = 1) = 13,3%

R_{10} = riesgo para (e = 1, f = 0) = 4,4%

R_{01} = riesgo para (e = 0, f = 1) = 4,8%

R_{00} = riesgo para (e = 0, f = 0) = 3,4%

Para evaluar la interacción a escala aditiva, calculamos: **$R_{11} - R_{10} - R_{01} + R_{00} = 13,3 - 4,4 - 4,8 + 3,4 = 7,5$**

que difiere relevantemente de cero y, por lo tanto, indica un "alejamiento de la ausencia de interacción a escala aditiva".

Para evaluar la interacción a escala multiplicativa, evaluamos si:

$RR_{11} - (RR_{10} \times RR_{01}) = (13,3/3,4) - (4,4/3,4) \times (4,8/3,4) = 3,912 - 1,827 = 2,1$

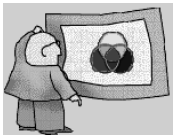
Dos tipos de interacción: aditiva y multiplicativa (continuación)

Este resultado también difiere relevantemente de cero y, por lo tanto, indica un "alejamiento de la ausencia de interacción a escala multiplicativa".

De la operación algebraica surge que "la ausencia de interacción a escala aditiva" es equivalente a obtener "iguales medidas de efecto de la diferencia de riesgo", cuando se compara el efecto **e-d** estratificado por **f**.

También surge de la operación algebraica que "la ausencia de interacción a escala multiplicativa" es equivalente a obtener "iguales medidas de efecto de la razón de riesgos", cuando se compara el efecto **e-d** estratificado por **f**.

Referencia: Kleinbaum et al., *Epidemiologic Research: Principles and Quantitative Methods*, Capítulo 19, John Wiley and Sons, 1982.



Auto-evaluación (P10.9)

1. Contrariamente a **???,** cuando hay interacción, los cálculos de **???** difieren a varios niveles de la variable de control.
2. Cuando se evalúa la confusión y la interacción en el mismo estudio, es **???** hallar una sin la otra.

Opciones
confusión la medida de efecto la modificación de efecto imposible
posible la precisión la discrepancia

Para los conjuntos de datos 1-3 del siguiente cuadro, seleccione la mejor de las siguientes respuestas:

A. confusión
B. interacción
C. ausencia de confusión o interacción
D. error de cálculo (imposible)

3. Conjunto de datos 1 . **???**
4. Conjunto de datos 2 . **???**
5. Conjunto de datos 3 . **???**

	Crudo OR	Estrato 1 OR	Estrato 2 OR	Ajustado OR
Conjunto de datos 1	4.0	1.0	6.0	4.0
Conjunto de datos 2	4.0	5.0	5.0	5.0
Conjunto de datos 3	4.0	4.0	4.0	3.0

Interacción versus modificación de efecto

Los términos **modificación de efecto** e **interacción** generalmente se utilizan en forma indistinta. Desde el punto de vista epidemiológico, empleamos la expresión "modificación de efecto" para resaltar que el efecto de la exposición en el resultado de salud se modifica según el valor de una o más variables de control.

Dichas variables de control se denominan **modificadores de efecto** de la relación entre la exposición y el resultado. Desde el punto de vista estadístico, utilizamos el término "interacción" para resaltar que la variable de exposición y la variable de control están interactuando de alguna forma dentro de un modelo matemático para determinar el resultado de salud. Para ejemplificar la modificación de efecto, considere los datos de casos y controles para evaluar la relación entre el consumo de alcohol y el cáncer de vejiga. Los datos mostraron evidencia clara de interacción, ya que el efecto calculado es mucho más fuerte en los asiáticos que en los negros o los blancos. Esta evidencia sugiere que la raza es un modificador de efecto de la relación entre el consumo de alcohol y el cáncer de vejiga. Dicha conclusión es respaldada por la literatura epidemiológica, que indica que los asiáticos metabolizan el alcohol de manera distinta a otros grupos raciales.

	Blancos		Negros		Asiáticos		Combinados	
	Alc	no Alc	Alc	no Alc	Alc	no Alc	Alc	no Alc
Caso	72	41	93	54	68	33	233	128
Control	106	105	113	113	78	142	297	360
	$\hat{OR} = 1.74$		$\hat{OR} = 1.72$		$\hat{OR} = 3.75$		$\hat{cOR}$	**=2.21**

La evaluación de la interacción o de la modificación de efecto normalmente se respalda con resultados de pruebas estadísticas. Recuerde que la confusión no involucra pruebas de significancia porque es un problema de validez. Sin embargo, las pruebas estadísticas para verificar la interacción se consideran adecuadas en los estudios epidemiológicos, porque la modificación de efecto conlleva la comprensión de los mecanismos causales subyacentes involucrados en la relación E→D, que no se considera un problema de validez. Una de las pruebas para datos estratificados que se ha incorporado a programas de computación disponibles se llama **prueba de Breslow-Day**.

Resumen

- Los términos modificación de efecto e interacción se utilizan muchas veces de manera indistinta.
- Si hay modificación de efecto, la variable o variables de control involucradas se llaman modificadores de efecto.
- La evaluación de la modificación de efecto normalmente se respalda mediante pruebas estadísticas para verificar la interacción significativa.
- La prueba estadística más conocida para verificar la interacción se llama prueba de **Breslow-Day**.

¿Hay realmente diferencia entre la modificación de efecto y la interacción?

Aunque los términos modificación de efecto e interacción se utilizan a menudo en forma indistinta, existe controversia en la literatura epidemiológica sobre sus definiciones precisas (véase Kleinbaum et al., *Epidemiologic Research: Principles and Quantitative Methods*, Capítulo 19, John Wiley and Sons, 1982).

Una distinción que se hace frecuentemente es que la modificación de efecto describe una característica clínica o biológica no cuantitativa de una población, mientras que la interacción normalmente es cuantitativa y específica de los datos y, en especial, depende de la escala sobre la que se miden las variables "que interactúan". Sin embargo, esta distinción conceptual muchas veces se pasa por alto en los estudios de investigación aplicados.

¿Por qué los epidemiólogos realizan pruebas estadísticas para verificar la interacción pero no la confusión?

Anteriormente (en la sección 10-1 de esta lección), hemos señalado que la evaluación de la confusión no debiera conllevar pruebas estadísticas, básicamente porque **la confusión es un problema de validez** que involucra algún error sistemático, en lugar de aleatorio. Además, si hay una diferencia relevante entre los cálculos de efecto crudo y ajustado, entonces se debe tomar una decisión con respecto a cuál de estos cálculos informar; en consecuencia, se debe utilizar el efecto ajustado, sin considerar una prueba estadística, porque controla las variables (es decir, los factores de riesgo) designadas para el ajuste.

Asimismo, no resulta obvio, ni siquiera si quisiéramos realizar una prueba estadística para verificar la confusión, cómo realizar exactamente una prueba para verificar la confusión en forma adecuada. Lo que se hace normalmente, aunque es incorrecto, es verificar si el factor potencial de confusión, por ejemplo, la edad, está significativamente asociado al resultado de salud, quizás controlando también la condición de exposición. Dicha prueba no evalúa realmente la confusión, ya que se refiere a un error aleatorio (es decir, la variabilidad) en lugar de considerar si los valores del cálculo crudo y ajustado son diferentes en los datos.

Con respecto a si se debería realizar una prueba estadística para evaluar la **interacción/modificación de efecto**, la respuesta no está tan definida. Si consideramos la interacción como una manifestación basada en los datos de un fenómeno basado en la población (es decir, una modificación de efecto), entonces puede justificarse una prueba estadística para explicar el error aleatorio asociado a un resultado basado en los datos. Además, a diferencia de la evaluación de la confusión, hay varios enfoques "legítimos" para verificar la interacción, uno de los cuales es la prueba de Breslow-Day para datos estratificados (descrita en la lección 14) y otro es una prueba para verificar la significancia de los términos de un producto en un modelo logístico.

Asimismo, puede sostenerse que la evaluación de la interacción/modificación de efecto no es un problema de validez, sino que se refiere a la comprensión/explicación conceptual de las relaciones entre las variables designadas para el control. Este último argumento, según la opinión de estos autores, es demasiado esotérico como para aceptarlo sin cuestionamientos. De hecho, se puede responder diciendo que la presencia de la interacción/modificación de efecto implica que los cálculos más "válidos" se obtienen mediante la estratificación de los modificadores de efecto, siempre que se pueda determinar qué variables son los "verdaderos" modificadores de efecto.

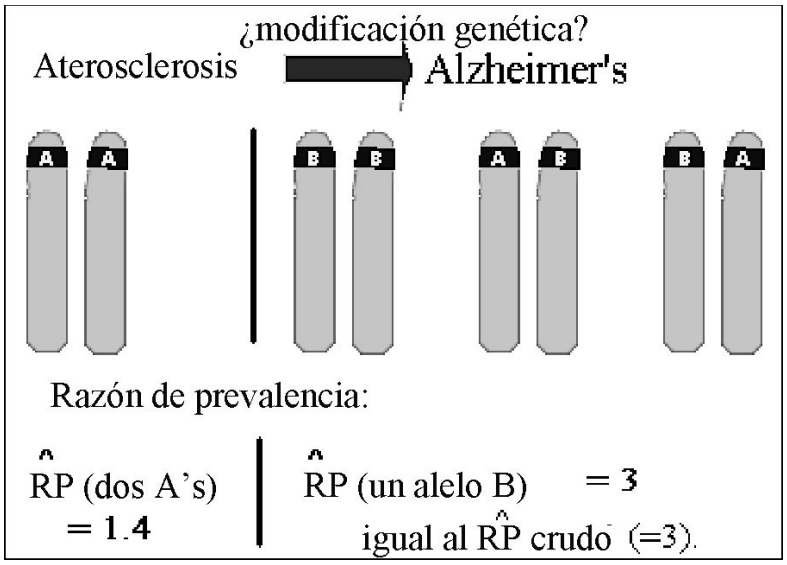
Al igual que con muchos problemas similares a éste que surgen al realizar investigaciones epidemiológicas, la mejor respuesta probablemente sea: "Depende". Es decir, depende del punto de vista del investigador si una prueba estadística de la interacción/modificación de efecto es adecuada. Sin embargo, estos autores tienden a ser de la opinión de que la modificación de efecto es un fenómeno de la población, que puede evaluarse mediante el empleo de pruebas estadísticas "legítimas" de la medida de la interacción basada en los datos.

Modificación de efecto: ejemplo

A principios de la década de los noventa, los investigadores del Estudio de Rotterdam realizaron tamizajes a 8000 ancianos y ancianas para detectar la presencia del mal de Alzheimer. Una de las preguntas de la investigación fue si la presencia de aterosclerosis aumentaba el riesgo de padecer la enfermedad. En un estudio transversal, los investigadores descubrieron que los pacientes con niveles altos de aterosclerosis tenían un riesgo tres veces mayor de padecer el mal de Alzheimer que los participantes con niveles bajos de aterosclerosis. Estos resultados sugerían una asociación entre la enfermedad cardiovascular y la enfermedad neurodegenerativa.

Los investigadores sabían, por investigaciones anteriores, que uno de los genes involucrados en el metabolismo lípido influye sobre el riesgo de padecer el mal de Alzheimer. Para este gen, hay dos formas alternativas, los alelos A y B. La composición genética de cada persona consta de dos de estos alelos. Las personas con, al menos, un alelo B corren mayor riesgo de padecer el mal de Alzheimer que quienes tienen dos alelos A.

Los investigadores plantearon la hipótesis de que la composición genética de un individuo podría modificar la asociación encontrada entre la aterosclerosis y el mal de Alzheimer. Por lo tanto, dividieron a la población de estudio en un grupo de participantes que tenían, al menos, un alelo B y un grupo de personas con dos alelos A. Obtuvieron los siguientes resultados: entre los que tenían al menos un alelo B, la prevalencia del mal de Alzheimer para los individuos con niveles altos de aterosclerosis se triplicó con respecto a la prevalencia de aquellos con niveles bajos. Este resultado es igual al crudo. Sin embargo, entre los que tenían dos alelos A, la prevalencia para los individuos con niveles altos de aterosclerosis fue sólo 1,4 veces mayor que la prevalencia de aquellos con niveles bajos.



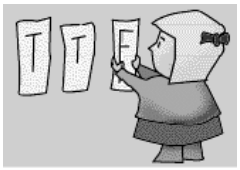
Estos resultados brindan un ejemplo de la modificación de efecto. La composición genética es un modificador de efecto. Los investigadores demostraron que la medida de la aterosclerosis está asociada al mal de Alzheimer, pero sólo en aquellos cuya composición genética los predispone a desarrollar esta enfermedad.

Preguntas de estudio (P10.10)

Prevalencia del mal de Alzheimer		
Alta ATE, 1 alelo B:	13.3%	$\hat{RP} = 3.0$
Baja ATE, 1 alelo B:	4.4%	
Alta ATE, dos alelos A:	4.8%	$\hat{RP} = 1.4$
Baja ATE, dos alelos A:	3.4%	

En este cuadro se enumeran la prevalencia del mal de Alzheimer y las razones de prevalencia para cada grupo de gen. Responda las siguientes preguntas suponiendo que la referencia es el grupo con baja aterosclerosis (ATE) y dos alelos A.

1. ¿Cuál es la razón de prevalencia comparando a quienes tienen alta ATE y dos alelos A con el grupo de referencia?
2. ¿Cuál es la razón de prevalencia comparando a quienes tienen baja ATE y, al menos, un alelo B con el grupo de referencia?
3. ¿Cuál es la razón de prevalencia comparando a quienes tienen alta ATE y, al menos, un alelo B con el grupo de referencia?
4. ¿Cuál es la diferencia entre las tres razones de prevalencia que acaba de calcular y las dos arriba enumeradas?



Auto-evaluación (P10.11)

Verdadero o **falso**.

1. El término "modificación de efecto" resalta que el efecto de la exposición en el resultado de salud se modifica según el valor de una o más variables de control. . . . . **???**
2. Hay evidencia de que existe modificación de efecto cuando las medidas de asociación específicas para cada estrato son aproximadamente iguales. . . . . **???**
3. Esta evaluación puede respaldarse mediante una prueba estadística llamada prueba de Breslow-Day. . . . . **???**

Una vacuna contra el sarampión puede ser muy efectiva para prevenir la enfermedad si se aplica a un niño después de que haya cumplido los 15 meses de edad, pero es menos efectiva si se suministra antes de cumplir 15 meses.

4. Este ejemplo ilustra **???**, donde la exposición es **???**, el resultado es **???** y el modificador de efecto es **???**.

Opciones
la edad al momento de la vacunación **la confusión** **la modificación de efecto** **el sarampión** **la vacuna contra el sarampión**

Nomenclatura

$a\hat{EOR}$	Cálculo del *odds ratio* de exposición ajustado
$a\hat{POR}$	Cálculo del *odds ratio* de prevalencia ajustado
$a\hat{RP}$	Cálculo de la razón de prevalencia
$a\hat{RR}$	Cálculo de una razón de riesgos ajustada
C	Variable de confusión
$c\hat{EOR}$	Cálculo del *odds ratio* crudo de exposición
$c\hat{POR}$	Cálculo del *odds ratio* de prevalencia
$c\hat{RP}$	Cálculo de la razón de prevalencia
$c\hat{RR}$	Cálculo de la razón de riesgos cruda
D	Enfermedad
E	Exposición
V	Variable interviniente
w o w_i	Ponderación; con un subíndice i, denota la ponderación para un estrato.

Referencias

Hofman A, Ott A, Breteler MM, Bots ML, Slooter AJ, van Harksamp F, van Duijn CN, Van Broeckhoven C, Grobbee DE. *Atherosclerosis, apolipoprotein E, and prevalence of dementia and Alzheimer's disease in the Rotterdam Study*. Lancet 1997; 349 (9046):151-4.

Greenland S, Morgenstern H. *Confounding in health research*. Annu Rev Public Health 2001; 22 189-212.

Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic Research: Principles and Quantitative Methods*. John Wiley and Sons Publishers, Nueva York, 1982.

Miettinen O. *Confounding and effect modification*. Am J Epidemiol 1974; 100(5):350-3.

Mundt KA, Shy CM. Interaction: *An epidemiological perspective for risk assessment*. En: Fan AM, Chang LW (eds.). Toxicology and risk assessment: Principles, methods, and applications. Marcel Dekker, Inc., Nueva York, NY, pág. 329-351, 1996

Whittemore AS. *Collapsibility of multidimensional contingency tables*. J R Stat Soc B 1978; 40:328-40.

Ejercicios
EPI-1. Confusión: tabaquismo

Se realizó un estudio de casos y controles para evaluar si el consumo de café (alto vs. bajo) está asociado a la enfermedad de úlcera péptica. A continuación se resumen los resultados, estratificados por la condición de tabaquismo.

	No fumadores	
	Alto consumo de café	Bajo consumo de café
Úlcera	40	65
Sin úlcera	45	150

	Fumadores	
	Alto consumo de café	Bajo consumo de café
Úlcera	150	45
Sin úlcera	65	40

a. Calcule los *odds ratios* específicos para cada estrato y el *odds ratio* crudo. Incluya el desarrollo de sus cálculos.
b. Basándose en estos cálculos, ¿la condición de tabaquismo es un factor de confusión en estos datos? Justifique su respuesta.
c. ¿Hay evidencia de interacción?
d. Evalúe si

    i. fumar está relacionado con el consumo de café entre los que no padecen úlcera péptica.
    ii. fumar está relacionado con la úlcera péptica entre los individuos con bajo consumo de café.

e. ¿Qué indican sus respuestas en el ítem d) con respecto a si fumar es un factor de confusión en este estudio?
f. ¿Su conclusión en el ítem e) sobre si la condición de tabaquismo es un factor de confusión es la misma que en el ítem a)?

EPI-2. Confusión e interacción

Debido a la presencia inusualmente alta de síndrome de ansiedad endógena entre los estudiantes de pregrado de la Universidad X, se realizó un estudio de casos y controles para determinar la conveniencia de hacer un curso de Introducción a la estadística (variable de exposición de 0,1, donde 1 = expuestos y 0 = no expuestos) podría ser una de las causas. El estudio incluyó a 57 estudiantes a los que se les diagnosticó el síndrome de ansiedad, y estos "casos" se compararon con una muestra de 750 controles "normales". Considere los siguientes cuadros de 2 x 2 que estratifican los antecedentes previos de trastorno mental (TM).

	TM = 1	
	E = 1	E = 0
D = 1	25	15
D = 0	50	50

	TM = 0	
	E = 1	E = 0
D = 1	2	15
D = 0	50	600

a. ¿Cuál es la medida de efecto calculada que describe la relación E-D que ignora el control de la variable TM (es decir, cuál es el efecto crudo calculado para estos datos)? Al responder esta pregunta, incluya el desarrollo de sus cálculos.
b. ¿Cuáles son las medidas de efecto calculadas para cada grupo con TM? Nuevamente, incluya el desarrollo de sus cálculos.
c. Basándose en sus cálculos para alguna de las preguntas anteriores o ambas, ¿debería controlarse el TM porque hay interacción significativa? Desarrolle su respuesta.
d. Basándose en sus cálculos para los ítems a) y b), ¿debería controlarse la variable TM porque es un factor de confusión? (Suponga que la variable TM ya satisface las condiciones *a priori* para la confusión). Desarrolle su respuesta con la información y lógica adecuada.
e. ¿Cuál es el valor del *odds ratio* que describe la asociación del TM con la exposición? ¿Qué dice este *odds ratio* sobre la distribución de antecedentes previos de trastorno mental (es decir, TM) cuando se compara a los estudiantes que estudian Introducción a la estadística con aquellos que no cursan esta materia? ¿Este resultado respalda la conclusión de que el TM es un factor de confusión en estos datos? [Sugerencia: para responder esta pregunta, debe considerar la siguiente forma alternativa de definir la confusión: si la medida de efecto es un *odds ratio*, una covariable (por ejemplo, TM) es un factor de confusión, siempre que esté asociada tanto a la exposición (es decir, E) como a la enfermedad (es decir, D) en forma separada, es decir, $OR_{TM,E|D=0}$ es relevantemente distinto de 1 y $OR_{TM,\ D|E=0}$ es relevantemente distinto de 1].

EPI-3. Exposición y factores potenciales de confusión

Se realizó un estudio de cohorte para evaluar la relación entre el uso de la terapia de reemplazo de estrógenos y la mortalidad por cáncer de ovario entre mujeres perimenopáusicas y postmenopáusicas. Se brindó la siguiente información que relaciona la variable de exposición con varios factores de confusión posibles.

Distribución de los factores potenciales de riesgo de padecer cáncer de ovario y su asociación con el uso de la terapia de reemplazo de estrógenos (TRE).

Factor potencial de riesgo	*Odds ratio*	Intervalo de confianza de 95%
Edad al momento de la menarquia		
< 12 años	1,00	Referente
12 años	0,98	0,95-1,00
13 años	0,94	0,83-1,05
Cantidad de nacimientos vivos		
0	1,00	Referente
1	1,06	1,01-1,10
2-3	1,08	1,05-1,11
4	0,86	0,82-0,89

Educación		
< Secundario	1,00	Referente
Secundario	1,92	1,89-1,95
Universitario	3,48	2,26-5,59

Indique cuál de las siguientes afirmaciones es VERDADERA. [Elija sólo la mejor respuesta].

i. No es probable que la edad al momento de la menarquia sea un fuerte factor de confusión de la relación entre la TRE y el cáncer de ovario.
ii. No es probable que el número de nacimientos vivos sea un fuerte factor de confusión de la relación entre la TRE y el cáncer de ovario.
iii. La educación es, definitivamente, un factor de confusión de la relación entre la TRE y el cáncer de ovario.

[Sugerencia: para responder esta pregunta, debe considerar la siguiente forma alternativa de definir la confusión: si la medida de efecto es una razón de riesgos, una covariable (por ejemplo, ED) es un factor de confusión, siempre que esté asociada tanto a la exposición (es decir, E) como a la enfermedad (es decir, D) en forma separada, es decir, $OR_{ED,E}$ es relevantemente distinto de 1 y $OR_{ED,\ D|E=0}$ es relevantemente distinto de 1].

EPI-4. Tipo de error

Para cada una de las situaciones que se describen a continuación, indique el tipo de error que podría presentarse con mayor probabilidad. Cada una de las opciones numeradas puede utilizarse una vez, más de una vez o no utilizarse.

i. Sesgo de selección
ii. Clasificación errónea no diferencial
iii. Clasificación errónea diferencial
iv. Confusión
v. Falacia ecológica
vi. Error aleatorio

a. En un estudio de cohorte sobre la terapia de reemplazo de hormonas (TRH) y el riesgo de padecer aterosclerosis de las arterias coronarias (AAC), el nivel alto de ingresos está asociado tanto al uso de la TRH como al riesgo de padecer AAC.
b. En un estudio de casos y controles sobre la relación entre sucesos estresantes de la vida diaria y ataques de asma, los casos tienen más probabilidad de auto-declarar un nivel de estrés sobrestimado en comparación a los controles.
c. En un estudio de cohorte sobre el uso de terminales de video (VDT) y el riesgo de padecer síndrome del túnel carpiano, es más difícil rastrear a los usuarios de VDT que a los no usuarios; en consecuencia, hay mayor pérdida de seguimiento de los usuarios de VDT.
d. En un estudio de casos y controles sobre el betacaroteno y el riesgo de padecer cáncer de esófago, se comparan, entre los casos y controles, muestras de sangre congeladas y almacenadas 20 años atrás. Más tarde se descubre que las muestras se deterioraron durante el almacenamiento.

EPI-5. Validez externa

¿Cuáles de los siguientes enfoques para controlar la confusión es más probable que afecte la validez externa de un estudio?

a. Aleatorización
b. Restricción
c. Análisis estratificado
d. Análisis de regresión

EPI-6. Categorización de la exposición

Se realizó un estudio de casos y controles con base hospitalaria para determinar si la exposición residencial a los campos magnéticos estaba asociada al desarrollo de linfoma. La variable de exposición (la fuerza del campo magnético en el hogar) se midió originalmente como una variable continua, pero fue categorizada para propósitos del análisis de datos. Como se muestra en el cuadro a continuación, los investigadores presentaron los resultados para tres planes de categorización distintos:

Fuerza del campo	**OR crudo**	**IC 95%**	**OR ajustado***	**IC 95%**
Plan 1				
Baja	1,0 (referente)		1,0 (referente)	
Alta	1,59	0,87-2,93	1,77	0,91-3,41
Plan 2				
Baja	1,0 (referente)		1,0 (referente)	
Media	1,07	0,68-1,99	1,20	0,62-2,32
Alta	1,64	0,84-3,20	2,06	0,96-4,42
Plan 3				
Muy baja	1,0 (referente)		1,0 (referente)	
Baja	1,22	0,49-3,02	1,46	0,53-4,04
Media	1,01	0,50-2,04	1,19	0,56-2,64
Alta	1,24	0,61-2,56	1,48	0,66-3,32
Muy alta	6,81	1,63-28,5	13,43	1,76-102,7

- Ajustado por edad, sexo, condición socioeconómica, área geográfica y años de residencia en el hogar.

a. ¿Qué plan de categorización arroja resultados que son más coherentes con un efecto dosis-respuesta?

_____ Plan 1 _____ Plan 2 _____ Plan 3

b. Indique si cada una de las siguientes afirmaciones es verdadera o falsa.

_____ El sesgo debido a la confusión (por los factores enumerados debajo del cuadro) se alejó consistentemente del valor nulo.

_____ El plan 3 de categorización de la exposición es menos susceptible a la confusión residual, ya que divide a los individuos del estudio en la mayor cantidad de categorías.

_____ Entre los individuos expuestos a campos magnéticos muy altos en el hogar, el ajuste de los factores de confusión produjo un aumento de la validez, pero una disminución de la precisión.

EPI-7. Confusión

La literatura ha indicado que la confusión está compuesta tanto de criterios basados en los datos como de criterios *a priori* (con respecto a los datos).

a. Exponga los criterios para la confusión basados en los datos.
b. Brinde un ejemplo numérico de la confusión basada en los datos que haga que el sesgo se aleje del valor nulo. (Nota: puede responder esta pregunta sin especificar completamente todas las frecuencias de celda en los cuadros de 2 x 2 utilizados para su ejemplo).
c. Brinde un ejemplo numérico de la confusión basada en los datos que haga que el sesgo sea hacia el valor nulo. (Nota: puede responder esta pregunta sin especificar completamente todas las frecuencias de celda en los cuadros de 2 x 2 utilizados para su ejemplo).
d. Cuando evalúa la interacción, ¿debería hacer una prueba para determinar si hay interacción significativa? ¿Por qué sí o por qué no?
e. Cuando evalúa la confusión, ¿debería hacer una prueba para determinar si hay confusión? ¿Por qué sí o por qué no?

Tres condiciones *a priori* que se han definido para la confusión son las siguientes:

I. El factor potencial de confusión debe ser un factor de riesgo para la enfermedad.

II. El factor potencial de confusión no puede ser una variable interviniente entre la exposición y la enfermedad.

III. El factor potencial de confusión debe estar asociado a la exposición en la población de origen que se está estudiando (nota: esta condición es la más controvertida).

f. Brinde un ejemplo de factor potencial de confusión que sea un factor de riesgo para alguna enfermedad.
g. Brinde un ejemplo de variable interviniente entre una exposición y una enfermedad.
h. Brinde un ejemplo de factor de riesgo para una enfermedad que no esté asociado con la exposición en la población de origen en estudio.
i. ¿Controlaría un factor de riesgo que no es una variable interviniente y que no satisface la condición *a priori* III, pero que, sin embargo, está asociado tanto a la exposición como a la enfermedad en los datos? Desarrolle su respuesta.
j. Para cada una de las condiciones I-III anteriores, brinde una justificación (ya sea conceptual o numérica) de la condición, si está de acuerdo con que la condición es adecuada.

Si, por otra parte, cree que una condición no es adecuada (o necesita mayor calificación), explique (ya sea conceptual o numéricamente) por qué cree que no es adecuada.

EPI-8. Modificación de efecto y confusión

Se llevó a cabo un estudio de cohorte sobre la actividad física (AF) y la incidencia de la diabetes durante un período de seis años entre hombres japoneses-estadounidenses en Honolulu. A continuación se resumen los datos estratificados de ese estudio, según el índice de masa corporal (IMC).

	IMC alto	
	Alta AF	**Baja AF**
Diabetes	48	62
Personas-años	1050	1067

	IMC bajo	
	Alta AF	**Baja AF**
Diabetes	54	71
Personas-años	1132	1134

a. Basándose en estos datos, ¿cuál es la razón de tasas observada (RDI) para alguien con IMC alta y actividad física alta (es decir, con relación a la tasa basal)?
b. ¿Hay evidencia de modificación de efecto por el IMC? (Incluya el desarrollo de los cálculos y justifique su respuesta).
c. ¿Hay evidencia de confusión por el IMC en estos datos? (Incluya el desarrollo de los cálculos y justifique su respuesta).

EPI-9. Interacción: epidemiólogos y dolores de cabeza

Los siguientes datos provienen de un estudio de cohorte con incidencia acumulada que tiene en cuenta la exposición a manuscritos escritos por un cierto epidemiólogo y los dolores de cabeza (d) entre estudiantes de postgrado de epidemiología. Los datos están estratificados por el grado de experiencia previa en idiomas extranjeros (f).

f = 1 (sí)

	e = 1	e = 0
d = 1	25	20
d = 0	975	980
	1000	1000

f = 0 (no)

	e = 1	e = 0
d = 1	10	5
d = 0	990	995
	1000	1000

a. Analice si hay interacción cuando se comparan las razones de riesgos y cuando se comparan las diferencias de riesgos entre las dos categorías de la variable f.
b. Complete el siguiente cuadro:

	e = 1, f = 1	e = 1, f = 0	e = 0, f = 1	e = 0, f = 0
d = 1				
d = 0				

c. Calcule lo siguiente:

R_{11} = riesgo para (e = 1, f = 1) =

R_{10} = riesgo para (e = 1, f = 0) =

R_{01} = riesgo para (e = 0, f = 1) =

R_{00} = riesgo para (e = 0, f = 0) = (riesgo basal)

d. Decimos que hay "ausencia de interacción a escala aditiva" si

$$R_{11} - R_{10} - R_{01} + R_{00} = 0$$

Basándose en los datos, ¿hay interacción a escala aditiva?

e. Decimos que hay "ausencia de interacción a escala multiplicativa" si

$RR_{11} = RR_{10} \times RR_{01}$ donde

$RR_{11} = R_{11} / R_{00}$ $RR_{10} = R_{10} / R_{00}$ y $RR_{01} = R_{01} / R_{00}$

Basándose en los datos, ¿hay interacción a escala multiplicativa?

f. Sobre la base de estos resultados, ¿cuál de las siguientes acciones sería adecuada como parte de su plan analítico general? (Puede elegir más de una opción).

1. Informar RR específicas para cada estrato (riesgos relativos).
2. Informar DR (diferencias de riesgo) específicas para cada estrato.
3. Hacer un cálculo general de resumen de la RR.
4. Hacer un cálculo general de resumen de la DR.
5. Calcular la RR cruda y evaluar la confusión.
6. Calcular la DR cruda y evaluar la confusión.

EPI-10. Sesgo de clasificación errónea

El siguiente cuadro resume los datos de un estudio de casos y controles.

	C = 1		C = 0	
	E = 1	E = 0	E = 1	E = 0
Casos	120	100	45	200
Controles	80	200	30	400

E = condición de exposición C = factor potencial de confusión/modificador de efecto

a. Suponga que se clasificó correctamente a todos los individuos en el estudio y que no hay sesgo debido a la selección. ¿Hay evidencia de modificación de efecto y/o confusión debido a C en estos datos? Justifique su respuesta.
b. Ahora suponga que los datos anteriores observados no habían sido clasificados correctamente y que, además, la exposición se clasificó erróneamente con una sensibilidad de 0,8 y una especificidad de 0,75. La clasificación errónea fue no diferencial tanto con respecto a la condición del resultado como de la categoría de C.
c. ¿Hay evidencia de sesgo en los datos clasificados erróneamente? De ser así, ¿puede corregirlo? Justifique su respuesta.
d. ¿Hay evidencia de modificación de efecto en los datos clasificados erróneamente? Justifique su respuesta.
e. ¿Cómo afecta la prevalencia de la exposición a la magnitud del sesgo debido a la clasificación errónea en este ejemplo? Desarrolle su respuesta.

Respuestas a preguntas de estudio y auto-evaluaciones

P10.1

1. La RR calculada de 1,3 para los fumadores indica que hay un riesgo 30% mayor para los expuestos al TCX, en comparación con los no expuestos. Algunos investigadores podrían considerar que esta cantidad de aumento indica una asociación moderada, en lugar de una falta de asociación.

P10.2

1. distorsión, medida de asociación, controlamos, factores de riesgo
2. 1,0
3. 1,0
4. sí
5. sí; la evaluación de la confusión basada en los datos se realiza determinando si el cálculo crudo de la medida de asociación es relevantemente distinto de la medida de asociación ajustada.

P10.3

1. 0,69
2. 0,65
3. 0,61
4. sí; la evaluación de la confusión basada en los datos se realiza determinando si el cálculo crudo de la medida de asociación es relevantemente distinto de la medida de asociación ajustada.
5. 0,65; 0,89. Para determinar si hay una diferencia importante entre los valores del cálculo crudo y ajustado sobre la base de un cambio en un porcentaje especificado y requerido entre los valores del cálculo crudo y ajustado, multiplique el cálculo crudo por el cambio del porcentaje especificado y luego sume y reste ese valor del cálculo crudo. Si el intervalo obtenido contiene el cálculo ajustado, entonces no hay una diferencia importante.
6. sí
7. no

P10.4

1. Verdadera
2. Falsa. El cálculo ajustado puede ser mayor o menor que el cálculo crudo correspondiente.
3. Falsa. La mayoría de los epidemiólogos prefieren asignar distinto peso a cada cálculo específico para cada estrato. En general, las ponderaciones se determinan sobre la base del tamaño o precisión de la muestra.
4. Falsa. Debido a que la confusión es un problema de validez, no debería evaluarse mediante una prueba estadística, sino, más bien, buscando una diferencia relevante entre los valores del cálculo crudo y ajustado.
5. 1,08
6. 2,17

P10.5

1. No

P10.6

1. se satisface
2. se satisface
3. no se satisface; la edad no puede estar asociada al tipo sanguíneo.
4. no; la edad no puede estar asociada al tipo sanguíneo.

P10.7

1. C; el *odds ratio* crudo que relaciona el alcohol con el cáncer bucal se calcula mediante la fórmula (a x d)/(b x c), que en este caso es igual a 2,8.
2. A; el *odds ratio* específico para cada estrato que relaciona el alcohol con el cáncer bucal puede calcularse para cada estrato mediante la fórmula (a x d)/(b x c). Los *odds ratios* específicos para cada estrato son 2,2 para el grupo de 40 a 49 años; 2,2 para el grupo de 50 a 59 años y 2,0 para el grupo de 60 años o más.
3. Sí. Hay confusión basada en los datos porque el *odds ratio* crudo de 2,8 es relevantemente distinto de cualquier promedio ponderado de los *odds ratios* específicos para cada estrato, los cuales son aproximadamente iguales a 2.
4. La edad es un factor de confusión porque se supone que se satisfacen las tres condiciones *a*

priori para la confusión y también se satisface el criterio basado en los datos.

P10.8

1. 4,00
2. 4,69
3. Negros o hispanos. Puede pensar que el cuadro para el grupo de blancos debe recibir un peso mayor porque tiene un tamaño de muestra levemente más grande; sin embargo, el cuadro para el grupo de negros o hispanos, en verdad, está más equilibrado. Véase la lección 14 para una explicación más completa sobre datos equilibrados.
4. quizás; depende de si el investigador considera que la diferencia entre 3,59 y 4,16 es importante.

P10.9

1. confusión, la medida de efecto
2. posible
3. B
4. A
5. D. Conjunto de datos 3: recuerde que el cálculo ajustado es un promedio ponderado de los dos cálculos específicos para cada estrato y, por lo tanto, debe ser un valor entre ellos.

P10.10

1. 4,8% / 3,4% = 1,4
2. 4,4% / 3,4% = 1,3
3. 13,3% / 3,4% = 3,9
4. Las dos RP anteriores son RP específicas para cada estrato, para cada uno de los dos grupos de genes. Las tres calculadas aquí utilizan un grupo como referencia y comparan las otras tres con ese grupo. En este ejemplo, tener baja ATE y 2 alelos A es el grupo de referencia comparado con aquellos que tienen uno o ambos factores de riesgo (alta ATE, 1 alelo B). Para ver cómo pueden utilizarse estas 3 RP para definir dos tipos diferentes de interacción, véase el primer suplemento en la sección 10-4 de esta lección o el recuadro titulado **Dos tipos de interacción: aditiva y multiplicativa**.

P10.11

1. Verdadero
2. Falso. Hay evidencia de que existe modificación de efecto cuando las medidas de asociación específicas para cada estrato son diferentes.
3. Verdadero
4. la modificación de efecto, la vacuna contra el sarampión, el sarampión, la edad al momento de la vacunación

LECCIÓN 11

CONFUSIÓN QUE INVOLUCRA VARIOS FACTORES DE RIESGO

11-1 Confusión que involucra varios factores de riesgo

*Esta lección considera cómo la evaluación de **confusión** se torna un poco más complicada cuando se controla más de un factor de riesgo. En especial, cuando se controlan varios **factores de riesgo**, podemos ver que considerar simultáneamente todos los factores de riesgo puede llevar a una conclusión distinta que cuando éstos se analizan en forma separada. Anteriormente (en la lección 10), sostuvimos que la evaluación de confusión no es adecuada para las variables que son **modificadores del efecto** de la relación entre la exposición y la enfermedad en estudio. En consecuencia, a lo largo de esta lección, nuestra discusión sobre confusión supondrá que **ninguna** de las variables consideradas para el control son modificadores de efecto (es decir, no hay interacción entre la exposición y ninguna de las variables que se están controlando).*

Evaluación de la confusión cuando hay interacción

Hemos limitado nuestra discusión a la confusión que involucra varias variables en la situación donde **ninguna** de las variables consideradas para el control es un modificador del efecto de la relación entre la exposición y la enfermedad en estudio. Esta restricción se definió principalmente por razones pedagógicas, ya que es más fácil analizar la confusión entre varias variables cuando no hay modificación del efecto.

Sin embargo, **a menudo es bastante adecuado considerar la confusión, incluso cuando hay interacción**. Por ejemplo, si sólo estamos controlando una variable, por ejemplo, sexo, y descubrimos que el *odds ratio* para los hombres es 1,3, mientras que el *odds ratio* para las mujeres es 3,6 y el *odds ratio* crudo es 10,1, entonces hay evidencia de confusión y de interacción , y ambas pueden ser abordadas en un análisis. Se puede presentar una situación similar cuando se están controlando dos o más variables.

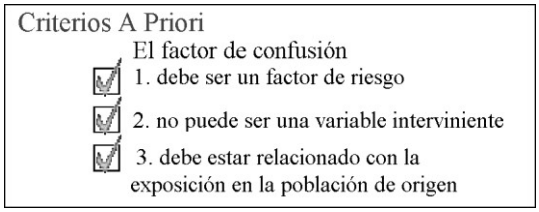
Asimismo, cuando se controlan diversas variables y hay interacción de, por ejemplo, sólo una de estas variables con la variable de exposición, entonces las demás variables consideradas para el control pueden evaluarse como factores potenciales de confusión. Por ejemplo, si en un estudio de cohorte sobre los factores de riesgo de la enfermedad coronaria (EC) se determinase que el nivel de colesterol (COL) es el único modificador del efecto de la variable de exposición (por ejemplo, el nivel de actividad física) entre los factores de riesgo que incluyen edad, tabaquismo, sexo y presión arterial, entonces estas últimas variables pueden evaluarse como posibles factores confusores.

En esta última situación, un método para realizar la evaluación de confusión involucra **estratificar según el modificador del efecto** (COL) y evaluar la confusión usando las otras variables en forma separada dentro de las diferentes categorías de COL.

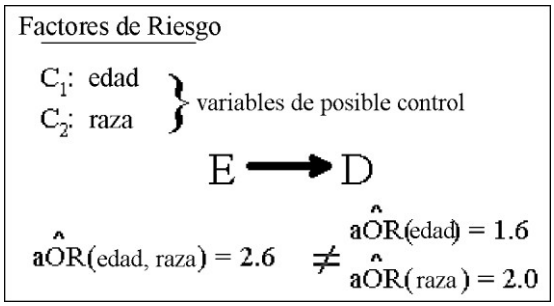
Otro abordaje es utilizar un **modelo matemático** (por ejemplo, mediante la regresión logística) que contenga todos los factores de riesgo considerados como los principales efectos, así como un parámetro del producto de la exposición y el colesterol (E x C). Esos factores de riesgo distintos del COL se pueden a su vez evaluar para analizar si hay confusión, siempre que el efecto principal del colesterol , la variable de exposición, y el producto de la exposición y el COL, continúen en el modelo en todos los pasos de la evaluación.

Dos principios importantes

Hasta ahora, hemos considerado sólo el control de un único factor de confusión en un estudio epidemiológico. Pero, generalmente, se identifican y miden varios factores de riesgo para un posible control. Recuerde los criterios *a priori* para la confusión. Cuando varios factores cumplen estos criterios, ¿cómo determinamos cuál debemos controlar en el análisis?

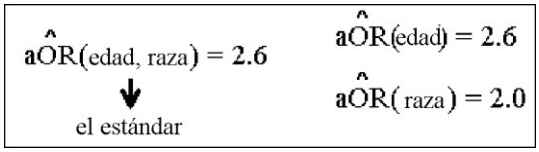


Suponga que la edad y la raza son dos factores de riesgo identificados y medidos para un posible control en un estudio de casos y controles cuyo objetivo es evaluar la relación entre una exposición y una enfermedad. Es posible que el *odds ratio* ajustado, que controla simultáneamente tanto la edad como la raza, arroje resultados diferentes de los obtenidos cuando se controla cada variable en forma separada.



Si nuestro estándar es un *odds ratio* que controla todos los factores potenciales de riesgo, entonces, ¿**siempre** deberíamos controlar **todos** los factores de riesgo? No necesariamente. Es posible que sólo se necesite controlar un subconjunto de estos factores para lograr resultados válidos.

Suponga que estos resultados se obtuvieron de nuestro estudio de casos y controles:



Aquí, el *odds ratio* que controla sólo la edad es equivalente al *odds ratio* que controla tanto la edad como la raza. En este caso, no perderíamos nada en términos de validez si seleccionásemos sólo el control de la edad.

Estos ejemplos muestran dos principios fundamentales sobre el control de la confusión cuando se han identificado y medido varios factores de riesgo. En primer lugar, el control conjunto (o simultáneo) de dos o más variables puede arrojar resultados diferentes de los obtenidos mediante el control de cada variable en forma

separada. El cálculo ajustado (denotado aquí como **theta** con acento circunflejo, $\hat{\theta}$) que controla simultáneamente todos los factores de riesgo considerados, debe ser el estándar sobre el cual se deben basar todas las conclusiones sobre la confusión y la identificación de factores específicos de confusión.

En segundo lugar, es posible que no todas las variables en una lista dada de factores de riesgo deban controlarse, ya que es factible que diferentes subconjuntos de dichas variables puedan corregir la confusión. Analizaremos estos dos principios en las actividades siguientes.

Dos principios fundamentales sobre el control de la confusión

1. El control conjunto o simultáneo de dos o más variables puede arrojar resultados diferentes de los obtenidos mediante el control de cada variable en forma separada.

$$a\hat{\theta}\ (C_1, C_2, \ldots, C_p) \rightarrow \text{estándar}$$

2. Puede que no todas las variables en una lista dada de factores de riesgo deban controlarse.

Note: θ denota cualquier medidada de efecto de interés

Preguntas de estudio (P11.1)

Suponga que edad, raza, sexo y tabaquismo son los únicos factores de riesgo que quedarán bajo control al evaluar la relación entre una exposición y una enfermedad.

1. Describa el cálculo ajustado que debería ser el estándar sobre el cual se basen todas las conclusiones sobre confusión e identificación de factores específicos de confusión.
2. Si no se consideran todos los factores potenciales de confusión en forma simultánea, ¿qué problemas podrían surgir?
3. Si el cálculo ajustado considerado estándar de oro controla todos los factores de riesgo, ¿es posible que un subconjunto de dichos factores de riesgo pueda también controlar la confusión?
4. ¿Por qué el uso de dicho subconjunto de variables podría ser ventajoso, en comparación con el uso del cálculo ajustado que controla todos los factores potenciales de confusión?

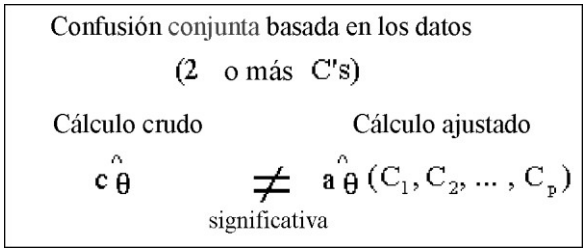
Resumen

- Hay dos principios fundamentales sobre el control de confusión cuando se han identificado y medido dos o más factores de riesgo de posible control.
  1. El control conjunto o simultáneo de dos o más variables puede arrojar resultados diferentes de los obtenidos mediante el control de cada variable en forma separada.
  2. Es posible que no todas las variables en una lista dada de factores de riesgo deban controlarse.

- Asimismo, según las relaciones entre estos factores de riesgo, es posible que la confusión se pueda corregir si se utilizan diferentes subconjuntos de factores de riesgo de la lista.

Confusión conjunta versus marginal

Ya dijimos que la **confusión basada en los datos** que involucra un único factor potencial de confusión significa que hay una diferencia relevante entre el cálculo crudo de efecto (que ignora completamente el factor potencial de confusión) y el **cálculo ajustado de efecto** (que controla el factor potencial de confusión). Ahora nos ocupamos de la **confusión conjunta basada en los datos** en presencia de dos o más factores potenciales de confusión. Esta situación ocurre cuando hay una diferencia relevante entre el cálculo crudo y el cálculo ajustado de efecto, que controla en forma simultánea toda la confusión potencial.



Preguntas de estudio (P11.2)

Suponga que se realizó un estudio de seguimiento para evaluar una relación E→D. Se determinó que la edad y la condición de tabaquismo eran posibles variables de control. Suponga además que:

aRR(edad, tabaquismo)	=	2,4
aRR(edad)	=	1,7
aRR(tabaquismo)	=	1,9
cRR	=	1,5

1. ¿Hay evidencia de confusión conjunta? ¿Por qué sí o por qué no?

Suponga que un estudio de seguimiento distinto fue realizado sobre la misma relación E→D en el que, nuevamente, edad y tabaquismo eran posibles variables de control. Suponga además que:

aRR(edad, tabaquismo)	=	1,4
aRR(edad)	=	2,4
aRR(tabaquismo)	=	2,4
cRR	=	1,5

2. ¿Hay evidencia de confusión conjunta? ¿Por qué sí o por qué no?

Por otra parte, decimos que la **confusión marginal basada en los datos** significa que hay una

diferencia relevante entre el cálculo crudo y el cálculo ajustado del efecto que controla sólo uno de varios factores potenciales de confusión.

Confusión marginal basada en los datos

$\hat{c\theta} \neq \hat{a\theta}(C_j)$

significativa

donde C_j es uno de varios factores potenciales de confusión

Preguntas de estudio (P11.2) (continuación)

Suponga que se realizó un estudio de seguimiento para evaluar una relación E→D. Se determinó que edad y tabaquismo eran posibles variables de control. Suponga que:

aRR(edad, tabaquismo)	=	2,4
cRR	=	1,5

3. ¿Hay evidencia de confusión marginal? ¿Por qué sí o por qué no?
4. Si aRR(edad) = 1,4, ¿ésto es evidencia de la existencia de confusión marginal?
5. ¿Ésto significa que no deberíamos controlar la edad como un factor de confusión?

La confusión conjunta es el principal criterio para determinar la presencia de confusión basada en los datos cuando todas las variables cumplen los requisitos para ser elegidas para el control. Sin embargo, la confusión marginal basada en los datos puede ayudar a determinar si **algunos** factores potenciales de confusión **no necesitan** ser controlados.

Preguntas de estudio (P11.2) (continuación)

6. En el estudio de seguimiento descrito en la pregunta de estudio anterior, se reportó:

aRR(edad, tabaquismo)	=	2,4
aRR(edad)	=	1,5
aRR(tabaquismo)	=	2,4
cRR	=	1,5

¿Ésto significa que no debemos controlar la edad?

7. ¿Qué problema podría haber en la práctica que impediría el cálculo del efecto que controla todos los factores de riesgo (por ejemplo, C_1, C_2, …, C_k)?
8. ¿Qué deberíamos hacer si hay demasiados factores potenciales de confusión en nuestra lista y no podemos determinar el cálculo ajustado adecuado?

9. ¿Qué hacer si la elección de dicho subconjunto se torna complicada?

Resumen

- La **confusión conjunta basada en los datos** surge cuando hay una diferencia relevante entre el cálculo crudo y el cálculo ajustado del efecto, que controla en forma simultánea *todos* los factores potenciales de confusión.
- La **confusión marginal basada en los datos** se produce cuando hay una diferencia relevante entre el cálculo crudo y el cálculo ajustado del efecto, que controla sólo *uno* de varios factores potenciales de confusión.
- Nuestras conclusiones sobre confusión deben basarse en la confusión conjunta siempre que sea posible,.

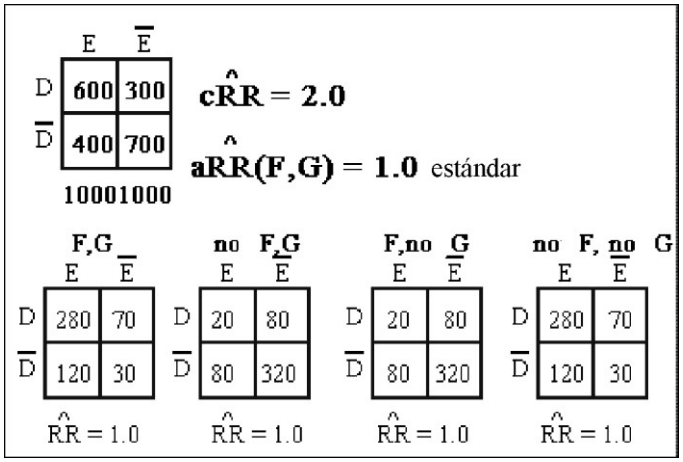
Confusión conjunta versus marginal: ejemplo

Suponga que se realiza un estudio de seguimiento para evaluar la relación entre una exposición y una enfermedad, y que la razón de riesgos cruda para estos datos es 2,0.

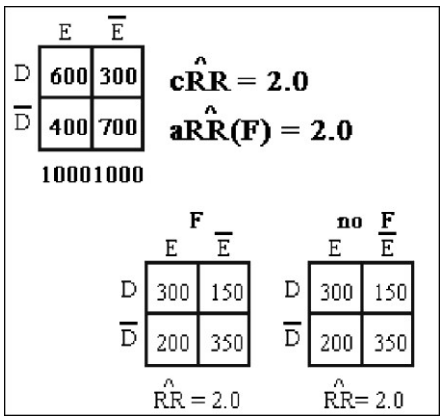
	E	$\overline{E}$	
D	600	300	$\hat{cRR} = 2.0$
$\overline{D}$	400	700	
	1000	1000	

Suponga además que se han identificado y medido dos variables dicotómicas **F** y **G** de posible control. Nos gustaría saber si podemos controlar **F** o **G** en forma separada o si debemos controlar **F** y **G** en forma simultánea para poder controlar adecuadamente la confusión.

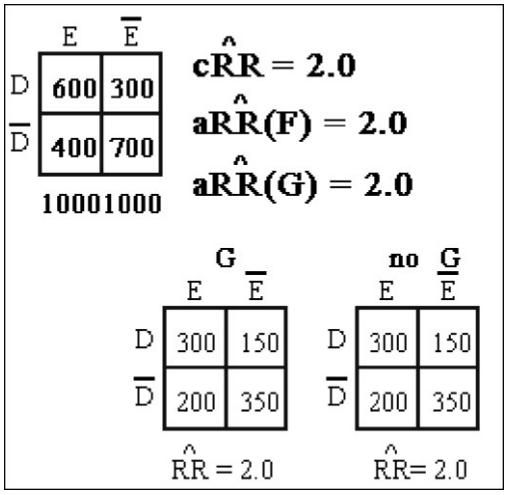
Recuerde el primer principio fundamental sobre confusión con varias variables. El cálculo ajustado que controla simultáneamente todos los factores de riesgo que se tienen en cuenta es el estándar sobre el cual deben basarse todas las conclusiones sobre la confusión. En nuestro ejemplo, el estándar es la razón de riesgos ajustada que controla **F y G**. Si estratificamos los datos tanto por **F** como por **G**, hallamos que cada razón de riesgos específica para cada estrato es igual a 1,0. Por lo tanto, se deduce que la razón de riesgos ajustada que controla **F y G** es 1,0. Esto difiere de la razón de riesgos cruda de 2,0. Entonces, podemos concluir de que hay confusión conjunta basada en los datos debido a **F** y a **G**.



Suponga que ignoramos el control conjunto de estas dos variables y evaluamos cada variable en forma separada. Cuando estratificamos los datos por **F**, vemos que las razones de riesgos específicas para cada estrato son iguales a 2,0, tanto cuando **F** está presente como cuando está ausente y, por lo tanto, la razón de riesgos ajustada que controla **F** es 2,0.



De la misma forma, si estratificamos los datos por **G**, vemos que las razones de riesgos específicas para cada estrato son iguales a 2,0, tanto cuando **G** está presente como cuando está ausente y, por lo tanto, la razón de riesgos ajustada es 2,0.



Ya que los cálculos ajustados que controlan **F** y **G** en forma separada son iguales a la razón de riesgos cruda de 2,0, podemos concluir que no hay confusión marginal basada en los datos debido a **F** o a **G**. Este ejemplo ilustra que pueden derivarse conclusiones incorrectas sobre confusión, si no se consideran las condiciones para la confusión conjunta.

Preguntas de estudio (P11.3)

1. En general, ¿cuál es el mejor indicador de la confusión basada en los datos en presencia de dos o más factores de riesgo: la confusión marginal o la conjunta?
2. En el ejemplo descrito en esta presentación, ¿había alguna interacción entre **F** y **E** o entre **G** y **E**?
3. Cuando se consideran dos o más factores de riesgo para un posible control, ¿es posible que no sea necesario controlar todos los factores de riesgo?
4. En este ejemplo, ¿es posible que sólo **F** o **G** sea un subconjunto adecuado para controlar la confusión?

Resumen

- El cálculo ajustado que controla simultáneamente todos los factores de riesgo en estudio es el estándar sobre el cual deben basarse todas las conclusiones sobre la confusión.
- Pueden derivarse conclusiones incorrectas sobre la confusión conjunta si se consideran sólo las condiciones para la confusión marginal.

Auto-evaluación (P11.4)

Suponga que **F** y **G** son dos factores de riesgo distintos para una enfermedad con niveles dicotómicos $\mathbf{F_1}$, $\mathbf{F_0}$ y $\mathbf{G_1}$, $\mathbf{G_0}$, respectivamente. A continuación se enumeran los cálculos de las razones de riesgos que describen la asociación entre la enfermedad y algún tipo de exposición, para varias combinaciones de niveles de **F** y **G**. Suponga que los cálculos de las razones de riesgos tienen alta precisión (es decir, están basados en tamaños de muestra grandes).

$\hat{RR}(F_1G_1) = 3.0$	$\hat{RR}(F_1) = 1.0$	$c\hat{RR} = 1.0$
$\hat{RR}(F_1G_0) = 3.0$	$\hat{RR}(F_0) = 1.0$	
$\hat{RR}(F_0G_1) = 0.3$	$\hat{RR}(G_1) = 1.0$	
$\hat{RR}(F_0G_0) = 3.0$	$\hat{RR}(G_0) = 1.0$	

Coloque **verdadero** o **falso** en las siguientes afirmaciones.

1. Hay evidencia de interacción en los datos. . . . . **???**
2. Hay evidencia de confusión en los datos. . . . . **???**
3. En el nivel $\mathbf{G_0}$ no hay confusión debido al factor **F**. . . . **???**
4. En el nivel $\mathbf{F_1}$ no hay interacción debido al factor **G**. . . . . **???**
5. En el nivel $\mathbf{F_0}$ no hay interacción ni confusión debido al factor **G**. . **???**
6. En el nivel $\mathbf{G_0}$ hay confusión, pero no hay interacción, debido al factor **F**.. . **???**
7. No es necesario controlar **F** o **G** (o ambos) para comprender la relación entre D y E. **???**

11-2 Confusión que involucra varios factores de riesgo

Confusión debido a uno de dos factores potenciales de riesgo

El segundo principio fundamental de la confusión que involucra varios factores de riesgo dice que puede no ser necesario controlar todos los factores de riesgo en un estudio específico. Un subconjunto de estas variables puede controlar la confusión. ¿Cómo determinamos cuáles son los subconjuntos candidatos?

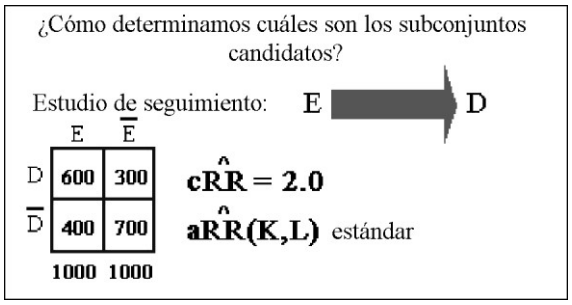
Confusión que involucra varios factores de riesgo

2. Puede no ser necesario controlar todos los factores de riesgo en un estudio específico.

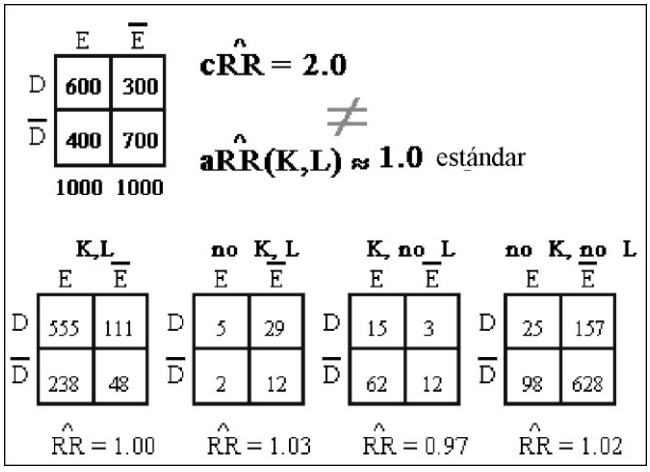
C_1
C_2
:
C_p
} Un subconjunto de estas variables puede controlar la confusión.

¿Cómo determinamos cuáles son los subconjuntos candidatos?

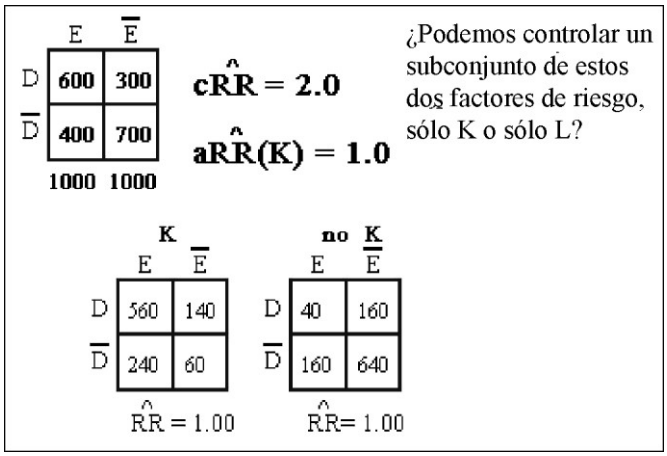
Suponga que se realizó un estudio de seguimiento para evaluar la relación entre una exposición y una enfermedad. La razón de riesgos cruda es igual a 2,0. Se han identificado y medido dos variables dicotómicas, **K** y **L**, para un posible control. La razón de riesgos ajustada que controla tanto **K** como **L** es el estándar sobre el cual deben basarse las conclusiones sobre confusión.



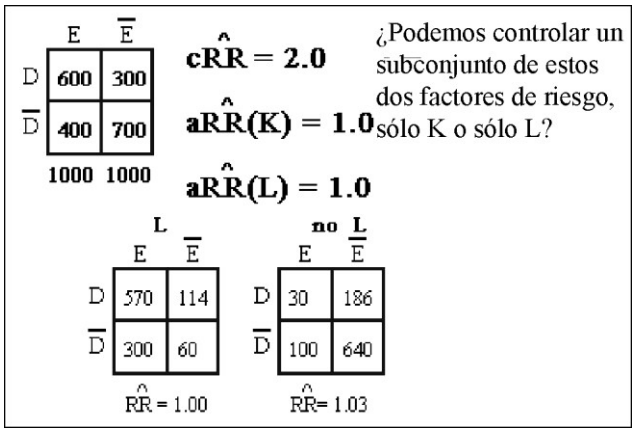
Si estratificamos los datos por **K** y por **L**, descubrimos que cada razón de riesgos específica para cada estrato es aproximadamente 1. Por lo tanto, la razón de riesgos ajustada, que controla tanto **K** como **L,** es aproximadamente 1. Esto difiere de la razón de riesgos cruda de 2. Entonces, podemos concluir que hay confusión conjunta basada en los datos debido a **K** y a **L**.



¿Podemos controlar un subconjunto de estos dos factores de riesgo (p. ej., sólo **K** o sólo **L**)? Si estratificamos los datos por **K**, vemos que las razones de riesgo específicas para cada estrato son iguales a 1,0, tanto para cuando **K** está presente como para cuando está ausente y, por lo tanto, la razón de riesgos ajustada que controla **K** es 1,0.



De la misma forma, si estratificamos los datos por **L**, vemos que las razones específicas para cada estrato son aproximadamente iguales a 1, tanto cuando **L** está presente como cuando está ausente y, por lo tanto, la razón de riesgos ajustada es 1,0.

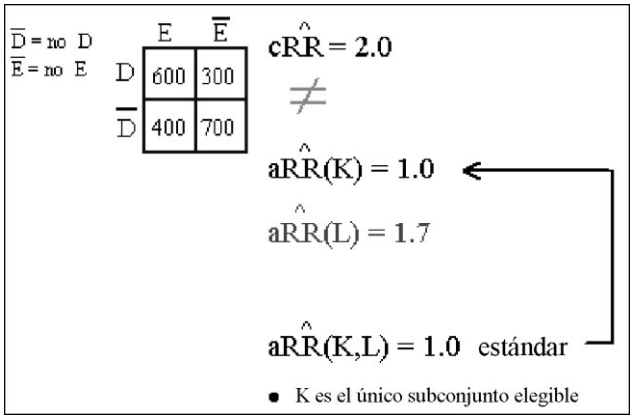


Como los cálculos ajustados que controlan tanto **K** como **L** en forma separada son diferentes de la razón de riesgos cruda, estos resultados ilustran la confusión marginal. En este ejemplo, no sería necesario controlar **K** y **L en conjunto**, ya que el control de cualquiera de estas dos variables arroja los mismos resultados que el estándar. Basta con controlar una de estas dos variables en forma separada.

Preguntas de estudio (P11.5)

1. ¿Cómo determinaría qué variable controlar (**K, L** o ambas) en el estudio mencionado anteriormente?
2. Si los resultados son los mismos, ¿hay alguna ventaja en controlar ambas variables?

Suponga ahora que **K** y **L** son los únicos factores de riesgo que se pueden elegir para el control y que estos son los resultados. Aquí, la razón de riesgos ajustada que controla tanto **K** como **L** no es igual a la razón de riesgos ajustada que controla sólo **L**. En este caso, **K** es el único subconjunto candidato de factores de confusión que puede elegirse para el control.



Preguntas de estudio (P11.5) (continuación)

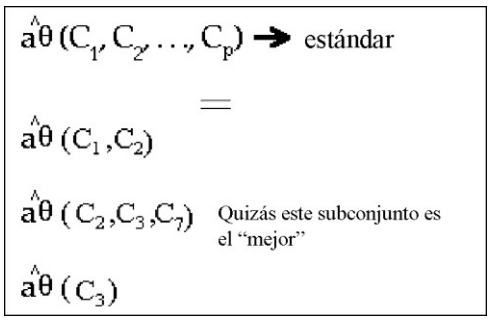
3. ¿Es adecuado controlar sólo **L** en el estudio anterior?
4. ¿Por qué es adecuado controlar sólo **K** en este estudio?
5. ¿Es adecuado controlar tanto **K** como **L** en este estudio?

Resumen

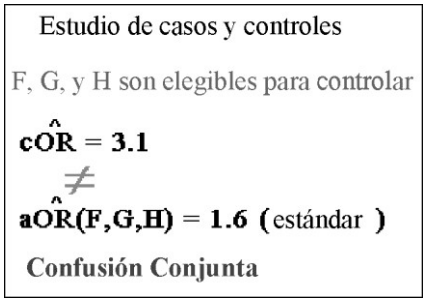
- Probablemente no sea necesario controlar todos los factores de riesgo en un estudio específico.
- Un subconjunto candidato de estas variables puede controlar la confusión.

Selección de las variables y control de la confusión

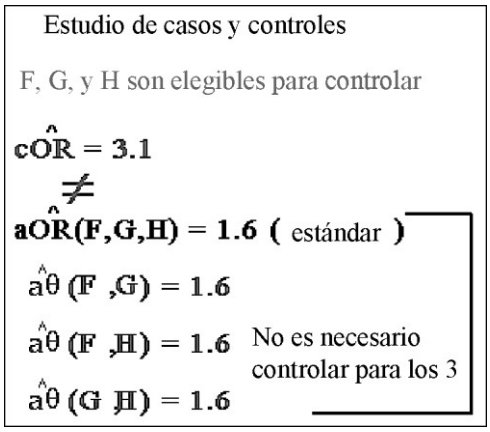
El cálculo ajustado que controla todos los factores potenciales de riesgo es el estándar sobre el cual deben basarse las conclusiones sobre la confusión. Sin embargo, si el cálculo ajustado que controla sólo un subconjunto de factores de riesgo es equivalente a este estándar, entonces podemos elegir dicho subconjunto para el control.



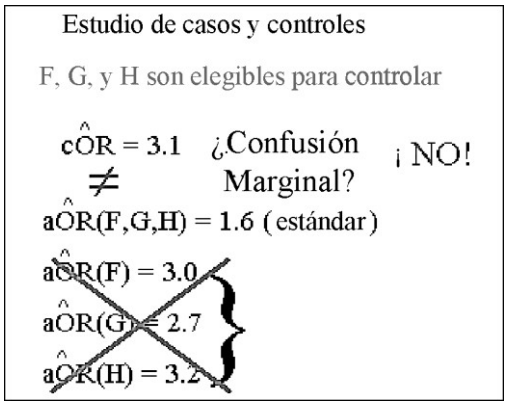
Considere un estudio de casos y controles en el que se quiere controlar tres factores de riesgo: **F**, **G** y **H**. El *odds ratio* crudo difiere del *odds ratio* ajustado que controla los tres factores. Como los valores del cálculo crudo y ajustado difieren, en estos datos tenemos evidencia de confusión conjunta basadas en los datos.



Suponga ahora que controlar dos de estos factores arroja los mismos resultados que controlar los tres. No es necesario controlar **todos** (**F**, **G** y **H**) en forma simultánea. El control de dos de estos tres factores de riesgo arrojará los mismos resultados que el estándar.



También podríamos considerar la confusión marginal para ver si cualquier variable individual está en el subconjunto adecuado. Estos resultados indican que no hay confusión marginal porque cada uno de estos resultados difiere del cálculo estándar; ninguna de estas variables en forma individual sería un subconjunto adecuado para el control.



Preguntas de estudio (P11.6)

$\hat{aOR}(F,G,H) = 1.6$ (estándar)	
$\hat{aOR}(F,G) = 1.6$	$\hat{aOR}(F) = 3.0$
$\hat{aOR}(F,H) = 1.6$	$\hat{aOR}(G) = 2.7$
$\hat{aOR}(G,H) = 1.6$	$\hat{aOR}(H) = 3.2$

1. ¿Cuál puede ser la ventaja de controlar sólo dos de estos factores de riesgo, en lugar de los tres, aun cuando se trate del estándar?
2. ¿Cómo podría determinar qué dos variables controlar?
3. ¿Por qué no podemos controlar **F**, **G** o **H** en forma separada?

Suponga que F, G, H, I y J son los únicos factores de riesgo en un estudio de casos y controles. Suponga, además, que:

$cOR \neq aOR(F, G, H, I, J)$

pero

$aOR(F, G, H, I) = aOR(G, J) = aOR(I) = aOR(F, G, H) = aOR(F, G, H, I, J)$

y que:

$aOR(\text{cualquier otro subconjunto de factores de riesgo}) \neq aOR(F, G, H, I, J)$

Determine si cada uno de los siguientes es un subconjunto adecuado de factores de confusión que controla la confusión (conjunta) en este estudio, respondiendo **sí** o **no**:

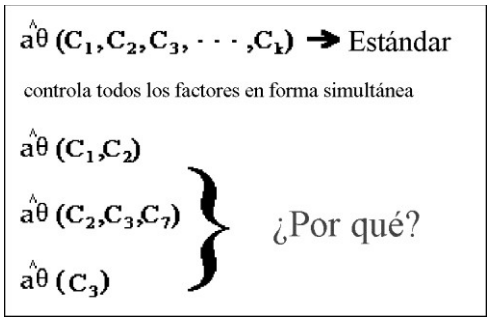
4. ¿{G, J}?
5. ¿{I}?
6. ¿{G, H, I}?
7. ¿{F, G, H}?

Resumen

- Cuando se considera controlar dos o más factores de riesgo, podemos seleccionar un subconjunto adecuado de factores de confusión para el control.
- Cuando los resultados de controlar varios subconjuntos de factores de riesgo son equivalentes al control conjunto de todos los factores de riesgo, podemos seleccionar cualquiera que brinde resultados válidos y precisos.

Confusión: validez versus precisión

El cálculo totalmente ajustado que controla todos los factores en forma simultánea es el estándar sobre el cual deben basarse todas las decisiones. ¿Por qué entonces quisiéramos pasar por el proceso de buscar subconjuntos candidatos de factores de confusión? Si no podemos mejorar la validez de la medida del efecto, ¿por qué simplemente no utilizar el cálculo totalmente ajustado?



La precisión del cálculo puede justificar dicho esfuerzo. El control de una menor cantidad de variables puede arrojar un cálculo más preciso del efecto. La identificación del subconjunto de factores de confusión que ofrece el cálculo más preciso es lo suficientemente importante como para hacer que dicho examen valga la pena. Considere el siguiente ejercicio para ejemplificar este punto.

Preguntas de estudio (P11.7)

Se realizó un ensayo clínico para determinar la efectividad de un tratamiento específico sobre la supervivencia de pacientes de cáncer en la etapa III. En el análisis, se consideraron las siguientes variables:

RX = exposición **EDAD** = edad de entrada al ensayo

SERH = nivel de hemoglobina en suero **TSZ** = tamaño del tumor primario

INSG = índice combinado que mide la etapa y el grado del tumor

1. La cRR = 6,28 y la aRR(EDAD, SERH, TSZ, INSG) = 8,24. ¿Esto indica confusión conjunta en el estudio? (Suponga que todas las cantidades anteriores son cálculos).

Calculamos la aRR para todos los subconjuntos posibles de los cuatro factores potenciales de confusión. Si se excluyen los resultados crudos y el estándar de oro, hay 14 subconjuntos posibles de estos 4 factores de confusión.

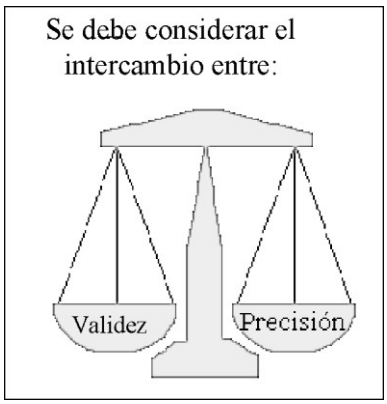
2. ¿Qué criterios podemos utilizar para reducir la cantidad de subconjuntos candidatos?

A continuación se presentan los resultados del estándar de oro y los 4 subconjuntos candidatos cuya aRR está dentro de un margen de ± 10% del estándar de oro.

Covariables	aRR	95% IC	IC Amplitud
INSG, EDAD, SERH, **TSZ**	8.24	(3.59, 18.91)	15.32
EDAD,**SERH**, **TSZ**	8.26	(3.64, 18.75)	15.11
INSG, EDAD	7.63	(3.75, 15.56)	11.81
INSG, **SERH**	7.63	(3.64, 15.97)	12.33
SERH	8.25	(4.06, 16.76)	12.70

3. ¿Qué covariables se controlan para obtener resultados más válidos de estos cálculos?
4. ¿Qué covariables se controlan para obtener resultados más precisos de estos cálculos?
5. ¿Qué covariables cree que son las más adecuadas para controlar?

Este ejercicio ha demostrado que necesitamos considerar tanto la validez como la precisión cuando evaluamos la relación entre una exposición y una enfermedad. Lo más importante es obtener un cálculo válido del efecto. Sin embargo, también se debe considerar el intercambio entre el control de suficientes factores de riesgo para mantener la validez y la posible pérdida de precisión al controlar demasiadas variables.



Resumen

- El motivo para buscar subconjuntos candidatos de todos los factores potenciales de confusión es la posibilidad de mejorar la precisión del cálculo de efecto.
- El control de una menor cantidad de variables puede (o no) arrojar un cálculo más preciso del efecto.
- Cuando se controlan varios factores potenciales de confusión, se deben considerar los posibles intercambios entre:
  - el control de suficientes factores de riesgo para mantener la validez
    versus
  - una posible pérdida de precisión por el control de demasiadas variables.

Auto-evaluación (P11.8)

Suponga que en determinado estudio se han medido las variables **F**, **G** y **H** y que sólo **F** y **G** se consideran factores de riesgo para alguna enfermedad (D). Suponga que es de interés describir la relación entre esta enfermedad y algún factor de estudio (E) y que no hay interacción de ningún tipo presente en los datos. Finalmente, suponga que las siguientes relaciones son válidas entre varios *odds ratios* calculados a partir de los datos:

$\widehat{cOR} \neq \widehat{aOR}(F)$	$\widehat{aOR}(F,G) = \widehat{aOR}(F)$
$\widehat{cOR} = \widehat{aOR}(G)$	$\widehat{aOR}(F,G) \neq \widehat{aOR}(G)$
$\widehat{cOR} = \widehat{aOR}(H)$	$\widehat{aOR}(F,H) = \widehat{aOR}(H)$
$\widehat{cOR} = \widehat{aOR}(F,G,H)$	$\widehat{aOR}(F,G,H) = \widehat{aOR}(G,H)$
$\widehat{cOR} \neq \widehat{aOR}(F,G)$	

Coloque **verdadero** o **falso** en las siguientes afirmaciones.

1. Hay confusión en los datos. . . . . **???**

2. Es necesario controlar la variable **F** para evitar confusión. . . **???**
3. Es necesario controlar la variable **G** para evitar confusión. . . **???**
4. Es necesario controlar la variable **H** para evitar confusión. . . **???**
5. No es necesario controlar las variables **F** y **G** simultáneamente para evitar confusión. . . . . **???**

Se realizó un estudio de seguimiento durante diez años para determinar si quien sufre de alergia alimenticias tiene mayor riesgo de padecer enfermedad coronaria. En el análisis, se consideraron las siguientes covariables: **EDAD** = edad de entrada al estudio, **IMC** = índice de masa corporal y **TAB** = tabaquismo. A continuación se enumeran los resultados del estándar y todos los subconjuntos posibles. La razón de riesgos cruda = 1,02.

	Covariables	aRR	95% IC	
1.	EDAD, IMC, TAB	4.10	(1.7, 8.8)	(estándar)
2.	EDAD, IMC	2.70	(1.6, 7.8)	
3.	EDAD, TAB	4.12	(1.4, 9.5)	
4.	IMC, TAB	2.50	(1.8, 6.2)	
5.	EDAD	4.01	(1.6, 8.7)	
6.	IMC	2.50	(1.4, 7.5)	
7.	TAB	4.16	(1.5, 8.6)	

6. ¿Hay evidencia de confusión? . . . . . **???**
7. Además del estándar, ¿cuáles son los subconjuntos candidatos de control? **???**
8. ¿Cuál de los subconjuntos candidatos corresponde al cálculo más válido (incluso el estándar)? . **???**
9. ¿Hay más de un subconjunto candidato que sea el más preciso? . . . **???**
10. ¿Qué cálculo debería utilizarse? . . . . . **???**

Opciones

N.° 2 **N.° 3** **N.° 4** **N.° 5** **N.° 6** **N.° 7** **no** **sí**

Nomenclatura

$\hat{\theta}$	Medida de efecto calculada
aOR	*Odds ratio* ajustado
aRR	Razón de riesgos ajustada
$\mathbf{C_i}$	Variable de confusión
IC	Intervalo de confianza
cOR	*Odds ratio* crudo
cRR	Razón de riesgos cruda
D	Enfermedad
E	Exposición

Referencia

Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic Research*: *Principles and Quantitative Methods.* John Wiley and Sons Publishers, Nueva York, 1982 (Capítulo 14).

Ejercicios

EPI-1. Identificación de los factores de riesgo para el control

Suponga que realiza un estudio de cohorte para evaluar la relación entre una variable de exposición dicotómica y una variable de enfermedad dicotómica. Identifique y mida tres factores de riesgo (f1, f2 y f3) que quiere controlar al evaluar la relación E-D. Suponga, además, que el análisis de sus datos arroja los siguientes cálculos de efecto:

aRR(f1, f2, f3) = 4,1	cRR = 1,3
aRR(f1, f2) = 2,5	aRR(f1) = 2,1
aRR(f2, f3) = 4,0	aRR(f2) = 4,0
aRR(f1, f3) = 2,9	aRR(f3) = 2,8

Suponiendo que no hay interacción de ningún tipo y que todos los cálculos anteriores son muy precisos, responda las siguientes preguntas:

a. ¿Hay confusión? Desarrolle su respuesta.
b. ¿Cómo debería decidir cuál de las variables anteriores necesita controlar? (Provea una estrategia, no una conclusión).
c. ¿Cuáles de las variables no son factores de confusión? Desarrolle su respuesta.
d. ¿Qué conclusiones puede sacar sobre qué variables debería controlar en este estudio?
e. Si usa estrictamente una regla de margen de cambio de 10% para determinar si un cálculo ajustado difiere del cálculo ajustado que se considera estándar de oro, ¿qué conclusiones sacaría sobre qué variables deberían controlarse en este estudio?
f. Si usa estrictamente una regla de margen de cambio de 20% para determinar si un cálculo ajustado difiere del cálculo ajustado que se considera estándar de oro, ¿qué conclusiones sacaría sobre qué variables deberían controlarse en este estudio?

EPI-2. Selección de las variables

Suponga que realiza un estudio de casos y controles para evaluar la relación entre una variable de exposición dicotómica y una variable de enfermedad dicotómica. Identifique y mida cuatro factores de riesgo, f1 (antecedentes previos de trastorno mental), f2 (tipo de personalidad), f3 (edad) y f4 (sexo) que quiere controlar al evaluar la relación E-D. Suponga, además, que el análisis de sus datos arroja los siguientes cálculos de efecto:

aOR(f1, f2, f3, f4) = 1,68
aOR(f2, f3, f4) = 1,20
aOR(f1, f3, f4) = 1,20
aOR(f1, f2, f4) = 2,78
aOR(f1, f2, f3) = 1,18
aOR(f3, f4) = 1,69
aOR(f1, f4) = 2,78
aOR(f1, f3) = 1,68
aOR(f2, f4) = 1,70
aOR(f2, f3) = 5,88
aOR(f1, f2) = 5,95
aOR(f4) = 2,75

aOR(f3) = 5,85
aOR(f2) = 1,20
aOR(f1) = 1,69

Dada la información anterior y suponiendo que no hay interacción de ningún tipo, utilice una regla de margen de cambio de 10% para determinar qué variables se deberían incluir para corregir la confusión. ¿Cómo podría la precisión cumplir una función en términos de qué variables se seleccionan para el control?

EPI-3. Confusión conjunta y marginal

Se realizó un estudio de casos y controles para evaluar si el consumo de alcohol era un factor de riesgo para el desarrollo de cáncer de mama en las mujeres. La variable de exposición se indicó como ALC y se categorizó en 3 grupos (1= sin consumo de alcohol, 2= consumo de alcohol bajo a moderado y 3= consumo de alcohol alto). Se consideraron tres factores de riesgo como variables de control: EDAD(1= menos de 50, 2= 50 o más), tabaquismo TAB(1= alguna vez, 2= nunca), y OBESIDAD (0= no, 1= sí). Los siguientes *odds ratio*s ajustados se obtuvieron cuando las personas que consumían alcohol de manera moderada (ALC=1) se compararon con las que no bebían alcohol (ALC=0), y a las personas con consumo de alcohol alto (ALC=2) con las que no bebían alcohol (ALC=0):

Variables controladas	aOR(1 a 0)	aOR(2 a 0)
Ninguna	4,5	6,0
EDAD	3,4	5,1
TAB	2,0	2,8
OBESIDAD	2,9	4,1
EDAD, TAB	1,8	3,3
EDAD, OBESIDAD	4,6	5,6
TAB, OBESIDAD	3,2	5,4
EDAD, TAB, OBESIDAD	1,9	3,1

Responda las siguientes preguntas suponiendo que no hay interacción de ninguna de las variables de control con ALC y utilizando una regla de margen de cambio de 10% para determinar una diferencia relevante en los *odds ratios* ajustados:

a. ¿Hay confusión? Justifique su respuesta.
b. ¿Qué subconjuntos de variables brindan el mismo *odds ratio* ajustado que el que se considera estándar de oro? Justifique su respuesta.
c. ¿Cómo consideraría la precisión para determinar qué subconjunto de variables controlar?
d. Suponga que no hay ninguna ganancia importante de precisión cuando se controla un subconjunto adecuado de las tres variables de control. ¿Qué variables controlaría? Justifique su respuesta.

EPI-4. Confusión marginal

La tabla a continuación muestra los resultados de un análisis estratificado de datos recabados en un estudio de casos y controles. La variable del resultado es dicotómica y se denomina A. Las variables predictoras están indicadas del 1 al 10, donde la variable 1 es la variable de exposición de interés. Las variables 2 a 10 son las variables de control.

Resultados del análisis estratificado en el examen de la asociación entre la variable 1 y el resultado (variable A)

Variable de riesgo controlada	Cálculo del OR por sub-estratos		OR_{M-H} ajustado	Intervalo de confianza de 95%	Prueba de Breslow-Day de homogeneidad (valor p)
	Estrato 1	Estrato 2			
Variable 2	1,550	2,360	1,998	(1,070; 3,730)	0,522
Variable 3	5,758	1,840	3,083	(1,517; 6,268)	0,136
Variable 4	3,300	1,875	2,040	(1,088; 3,822)	0,534
Variable 5	1,250	3,829	2,711	(1,357; 5,415)	0,151
Variable 6	0,563	2,972	1,813	(0,984; 3,341)	0,022
Variable 7	2,000	1,648	1,711	(0,889; 3,293)	0,819
Variable 8	1,125	2,134	2,032	(1,082; 3,819)	0,603
Variable 9	1,333	2,146	1,964	(1,044; 3,693)	0,570
Variable 10	1,9028	1,950	1,931	(1,026; 3,632)	0,970

cOR para la variable 1 vs. la variable A (resultado) = 2,022 (1,084 – 3,772)

a. Suponiendo que no hay interacción de ningún tipo, ¿cómo evaluaría si hay confusión o no? ¿Se ha brindado suficiente información en la tabla anterior para permitirle responder esta pregunta?
b. Nuevamente, suponiendo que no hay interacción, ¿hay confusión marginal debido a alguna de las variables de control? Desarrolle su respuesta.
c. Nuevamente, suponiendo que no hay interacción, ¿cómo determinaría qué variables controlar? ¿Cuál es la principal razón de por qué no puede responder esta pregunta sobre la base de los datos brindados anteriormente (suponiendo que no hay interacción)?
d. Suponiendo que hay interacción de las variables 9 y 10 con la exposición (variable 1), ¿cómo modificaría su respuesta al ítem c) para determinar qué variables controlar?

Respuestas a las preguntas de estudio y auto-evaluaciones

P11.1

1. El cálculo ajustado que controla simultáneamente los 4 factores de riesgo considerados.
2. La confusión no se podría controlar si no hay un subconjunto de factores potenciales de confusión que dé, básicamente, el mismo cálculo ajustado que el obtenido cuando se controlan todos los factores de confusión.
3. Sí, siempre que el subconjunto arroje básicamente el mismo cálculo ajustado que el estándar de oro.
4. El ajuste por una menor cantidad de variables puede aumentar la precisión. Asimismo, dicho subconjunto ofrece una descripción más exigua de la relación entre la exposición y la enfermedad.

P11.2

1. Sí. La cRR de 1,5 difiere de la aRR(edad, tabaquismo) de 2,4 que controla ambos factores potenciales de confusión.
2. No. La cRR de 1,5 es básicamente igual a la aRR(edad, tabaquismo) de 1,4 que controla ambos factores potenciales de confusión.
3. No. La cRR de 1,5 difiere de la aRR(edad, tabaquismo) de 2,4 que controla todos los factores potenciales de confusión. Esto indica que hay confusión conjunta.
4. No. Como la cRR de 1,5 es aproximadamente igual a la aRR(edad) de 1,4, no hay evidencia de confusión marginal debido a la edad.
5. No necesariamente. Nuestras conclusiones sobre la confusión deben basarse en el control conjunto de todos los factores de riesgo.
6. Sí. Controlar sólo el tabaquismo nos da el mismo resultado que controlar ambos factores de riesgo. De todos modos, podríamos querer evaluar la precisión de los cálculos antes de sacar una conclusión final.
7. Puede haber tantos factores de riesgo en nuestra lista en relación con la cantidad de datos disponibles, que el cálculo ajustado no puede determinarse con ninguna precisión.
8. Entonces, podemos estar forzados a tomar decisiones mediante un subconjunto de este gran conjunto inicial de factores de riesgo.
9. El uso de la confusión marginal puede ser la única alternativa.

P11.3

1. Debería utilizarse la confusión marginal, siempre que sea posible, como la referencia a partir de la cual deberían examinarse todos los otros temas relacionados con la confusión.
2. No. Las dos RR específicas para cada estrato que comparan **F** con **no F** son iguales y las dos RR específicas para cada estrato que comparan **G** con **no G** son iguales.
3. Sí. El segundo principio fundamental de la confusión dice que "no todas las variables en una lista dada de factores de riesgo deben controlarse"; es posible que diferentes subconjuntos de dichas variables puedan corregir la confusión en forma alternativa.
4. No. El control de dos cualesquiera de estos tres factores de riesgo en forma separada arroja los mismos resultados que los datos crudos. Sólo observamos confusión en el control conjunto de estos factores.

P11.4

1. Verdadero. Hay interacción porque la razón de riesgos calculada en un estrato (F_0G_1) es 0,3, que es bastante diferente de las razones de riesgos específicas para cada estrato de 3,0 en los otros estratos.
2. Verdadero. La presencia de fuerte interacción puede impedir la evaluación de la confusión. Asimismo, el valor de un cálculo ajustado puede variar según las ponderaciones elegidas para los diferentes estratos.
3. Falso. Las RR para F_1 y F_0 en el nivel G_0 son ambas iguales a 3,0. Éstas difieren de la RR general en el nivel G_0 de 1,0. Por lo tanto, en el nivel G_0, hay confusión debido al factor F.
4. Verdadero
5. Falso. Hay interacción y, posiblemente, confusión. En el nivel F_0, las RR para G_1 y G_0 son muy diferentes, y ambas son muy distintas de la razón de riesgos general en el nivel F_0.
6. Verdadero
7. Falso. Hay tanto confusión como interacción y se debería abordar cada una de ellas.

P11.5

1. Es posible que desee controlar la(s) variable(s) que arroja(n) el cálculo más preciso.
2. Sí. Es posible que desee controlar ambas si no gana nada en términos de precisión al dejar una de lado. Aunque estos resultados pueden ser iguales, si deja de lado una variable en el análisis, no queda claro, para quien lo revise, que controló ambas.
3. No. Los resultados cuando se controla sólo L difieren de los resultados cuando se controlan tanto K como L, el estándar sobre el cual deben basarse todas las conclusiones sobre la confusión.
4. El control de K arroja los mismos resultados que el control de los dos factores K y L en forma simultánea.
5. Sí. Si no gana nada en términos de precisión al dejar L de lado.

P11.6

1. El control de menor cantidad de variables probablemente aumentará la precisión de los resultados.
2. Las dos que brinden el cálculo ajustado más preciso.
3. El control de sólo uno de estos tres factores arroja resultados diferentes que el control de los tres, lo cual es el estándar sobre el cual deben basarse nuestras conclusiones.
4. Sí
5. Sí
6. No
7. Sí

P11.7

1. Sí. La cRR difiere de la aRR que controla todos los factores potenciales de confusión, la cual es el estándar de oro.
2. Podemos optar por considerar sólo aquellos resultados dentro del 10% del estándar de oro. En este caso, eso sería 8,24 ± 0,82, que es un rango de valores entre 7,42 y 9,06.
3. El control de todas las covariables ofrece el cálculo más válido. Es el estándar de oro.
4. El control tanto de INSG como de EDAD ofrece el intervalo de confianza más acotado y, por lo tanto, es el más preciso.
5. Discutible: el control sólo de SERH arroja una aRR casi idéntica al estándar de oro, aumenta la precisión y es el subconjunto más reducido. El control de INSG y EDAD ofrece un aumento levemente mayor de la precisión (que controlar sólo SERH) y su aRR está dentro del 10% del estándar. Considere el intercambio entre la insuficiencia y las implicaciones políticas/científicas de no controlar todos los factores de riesgo, y una mayor precisión al controlar menor cantidad de factores de riesgo.

P11.8

1. Verdadero
2. Verdadero
3. Falso. No es necesario controlar la variable G, ya que aOR(F,G) = aOR(F). En otras palabras, el control sólo de F arroja los mismos resultados que el estándar de oro, que controla tanto F como G.
4. Falso. La variable H no es un factor de riesgo en este estudio y, por lo tanto, no debería considerarse una confusión.
5. Verdadero
6. Sí. La cRR de 1,02 difiere de la RR estándar de 4,10 que controla todos los factores potenciales de confusión.
7. N.° 3, n.° 5, n.° 7
8. N.° 1. El cálculo más válido controla todos los factores de riesgo medidos.
9. Sí. Los subconjuntos candidatos 1 y 7 son igualmente precisos.
10. N.° 1. El estándar de oro es el cálculo más válido; tiene la misma precisión que la obtenida para el candidato 7. No se gana precisión alguna al dejar de lado alguno de los factores de riesgo, por lo cual puede argumentarse que el estándar de oro es la elección “política”, ya que controla todos los factores de riesgo considerados. Controlar sólo TAB es la mejor opción, ya que ofrece el subconjunto más reducido y más preciso de variables.

LECCIÓN 12

ANÁLISIS SIMPLES

12-1 Inferencia estadística para los análisis simples

*Esta lección examina los métodos para realizar procedimientos de **inferencia estadística** para datos epidemiológicos dados en un cuadro simple de 2 x 2. Dichos procedimientos se denominan **análisis simples**, porque aquí se restringe la discusión sólo a variables de enfermedad y exposición dicotómicas y se ignora la situación de análisis típica que considera el control de **otras** variables cuando se estudia el efecto de una exposición sobre la enfermedad.*

¿QUÉ ES EL ANÁLISIS SIMPLE?

Cuando se analizan los datos crudos que describen la relación entre una variable de exposición dicotómica y una variable de enfermedad dicotómica, normalmente se busca hacer inferencias estadísticas sobre esta relación. Es decir, se procura determinar si la medida de efecto que se calcula es estadísticamente significativa y se desea obtener una estimación de intervalos que considere la variabilidad muestral de la medida de efecto.

Los cuadros que se muestran a continuación se describieron en lecciones anteriores para ejemplificar datos de tres estudios distintos: un estudio de cohorte para evaluar si el hecho de dejar de fumar después de haber sufrido un ataque cardíaco reducirá el riesgo de morir que tiene un individuo; un estudio de casos y controles para determinar la fuente de un brote de diarrea en un club de vacaciones en Haití; y un estudio de cohorte con datos de persona-tiempo para evaluar la relación entre el nivel de colesterol sérico y la mortalidad.

Pacientes ataque cardíaco	Fuman	Dejaron de fumar	Total
Muerte	27	14	41
Supervivencia	48	67	115
Total	75	81	156

COHORTE

Dejar de fumar → Mortalidad

$\widehat{RR} = 2.1$

Estudio del Club	Hamburguesa Cruda Comió	No comió	Total
Casos	17	20	37
Controles	7	26	33
Total	24	46	70

CASOS Y CONTROLES

Fuente → Diarrea

$\widehat{OR} = 3.2$

Estudio de colesterol	Nivel de Colesterol Alto	Normal
Muertes	26	14
P-años	36,581	68,239

COHORTE CON DATOS DE PERSONA-TIEMPO

Colesterol → Mortalidad

$\widehat{RDI} = 3.5$

En cada estudio se calculó una medida de efecto para evaluar el grado de relación entre la variable de exposición y la variable del resultado de salud. En el estudio de cohorte donde se consideró el abandono del

tabaquismo, la medida de efecto era una razón de riesgos y su valor fue de 2,1. En el estudio sobre el brote de diarrea, la medida de efecto era un *odds ratio* y su valor fue de 3,2. Y en el estudio sobre mortalidad y colesterol, la medida de efecto era una razón de tasas, también denominada razón de densidades de incidencia, cuyo valor fue de 3,5.

Hemos examinado la manera de interpretar cada uno de estos cálculos en términos de la relación entre la exposición y la enfermedad que se estudia. Aun cuando se trate de diferentes tipos de estudios, que plantean preguntas distintas, los tres valores son similares, puesto que son mayores que el **valor nulo de 1**; además, todos indican que hay un efecto moderadamente grande de la exposición. Sin embargo, debemos tener la prudencia de darnos cuenta de que cada uno de estos tres valores se basa en datos muestrales. Si se hubiese tomado otra muestra, cualquiera de estos cálculos habría dado un valor distinto como resultado, quizás mucho más grande, tal vez más cercano al valor nulo de 1. Es decir, siempre hay **error aleatorio** asociado a cualquier cálculo muestral.

Estos cálculos se denominan **estimaciones puntuales** porque cada uno de ellos representa un número único o punto del rango posiblemente amplio de números que podrían haberse obtenido si se hubiesen tomado muestras diferentes. Por lo tanto, podríamos preguntarnos, dada la variabilidad inherente a la estimación puntual, ¿cómo podemos sacar conclusiones sobre los **parámetros poblacionales** que se están calculando?

Por ejemplo, en el primer estudio de cohorte, ¿qué podemos decir sobre la **razón de riesgos de la población** basada en la **razón de riesgos calculada**? O, en el estudio de casos y controles, ¿a qué conclusión podemos llegar sobre el ***odds ratio* de la población** basado en el ***odds ratio* calculado**? En el estudio de persona-tiempo, ¿qué podemos inferir sobre la **razón de tasas de la población** basada en la **razón de tasas calculada**? Al responder estas preguntas, normalmente tenemos uno de dos objetivos. Es posible que queramos determinar si tenemos pruebas en la muestra de que la razón de riesgos de la población, el *odds ratio* o la razón de tasas que se calculan, son diferentes del valor nulo de 1. Para tal fin, utilizamos la **prueba de hipótesis**. O tal vez deseemos establecer la precisión de nuestra estimación puntual justificando su **variabilidad muestral**. Para dicho propósito utilizamos la **estimación de intervalos**.

El método que se emplea para lograr estos objetivos comprende el tema general de la **inferencia estadística**. Cuando centramos nuestra atención en la situación relativamente simple que involucra sólo una variable de exposición dicotómica y una variable de enfermedad dicotómica, como se ejemplifica mediante estos tres estudios, llamamos a estos métodos: **análisis simples**.

Resumen

- Los cálculos de las medidas de efecto como, por ejemplo, la RR, el OR y la RDI son **estimaciones puntuales**, ya que cada valor representa un número único que puede variar de muestra en muestra.
- La **inferencia estadística** implica derivar conclusiones sobre el valor de una medida de efecto en una población, sobre la base de su cálculo obtenido de una muestra.
- Los dos tipos de procedimientos de **inferencia estadística** son la **prueba de hipótesis** y la **estimación de intervalos**.

INFERENCIAS ESTADÍSTICAS: REVISIÓN

Las actividades de esta sección examinan conceptos y métodos fundamentales de estadística. Nos concentramos principalmente en cómo sacar conclusiones sobre las poblaciones basándonos en los datos obtenidos de una muestra. Suponemos que usted ya ha tenido algún tipo de exposición previa a los conceptos estadísticos básicos, incluso la distinción entre una muestra y una población, un estadístico muestral y un parámetro poblacional, algunas distribuciones importantes como la normal, binomial, t de Student y chi cuadrado. También suponemos

que ya ha tenido alguna exposición previa a los conceptos subyacentes a las pruebas estadísticas de hipótesis y los intervalos de confianza, que son los dos tipos de inferencias estadísticas posibles. Aquí nos enfocaremos en el repaso de los conceptos de inferencia estadística en el contexto de las preguntas estadísticas que corresponden al análisis de un cuadro de 2 x 2, que es el tipo de datos que se considera en un análisis simple. Quizás desee obviar toda esta parte dedicada a la revisión y continuar con la sección 12-4, Estudios de cohorte. Comenzamos utilizando datos de un famoso "incidente" para distinguir entre los dos tipos de procedimientos de inferencia estadística: la prueba de hipótesis (significación) y la estimación de intervalos de confianza. Veamos si puede adivinar qué "incidente" usaremos como ejemplo.

Revisión general sobre inferencia estadística

A continuación se incluyen algunos datos de un incidente en el cual un grupo de personas corría riesgo de muerte. A partir de estos datos, podemos encontrar las proporciones de los que fallecieron, separándolas por sexo. Podemos ver que el 79,7% de los hombres perdieron la vida, pero sólo 25,6% de las mujeres murieron

UN INCIDENTE

Personas en Riesgo

	Hombres	Mujeres	Total
Murieron	1329	109	1438
Vivieron	338	316	654
Total	1667	425	2092

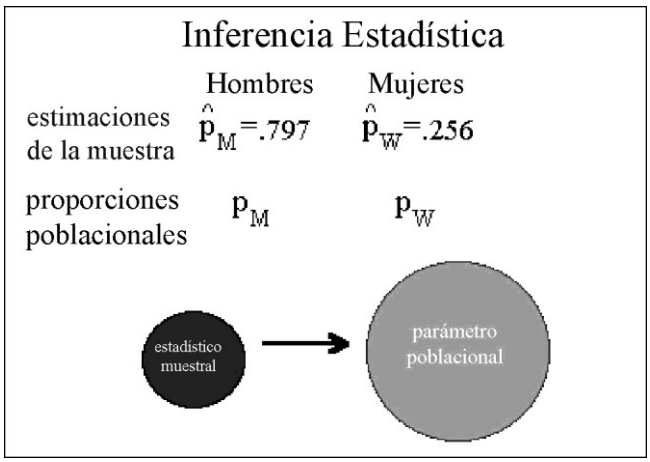
.797 .256 ¡Significativamente

79.7 % 25.6 % distintos!

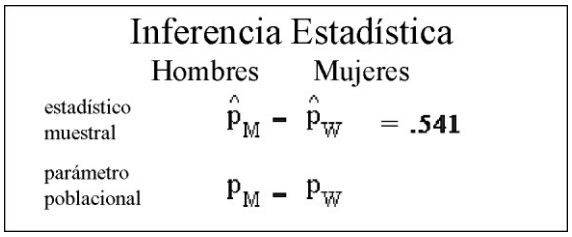
Claramente, estos dos porcentajes son significativamente distintos, ya que los hombres tenían un riesgo mucho mayor de morir que las mujeres. Pero, ¿podemos decir también que hay una disparidad en el riesgo de muerte entre hombres y mujeres en la población de donde provinieron estas muestras? En otras palabras, ¿la variación en los riesgos para los hombres y las mujeres es **estadísticamente significativa**?

Si deseamos sacar conclusiones sobre una población a partir de los datos recabados de una muestra, debemos considerar los métodos de **inferencia estadística**. En particular, tenemos que ver las dos proporciones, o porcentajes o binomios, como cálculos obtenidos de una muestra. Concentrémonos en las dos proporciones muestrales, que denotamos $\hat{p}_M$ y $\hat{p}_W$. Las respectivas proporciones poblacionales se denotan como p_M y p_W, sin acentos circunflejos.

La inferencia estadística nos lleva a conclusiones sobre un parámetro poblacional basándose en la información obtenida de un estadístico muestral. Por lo tanto, ¿cuál es el parámetro poblacional que se considera para estos datos y cuál es el estadístico muestral correspondiente?



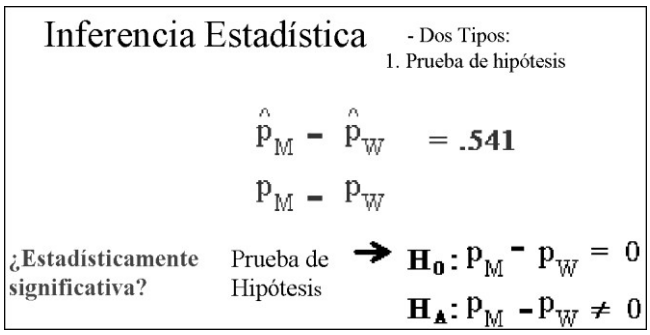
Ya que aquí nos enfocamos en **comparar** las proporciones para hombres y mujeres, una elección lógica para nuestro parámetro de interés es la diferencia entre las dos proporciones poblacionales. El estadístico muestral correspondiente es la diferencia entre las dos proporciones calculadas.



Preguntas de estudio (P12.1)

1. ¿Qué otros parámetros (epidemiológicos) podrían también considerarse como alternativas a la diferencia entre las dos proporciones?

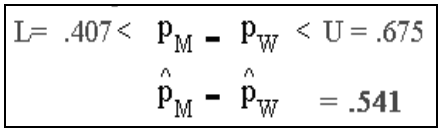
La prueba de hipótesis puede utilizarse aquí para determinar si la diferencia entre las dos proporciones es estadísticamente significativa. Éste es uno de los dos tipos de preguntas de inferencia estadística que podemos formular. Nuestra hipótesis en este caso, generalmente enunciada como lo que queremos refutar, es que la verdadera diferencia entre las dos proporciones poblacionales es cero. Esto se denomina **hipótesis nula**. En la prueba de hipótesis, buscamos justificaciones en la muestra para rebatir la hipótesis nula a favor de la **hipótesis alternativa** que indica que hay una diferencia entre las proporciones poblacionales.



Preguntas de estudio (P12.1) (continuación)

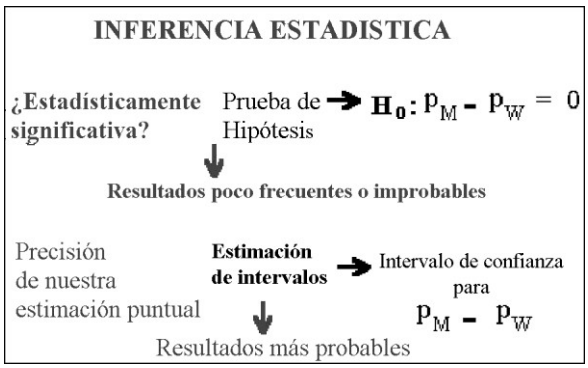
2. Si el parámetro de interés es la razón de riesgos (RR), ¿cómo expresaría la hipótesis nula?
3. Si el parámetro de interés es el *odds ratio* (OR), ¿cómo expresaría la hipótesis nula?

Podemos utilizar la **estimación de intervalos** para determinar la precisión de nuestra estimación puntual. Aquí, nuestro objetivo es emplear información muestral que tenemos para calcular dos números, digamos **L** y **U**, que definen un intervalo de confianza para la diferencia entre las dos proporciones poblacionales. Mediante un intervalo de confianza, podemos predecir con cierto grado de seguridad, por ejemplo, el 95%, que los límites, **L** y **U**, encierran el verdadero valor del parámetro. Para nuestros datos, surge que los límites inferior (L) y superior (U) correspondientes a la diferencia entre las dos proporciones son 0,407 y 0,675, respectivamente.



A partir de estos dos números, puede parecer que la estimación de intervalos es menos precisa que la estimación puntual. En realidad, sucede todo lo contrario. De hecho, el rango de valores especificados por la estimación de intervalos considera la falta de fiabilidad o varianza de la estimación puntual. Por lo tanto, es más preciso, ya que utiliza más información para describir la estimación puntual.

En general, la estimación de intervalos y la prueba de hipótesis se pueden contrastar por los diferentes enfoques que tienen para responder las preguntas. La prueba de hipótesis llega a una respuesta buscando resultados muestrales poco frecuentes o improbables. Por el contrario, la estimación de intervalos presenta respuestas observando los resultados más probables, es decir, aquellos valores con respecto a los cuales estamos seguros que son cercanos al parámetro que investigamos.



Resumen

- La **inferencia estadística** implica sacar conclusiones de un parámetro poblacional sobre la base de información obtenida de un estadístico muestral.
- Los dos tipos de inferencia estadística son la **prueba de hipótesis** y la **estimación de intervalos**.
- Cuando hacemos una prueba de hipótesis, normalmente queremos refutar una hipótesis nula.
- Cuando realizamos estimaciones de intervalos, queremos obtener un intervalo de confianza que brinde los límites superior e inferior de los que podamos estar, digamos 95% seguros de que abarcan el parámetro poblacional de interés.
- La prueba de hipótesis busca resultados poco frecuentes o improbables a partir de nuestra muestra, mientras que el intervalo de confianza apunta a encontrar resultados probables.

El incidente

La historia del Titanic es famosa. El barco más grande que se había construido hasta ese momento zarpó de Southampton, Inglaterra, en su viaje inaugural a Nueva York el miércoles 10 de abril de 1912, transportando a muchos de los ricos y famosos de Inglaterra y Estados Unidos, pero también a muchas personas menos pudientes. Debido a que el Titanic era tan grande y moderno, muchos pensaron que no podía hundirse.

Después de una escala en Cherbourg, Francia, donde abordaron muchos pasajeros de tercera clase en busca de otra vida en el Nuevo Mundo y, tras una breve escala en las afueras de Queenstown, Irlanda, emprendió viaje cruzando el Atlántico. A las 11:40 de la noche del 14 de abril el Titanic chocó contra un iceberg, hundiéndose a las 2:15 de la mañana siguiente.

De los 2.201 pasajeros y miembros de la tripulación, sólo sobrevivieron 710. Se pudieron conocer algunos datos sobre los pasajeros, quién sobrevivió y quién no. Un interrogante de interés que subyace en cualquier tragedia de este tipo es si todos tuvieron la misma oportunidad de sobrevivir. O, expresado en términos estadísticos, si la probabilidad de sobrevivir era independiente de otros factores.

12-2 Inferencia estadística para los análisis simples (continuación)

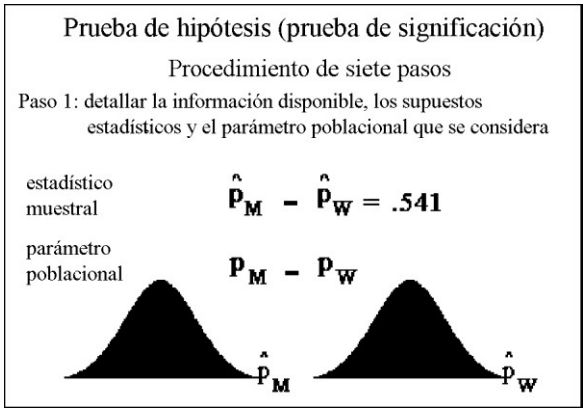
Prueba de hipótesis

Prueba de hipótesis, parte 1

Ejemplificamos cómo realizar una prueba estadística de hipótesis para comparar la supervivencia de los pasajeros de sexo masculino y femenino del Titanic. Queremos evaluar si la diferencia entre las proporciones muestrales para los hombres y las mujeres es estadísticamente significativa.

La **prueba de hipótesis**, también llamada **prueba de significación**, puede describirse como un procedimiento de siete pasos.

En el **paso 1** detallamos la información disponible, los supuestos estadísticos y el parámetro poblacional que se considera. En nuestro ejemplo, la información incluye las cantidades de pasajeros hombres y mujeres, al igual que las proporciones muestrales. Debemos suponer que los datos representan una muestra de una población más grande y que cada proporción muestral tiene aproximadamente una distribución normal. El **parámetro poblacional** de interés es la diferencia entre las dos proporciones poblacionales que se calculan (P_M - P_W). El **estadístico muestral** correspondiente es la diferencia entre las dos proporciones calculadas $(\hat{P}_M - \hat{P}_W)$.



En el **paso 2** especificamos las **hipótesis nula** y **alternativa**. Una hipótesis es una afirmación sobre el valor de un parámetro poblacional. La hipótesis que planeamos probar normalmente se denomina hipótesis nula, que a menudo es el estado aceptado de conocimiento que queremos cuestionar. La hipótesis nula en nuestro ejemplo es que la diferencia entre las dos proporciones poblacionales es cero. Básicamente, se trata a $\mathbf{H_0}$ como al acusado en un juicio. Se supone verdadera, o inocente, hasta que las pruebas obtenidas de los datos hacen que sea muy poco probable que haya ocurrido de casualidad. Nuestro objetivo aquí es ver si tenemos pruebas para refutar esta hipótesis nula. La hipótesis alternativa, normalmente llamada $\mathbf{H_A}$, da los valores que puede tomar el parámetro si la hipótesis nula es falsa. En nuestro ejemplo, la hipótesis alternativa es que la diferencia entre las proporciones poblacionales no es igual a cero. Esto se denomina **hipótesis alternativa de dos colas** porque indica que estamos interesados en valores tanto por arriba como por debajo de la hipótesis nula. Diremos que la hipótesis alternativa tiene **una cola** si, antes de observar nuestros datos, estuviésemos interesados en determinar sólo si los hombres tenían mayor riesgo de morir que las mujeres. Para evitar sesgar el análisis, tanto la hipótesis nula como la alternativa deben formularse sin observar los datos del estudio y deben basarse sólo en los objetivos *a priori* del estudio.

Paso 2: Especificar las hipótesis nula y alternativa

$$\text{APRIORI} \begin{cases} H_0: & p_M - p_W = 0 \\ H_A: & p_M - p_W \neq 0 \quad \text{Dos colas} \\ & p_M - p_W > 0 \quad \text{Una cola} \end{cases}$$

En el **paso 3** especificamos el **nivel de significación alfa**. Vamos a establecer el nivel de significación para nuestro ejemplo en 0,05 ó 5%. Esto implica que al llevar a cabo nuestro procedimiento, estamos dispuestos a asumir 5% de riesgo de rechazar la hipótesis nula, aun cuando en verdad sea cierta. De manera equivalente, el nivel de significación nos dice qué poco frecuentes o improbables deben ser los resultados de nuestro estudio conforme a la hipótesis nula para que la rechacemos a favor de la hipótesis alternativa. Un alfa equivalente a 5% significa que, si la hipótesis nula es, en verdad, cierta, tendremos 5% de probabilidad de rechazarla.

Paso 3: especificamos el nivel de significación α

$\alpha = .05$ (i.e., 5%)

Aun cuando H_0 sea cierta, asumir un riesgo del 5% de rechazar H_0

Qué poco frecuentes o improbables deben ser los resultados de nuestro estudio (conforme a la H_0) para que la rechacemos

Resumen

- La prueba estadística de hipótesis puede describirse como un procedimiento de siete pasos. Los primeros tres pasos son los siguientes:
  - Paso 1: describir la información disponible, los supuestos estadísticos y el parámetro poblacional que se considera.
  - Paso 2: especificar las hipótesis nula y alternativa.
  - Paso 3: especificar el nivel de significación alfa (α).

Prueba de hipótesis, parte 2

En el **paso 4** de nuestro procedimiento de prueba de hipótesis, debemos seleccionar el **estadístico de prueba** por utilizar y tenemos que describir su **distribución muestral** bajo el supuesto de que la hipótesis nula es verdadera. Como el parámetro de interés es la diferencia entre dos proporciones, el estadístico de prueba T está dada por la diferencia entre las dos proporciones muestrales, dividida por el error estándar calculado de esta diferencia muestral bajo la hipótesis nula. Aquí el denominador se calcula mediante una expresión que involucra el cálculo combinado, $\hat{p}$, de la proporción común para ambos grupos que surgiría bajo la hipótesis nula.

Paso 4: seleccionar el estadístico de prueba y describir su distribución muestral bajo la hipóteis nula

$$T = \frac{\hat{p}_M - \hat{p}_W}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_M} + \frac{1}{n_W}\right)}} \qquad \hat{p} = \frac{n_M \hat{p}_M + n_W \hat{p}_W}{n_M + n_W}$$

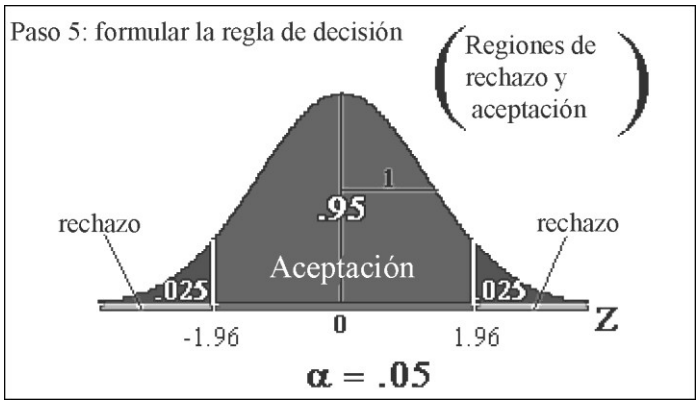
Preguntas de estudio (P12.2)

El cálculo combinado de la proporción común para dos grupos es un promedio ponderado de las dos proporciones muestrales, donde las ponderaciones son los tamaños de las muestras utilizados para calcular cada proporción. Las proporciones del tamaño de las muestras y sus correspondientes tamaños muestrales son 0,7972 y 1667, para los hombres, y 0,256 y 425 para las mujeres, respectivamente.

1. Realice el cálculo combinado a partir de la información anterior. (Necesitará una calculadora para obtener su respuesta).

La distribución muestral de este estadístico de prueba según la hipótesis nula es aproximadamente la distribución normal estándar, con media 0 y desvío estándar 1.

En el **paso 5** formulamos la regla de decisión que divide los resultados posibles del estadístico de prueba en regiones de aceptación y rechazo. Como nuestro estadístico de prueba tiene aproximadamente la distribución normal estándar o la distribución Z bajo la hipótesis nula, las regiones de aceptación y rechazo se especificarán como intervalos a lo largo del eje Z bajo la curva de esta distribución. En especial, debido a que nuestra hipótesis alternativa es de dos colas y puesto que nuestro nivel de significación es 0,05, estas dos regiones son las que se muestran aquí mediante las líneas rojas y verdes. (Nota: las líneas rojas están en las áreas de las colas < -1,96 y > 1,96; la línea verde está entre –1,96 y 1,96). El área bajo la curva normal estándar que se encuentra arriba del intervalo, descrita como la región de aceptación, es 0,95. El área bajo la curva en cada cola de la distribución, identificada como las regiones de rechazo, es 0,025. La suma de estas dos áreas es 0,05, que es nuestro nivel de significación elegido. El -1,96 del lado izquierdo bajo la curva es el punto porcentual 2,5 de la distribución normal estándar, y el 1,96 del lado derecho bajo la curva es el punto porcentual 97,5.



Nuestra regla de decisión puede describirse ahora de la siguiente manera: si el valor del **estadístico de**

prueba T calculado a partir de nuestros datos cae en la región de rechazo, rechazamos la hipótesis nula a favor de la hipótesis alternativa. Sin embargo, si el valor de estudio observado cae en la región de aceptación, no rechazamos la hipótesis nula.

El **paso 6** de nuestro proceso simplemente requiere que calculemos el valor del estadístico de prueba T a partir de los datos observados. Al valor calculado lo llamaremos **T*** para distinguirlo de del estadístico de prueba T. Nuevamente, estos son los resultados de la muestra:

Paso 6: calcular el estadístico de prueba utilizando los datos observados

$$T = \frac{\hat{p}_M - \hat{p}_W}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_M}+\frac{1}{n_W}\right)}} \qquad \hat{p} = \frac{n_M\hat{p}_M + n_W\hat{p}_W}{n_M + n_W}$$

Calculado $T = T^*$

$\hat{p} = .6874$ (hombres y mujeres)

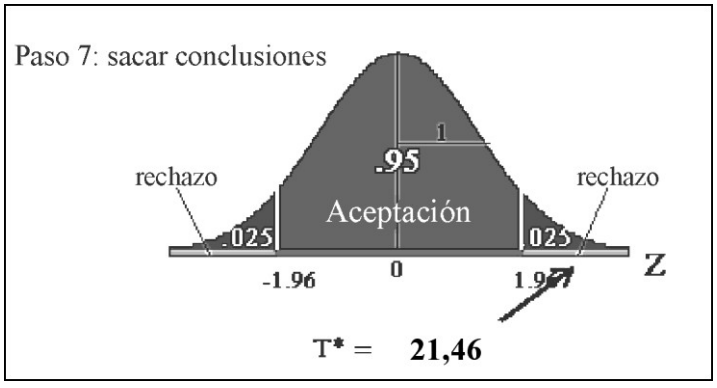
$\hat{p}_M = .7972$ $n_M = 1667$ (hombres)

$\hat{p}_W = .2565$ $n_W = 425$ (mujeres)

Al sustituir la información muestral en la fórmula de T, nuestro valor calculado T* resulta ser 21,46.

$$T^* = \frac{.7972 - .2565}{\sqrt{.6874\ (1 - .6874)\ \left(\frac{1}{1667}+\frac{1}{425}\right)}} = \mathbf{21,46}$$

Finalmente, en el **paso 7**, utilizamos nuestro estadístico de prueba calculado para sacar conclusiones sobre nuestra prueba de significación. En este ejemplo, el estadístico de prueba calculado cae en el extremo de la cola derecha que corresponde a la región de rechazo porque es mucho más grande que 1,96.



En consecuencia, rechazamos la hipótesis nula y llegamos a la conclusión de que tenemos una

diferencia estadísticamente significativa entre las dos proporciones al nivel de significación 0,05. Por lo tanto, concluimos que, de todos aquellos que podrían haber estado a bordo del Titanic, los hombres tenían mayor probabilidad de morir que las mujeres.

Resumen

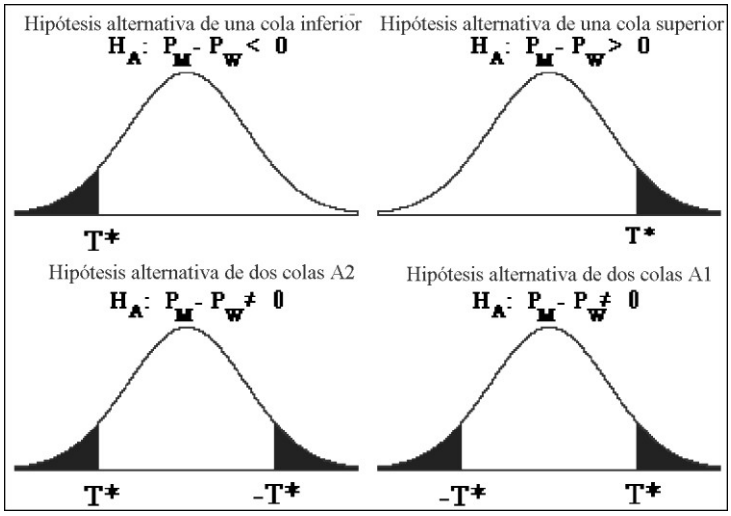
- La prueba estadística de hipótesis puede describirse como un procedimiento de siete pasos. Los primeros tres pasos son los siguientes:
  - Paso 1: describir la información disponible, los supuestos estadísticos y el parámetro poblacional que se considera.
  - Paso 2: especificar las hipótesis nula y alternativa.
  - Paso 3: especificar el nivel de significación alfa (α).
  - Paso 4: seleccionar el estadístico de prueba y describir su distribución muestral bajo la hipótesis nula.
  - Paso 5: formular la regla de decisión en términos de regiones de rechazo y aceptación bajo la hipótesis nula.
  - Paso 6: calcular el estadístico de prueba utilizando los datos observados.
  - Paso 7: sacar conclusiones, es decir, rechazar o no la hipótesis nula al nivel de significación alfa.

Prueba de hipótesis - El valor p

Hemos descubierto que la diferencia entre las proporciones muestrales de los hombres y las mujeres que murieron en el Titanic era estadísticamente significativa al nivel de significación 0,05. En especial, el valor calculado del estadístico de prueba cayó en el extremo de la cola derecha de la región de rechazo. Esto nos dice que si la hipótesis nula fuese verdadera, los resultados observados hubiesen tenido menos de 5% de probabilidad de ocurrir. Es decir, los resultados eran bastante improbables según la hipótesis nula.

Podríamos preguntarnos, entonces, exactamente qué tan improbables, o qué poco frecuentes, eran los resultados observados bajo la hipótesis nula. ¿También tenían una probabilidad menor a 1%, o menor a 0,1%, o incluso menor? La respuesta a estas preguntas es el **valor p**, el cual da la probabilidad de obtener el valor del estadístico de prueba que hemos calculado o un valor más extremo si la hipótesis nula es verdadera.

Supongamos, como en nuestro ejemplo, que el estadístico de prueba tiene la distribución normal estándar bajo la hipótesis nula. Para obtener el valor p debemos determinar un área bajo esta curva. Aquí mostramos cuatro áreas distintas que corresponden al lugar donde el valor calculado T* cae bajo la curva e indican si la hipótesis alternativa es de una o de dos colas (en otras palabras, unilateral o bilateral, respectivamente).

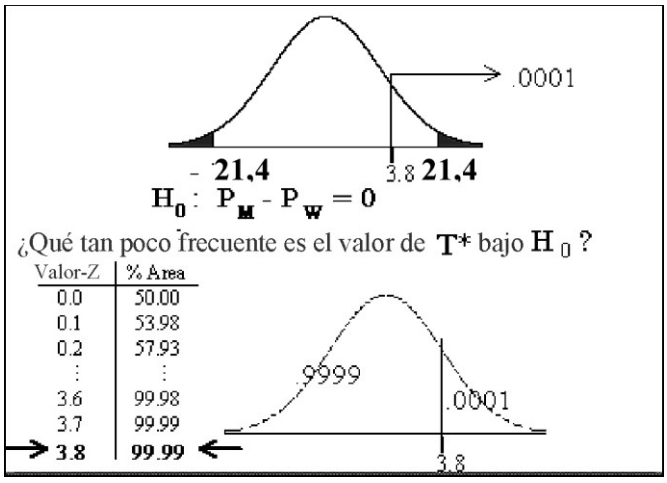


Si la hipótesis alternativa es una hipótesis unilateral superior, entonces el valor p es el área bajo la curva a la derecha de T* (distribución arriba a la derecha en la figura anterior). Si la hipótesis alternativa es una hipótesis unilateral inferior, entonces el valor p es el área bajo la curva a la izquierda de T* (distribución arriba a la izquierda en la figura anterior). Si la hipótesis alternativa es bilateral y T* cae en la cola derecha bajo la curva, entonces el valor p es la suma de las áreas bajo la curva a la derecha de T* y a la izquierda de -T*. Si la hipótesis alternativa es bilateral y T* cae en la cola izquierda bajo la curva, entonces el valor p es la suma de las áreas bajo la curva a la izquierda de T* y a la derecha de -T*. El valor p da el área bajo la curva que muestra la probabilidad de los resultados del estudio conforme a la hipótesis nula.

Preguntas de estudio (P12.3)

1. ¿Cuál de los cuatro escenarios anteriores corresponde al valor p para nuestro ejemplo del Titanic? (Sugerencia: T* = 12, H_A es de dos colas).
2. Para obtener el valor p para una H_A de dos colas, ¿por qué no es necesario calcular 2 áreas bajo la curva normal?

Ahora veamos qué poco frecuente es el valor de nuestro estadístico de prueba bajo la hipótesis nula. El valor del estadístico de prueba es 21,46, entonces debemos encontrar el área bajo la curva normal a la derecha del valor 21,46, y a la izquierda de -21,46. Una forma de determinar esta área es utilizar un cuadro de puntos porcentuales de la distribución normal estándar o de la distribución Z. En uno de dichos cuadros, como se muestra en la figura bajo este párrafo, el punto porcentual más alto es 3,8, correspondiente al punto porcentual 99,99. Aunque no podemos encontrar el área a la derecha de 21,46 bajo la curva normal en forma exacta, podemos decir que esta área es menor que 0,0001, lo que es claramente un valor muy pequeño.



Preguntas de estudio (P12.3) (continuación)

3. Si nuestra hipótesis alternativa hubiese sido de una cola, ¿cuál sería nuestro valor p?
4. Ya que nuestra hipótesis alternativa, en realidad, fue de dos colas, ¿cuál es nuestro valor p?
5. Sobre la base de su respuesta a la pregunta 2, ¿ha ocurrido un evento poco frecuente bajo la hipótesis nula?
6. En función del valor p aquí, ¿qué conclusión debería sacar sobre si la prueba de hipótesis es significativa o no?

Los valores p se utilizan a menudo como una forma alternativa para sacar conclusiones sobre una prueba de hipótesis, en lugar de especificar un nivel de significación fijo antes de calcular el estadístico de prueba. Si el valor p es suficientemente pequeño, de forma que ha ocurrido un evento poco frecuente, entonces rechazamos la hipótesis nula. Si el valor p no es pequeño, entonces no rechazaríamos la hipótesis nula.

Entonces, ¿qué tan pequeño debe ser el valor p para que nuestros resultados se consideren poco frecuentes? La respuesta aquí depende básicamente del nivel de significación alfa (α) que deseamos emplear. Una elección convencional de alfa es 0,05, aunque otra variación frecuente es 0,01. Por lo tanto, si el valor p es <0,05 ó <0,01, entonces los resultados de la prueba se consideran, por lo general, lo suficientemente poco frecuentes como para rechazar la hipótesis nula a favor de la hipótesis alternativa.

Preguntas de estudio (P12.3) (continuación)

Si su nivel de significación fue 0,05, ¿qué conclusiones sacaría sobre la hipótesis nula para los siguientes valores p?

7. a) P > 0,01 b) P = 0,023
8. c) P < 0,001 d) P = 0,54 e) P = 0,0002

Si su nivel de significación fue 0,001, ¿qué conclusiones sacaría sobre la H_0 para los siguientes valores p?

9. a) P > 0,01 b) P = 0,023
10. c) P < 0,001 d) 0,01 < P = 0,05 e) P = 0,0002

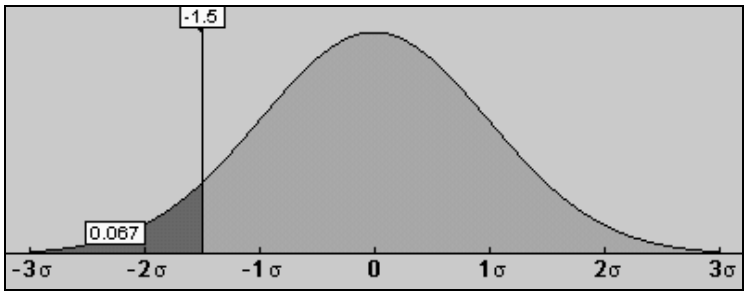
Resumen

- El valor p describe qué tan improbables, o qué poco frecuentes, son los resultados observados del estudio bajo la hipótesis nula.
- Para las hipótesis alternativas de una cola, el valor p está determinado por el área en la cola de la distribución, más allá de los estadísticos de pruebas calculados (es decir, a la derecha y a la izquierda) bajo la hipótesis nula.
- Para las hipótesis alternativas de dos colas, el valor p es el doble del área en la cola de la distribución, más allá del estadístico de prueba calculado bajo la hipótesis nula.
- El valor p se utiliza a menudo como una forma alternativa para sacar conclusiones sobre una hipótesis nula, en lugar de especificar un nivel de significación antes de calcular los estadísticos de prueba.
- Si los investigadores consideran que el valor p es pequeño, digamos, menor que 0,05, 0,01, rechazamos la hipótesis nula a favor de la hipótesis alternativa.
- Si el valor p no se considera pequeño, generalmente mayor que 0,10, no rechazamos la hipótesis nula.

Valores Z y frecuencias relativas – La función de densidad normal

La mayoría de las pruebas estadísticas incluyen un cuadro que le permite relacionar los valores z con las frecuencias relativas en una **densidad normal**. Los cuadros siempre brindan esta información para la **densidad normal estándar**, de forma que el eje x de la densidad se marca en valores z. La herramienta de densidad normal con la que hemos trabajado brinda la misma información de manera más sencilla. Por ejemplo, para encontrar la frecuencia relativa de valores con valores z menores a -1,5, simplemente arrastre la marca izquierda hasta el valor z -1,5 y lea la frecuencia relativa en el recuadro que se encuentra abajo a la izquierda, 0,067.

(Nota: utilice el CD-ROM ActivEpi para usar la **herramienta de densidad normal** a fin de responder varias preguntas de ejemplo).



Resumen

- La curva de densidad normal estándar, para la cual normalmente se brindan cuadros en los libros de texto de estadística, relaciona los valores z con las frecuencias relativas.
- La herramienta de densidad normal suministra la misma información en forma directa.



Auto-evaluación (P12.4)

Complete los espacios vacíos.

1. Utilizamos **???** para evaluar si el parámetro poblacional es diferente del valor nulo.
2. Cuando se determina la precisión de una estimación puntual, **???** representa la variabilidad muestral.
3. **???** busca resultados poco frecuentes o improbables.
4. Observando los resultados más probables, **???** encuentra aquellos valores con respecto a los cuales estamos seguros que se encuentran cerca del parámetro poblacional.

Opciones
la prueba de hipótesis **la estimación de intervalos**

5. El **???** muestra el riesgo que estamos dispuestos a correr para rechazar la hipótesis nula cuando ésta es falsa.
6. La **???** puede ser de una cola superior, de una cola inferior o de dos colas.
7. Si el valor calculado del estadístico de prueba cae en la **???**, rechazamos la **???** y llegamos a la conclusión de que los resultados son **???** significativos.

Opciones
región de aceptación **hipótesis alternativa** **significativamente** **hipótesis nula**
región de rechazo **niveles de significación** **estadísticamente**

8. El **???** describe qué poco frecuentes, o qué improbables, son los resultados observados del estudio bajo la **???**
9. Si el valor p satisface la desigualdad P>0,30, deberíamos **???** la hipótesis nula.
10. Si el valor p satisface la desigualdad P<0,005, deberíamos rechazar la hipótesis nula al nivel de significación de **???** , pero no al nivel de **???**

Opciones
0,001 **0,01** **valor p** **hipótesis alternativa** **no rechazar**
hipótesis nula **rechazar** **nivel de significación**

12-3 Análisis simples (continuación)

Revisión sobre los intervalos de confianza

*Un **intervalo de confianza (IC)** brinda dos números, **L** y **U**, entre los cuales se encuentra el parámetro poblacional con un nivel de confianza especificado. Aquí describimos cómo calcular el IC del 95% de muestra grande para la diferencia entre dos proporciones.*

Intervalo de confianza para comparar dos proporciones

Ahora mostramos cómo calcular el intervalo de confianza para la diferencia entre dos proporciones mediante los datos del Titanic. Nuestro objetivo es utilizar nuestra información muestral para calcular dos

números, **L** y **U**, sobre los cuales podemos afirmar con cierto grado de confianza, digamos del 95%, que encierran el verdadero valor del parámetro. A continuación incluimos la fórmula para este intervalo de confianza del 95%:

INTERVALO DE CONFIANZA para $p_M - p_W$

Los datos: $\hat{p}_M = .797$ $\hat{p}_W = .256$

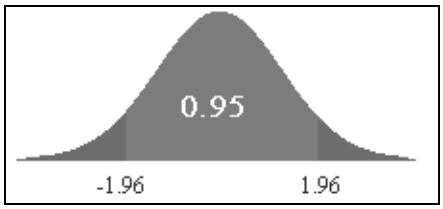
$n_M = 1667$ $n_W = 425$

Objetivo: $L < p_M - p_W < U$

es decir, confianza del 95%

$$(\hat{p}_M - \hat{p}_W) \pm 1.96 \sqrt{\frac{\hat{p}_M(1 - \hat{p}_M)}{n_M} + \frac{\hat{p}_W(1 - \hat{p}_W)}{n_W}}$$

El error estándar de la diferencia es la raíz cuadrada de la suma de las varianzas de las proporciones, donde cada varianza es de la forma ($\hat{p}$)(1- $\hat{p}$) / (tamaño de la muestra). El valor 1,96 es el punto porcentual 97,5 de la distribución normal estándar. Se elige este punto porcentual porque el área entre -1,96 y +1,96 bajo la curva normal es 0,95, lo que corresponde al nivel de confianza del 95% que especificamos. Aquí se emplea la distribución normal, porque la diferencia entre las dos proporciones muestrales tiene aproximadamente la distribución normal si los tamaños de las muestras en ambos grupos son razonablemente grandes, lo que de hecho sucede para estos datos. Éste es el motivo por el cual la fórmula del intervalo de confianza descrita aquí generalmente se denomina intervalo de confianza de **muestra grande**.



Preguntas de estudio (P12.5)

1. ¿Por qué la fórmula de error estándar utilizada aquí es diferente de la que se emplea cuando se prueba la hipótesis nula de diferencia cero entre las dos proporciones?

E.E. de prueba: $\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_M} + \frac{1}{n_W}\right)}$

Podemos calcular el intervalo de confianza para nuestros datos sustituyendo en la fórmula los valores de $\hat{p}_M$, $\hat{p}_W$, n_M y n_W.

INTERVALO DE CONFIANZA para $p_M - p_W$

Los datos: $\hat{p}_M = .797$ $\hat{p}_W = .256$

$n_M = 1667$ $n_W = 425$

$$(\hat{p}_M - \hat{p}_W) \pm 1.96 \sqrt{\frac{.797(1 - .797)}{1667} + \frac{.256(1 - .256)}{425}}$$

El error estándar resulta ser 0,0234. Los límites inferior y superior del intervalo del 95% son entonces 0,495 y 0,587, respectivamente. Por lo tanto, el intervalo de confianza de 95% para la diferencia entre las proporciones de hombres y mujeres que murieron en el Titanic está dado por el rango de valores entre 0,495 y 0,587.

$$(\hat{p}_M - \hat{p}_W) \pm (1.96)\ (.0234)$$

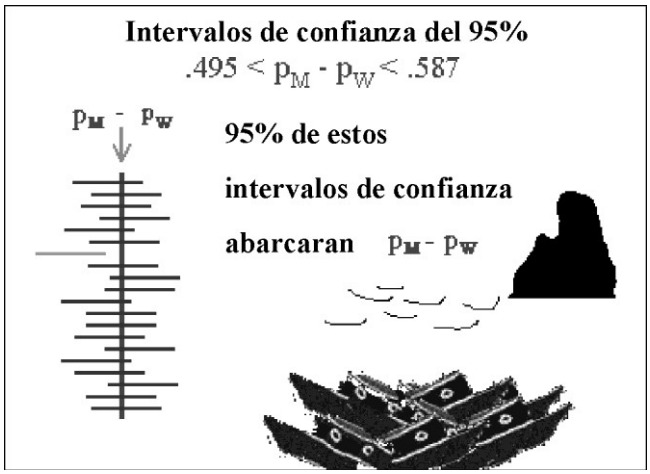
$$.495 < p_M - p_W < .587$$

Resumen

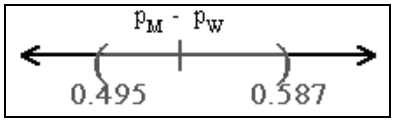
- Un **intervalo de confianza** (**IC**) ofrece dos números, **L** y **U**, entre los cuales se encuentra el parámetro poblacional con un nivel de confianza especificado.
- El IC de 95% de muestra grande para la diferencia entre dos proporciones está dado por la diferencia $\pm$1,96 multiplicada por el valor del error estándar para la diferencia calculada.
- El valor del error estándar está dado por la raíz cuadrada de la suma de las varianzas de cada proporción.

Interpretación del intervalo de confianza

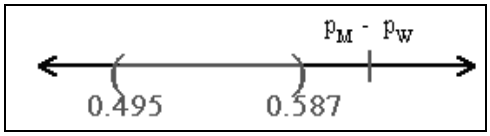
¿Cómo interpretamos este intervalo de confianza? Una correcta comprensión requiere que consideremos lo que podría suceder si repitiéramos el estudio, en este ejemplo, el viaje y el hundimiento del Titanic, varias veces. Si calculamos intervalos de confianza de 95% para los datos que surgen de cada repetición, entonces esperaríamos que aproximadamente 95% de estos intervalos de confianza abarquen la verdadera diferencia poblacional entre las proporciones.



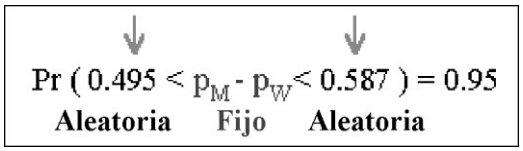
Esto equivale a decir que existe una probabilidad de 0,95 de que el intervalo entre 0,495 y 0,587 incluya la verdadera diferencia poblacional entre las proporciones.



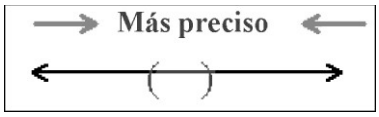
La verdadera diferencia podría caer, en realidad, fuera de este intervalo, pero sólo hay una probabilidad de 5% de que esto suceda.



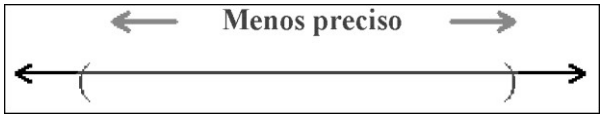
La declaración de probabilidad que describe el intervalo de confianza, que en el centro tiene el parámetro poblacional, $P_M - P_W$, sin ningún acento circunflejo, sugiere que este parámetro es una variable aleatoria. Esto no es verdad. El parámetro $P_M - P_W$ no varía en lo absoluto; tiene un único valor poblacional fijo. Los elementos aleatorios del intervalo son los límites 0,495 y 0,587, que se calculan a partir de los datos de la muestra y variarán de muestra en muestra.



En general, el intervalo de confianza es una medida de la precisión de un cálculo de algún parámetro de interés que, para nuestro ejemplo, es la diferencia entre dos proporciones poblacionales. Cuanto más acotado es el intervalo de confianza, más preciso es el cálculo.



Por el contrario, cuanto más amplio es el intervalo, menos preciso será el cálculo.



El caso extremo de falta total de precisión ocurriría para las medidas en diferencias (por ejemplo, la diferencia de riesgos y la diferencia de tasas de incidencia) donde el intervalo de confianza va desde menos infinito a infinito; para las medidas en razones (por ejemplo, el *odds ratio*, la razón de riesgos y la razón de densidades de incidencia), sería un intervalo de confianza desde cero hasta infinito; y para las proporciones, sería un intervalo de confianza de cero a 1.

Preguntas de estudio (P12.6)

1. Para un intervalo de confianza que va desde menos infinito hasta infinito, ¿qué seguros estamos de que el intervalo abarca el verdadero parámetro?
2. Si se obtuvieran intervalos de confianza de 90%, 95% y 99% para la diferencia entre dos proporciones basadas en los mismos datos, ¿qué IC sería el más amplio y cuál sería el más acotado?
3. Suponga que dos conjuntos de datos distintos mostraron intervalos de confianza de 95% para la diferencia entre dos proporciones. ¿Qué conjunto de datos (A o B siguientes) muestra el cálculo más preciso?

   Conjunto de datos A: $0{,}49 < p1 - p2 < 0{,}58$

   Conjunto de datos B: $0{,}40 < p1 - p2 < 0{,}52$

Resumen

- El IC de 95% puede interpretarse mediante la declaración de probabilidad P(L < el parámetro < U) = 0,95.
- Si el IC se calcula para varias repeticiones del mismo estudio, esperaríamos que aproximadamente 95% de los IC abarcaran el verdadero parámetro poblacional.
- Los elementos aleatorios del intervalo de confianza son los límites L y U.
- Es incorrecto suponer que el parámetro en el medio de una declaración de intervalo de confianza es una variable aleatoria.
- Cuanto mayor sea el nivel de confianza elegido, más amplio será el intervalo de confianza.

Debate: ¿los datos del Titanic representan una población o una muestra?

Como se indicó anteriormente, nuestro ejemplo describe los datos de supervivencia para todos los pasajeros hombres y mujeres del Titanic, el transatlántico "imposible de hundir" que chocó contra un iceberg y naufragó en 1912. Ya que estos datos consideran a los pasajeros de ambos sexos, se puede argumentar que las

proporciones que se comparan son, en verdad, proporciones poblacionales, de modo que no es apropiado realizar una prueba estadística de significación o calcular un intervalo de confianza con estos datos. Sin embargo, una respuesta a este argumento es que los 1667 hombres y las 425 mujeres representan una muestra de quienes cumplían los requerimientos para ser elegidos para el viaje del Titanic, mientras que la diferencia poblacional entre las proporciones hace referencia a las proporciones de todas las personas que reunían los requisitos para el viaje.

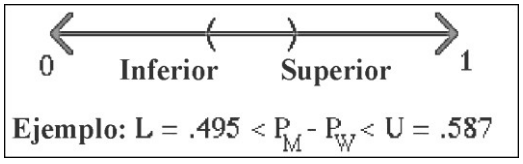
Estos dos argumentos son debatibles y, según nuestro punto de vista, no hay una razón clara que nos lleve a la conclusión de que uno u otro argumento es el correcto. De hecho, a menudo se presentan debates similares cuando se analizan datos de alguna investigación epidemiológica sobre un brote. Por ejemplo, cuando se busca la fuente de un brote de diarrea a partir de un picnic, las pruebas estadísticas generalmente se realizan sobre datos que representan a todas las personas que asistieron a tal evento. Dichas pruebas se justifican sólo si los datos analizados se consideran una muestra, en lugar de una población.

Auto-evaluación (P12.7)

Complete los espacios vacíos.

1. Un intervalo de confianza de 95% de muestra grande para la diferencia entre dos proporciones suma y resta 1,96 de la diferencia calculada entre las dos proporciones, multiplicado por el **???** de la diferencia calculada.
2. El ejemplo del intervalo de confianza que se muestra en la próxima página no contiene el valor nulo para la **???** de las dos proporciones.
3. El **???** dentro de un intervalo de confianza tiene un único valor fijo y no varía en lo absoluto.

Opciones
nivel de confianza **diferencia** **media calculada** **error estándar calculado**
varianza calculada **parámetro poblacional** **razón**



4. El **???** de un intervalo de confianza puede variar de una muestra a otra.
5. Para un intervalo de confianza del **???**, hay una probabilidad de 0,95 de que el intervalo entre los límites superior e inferior incluya el verdadero parámetro poblacional.
6. El verdadero parámetro poblacional podría, en verdad, caer fuera de este intervalo, pero sólo hay una probabilidad del **???** de que esto suceda.

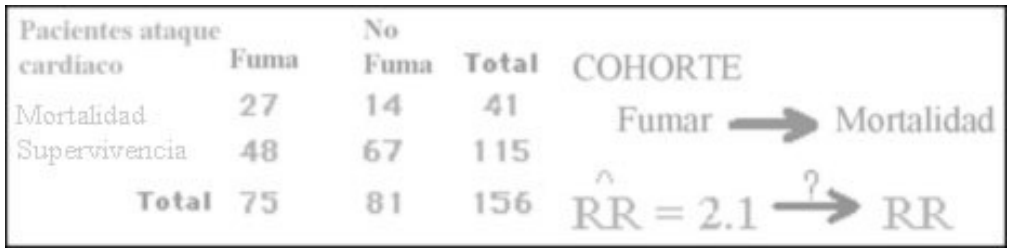
Opciones
5% **95%** **97,5%** **nivel de confianza** **parámetro poblacional** **límite superior**

12-4 Análisis simples (continuación)

ESTUDIOS DE COHORTE QUE INVOLUCRAN RAZONES DE RIESGOS

Prueba de hipótesis para el análisis simple en estudios de cohorte

Volvemos a los datos de un estudio de cohorte que evaluaba si dejar de fumar después de un ataque cardíaco reduciría el riesgo que tiene una persona de morir. La medida de efecto en este estudio era una razón de riesgos y su valor fue 2,1. ¿Qué podemos decir sobre la razón de riesgos de la población basada en la razón de riesgos obtenida de nuestra muestra?



Deseamos saber si tenemos pruebas en la muestra de que la razón de riesgos es estadísticamente diferente del valor nulo. Es decir, queremos realizar una **prueba de hipótesis** para ver si la razón de riesgos es significativamente diferente de 1. La hipótesis nula que se somete a prueba indica que la razón de riesgos de la población es 1. La hipótesis alternativa lógica aquí es que la razón de riesgos es >1, ya que antes de observar los datos los investigadores estaban interesados en saber si continuar fumando tendría más probabilidad de afectar la mortalidad que dejar de fumar.

Prueba de hipótesis	H_0 : RR = 1
	H_A : RR > 1

Como la razón de riesgos es la razón de incidencias acumuladas para el grupo expuesto (IA_1), dividido por las incidencias acumuladas para el grupo no expuesto (IA_0), en forma similar podemos expresar la hipótesis nula en términos de la diferencia entre las incidencias acumuladas de la población, como se muestra a continuación:

$H_0: RR = \frac{IA_1}{IA_0} = 1$

$H_0: P_1 - P_0 = 0$

$H_0: ROR = 1$

} Equivalente

Puesto que la razón de riesgos es igual a 1 si y sólo si el *odds ratio* de riesgo es igual a 1, también podemos expresar, en forma equivalente, la hipótesis nula en términos del *odds ratio* de riesgo.

Debido a que la incidencia acumulada es una proporción, la versión de la incidencia acumulada de la hipótesis nula implica que nuestra prueba sobre la razón de riesgos es equivalente a probar una hipótesis sobre la diferencia entre dos proporciones: H_0: $p_1 - p_0 = 0$.

El estadístico de prueba es la diferencia entre las dos incidencias acumuladas calculadas, dividida por el valor del error estándar de esta diferencia, según la hipótesis nula de que la razón de riesgos es igual a 1. Ya que los tamaños de las muestras en ambos grupos son razonablemente grandes, este estadístico de prueba tiene aproximadamente una distribución normal estándar según la hipótesis nula.

$$T = \frac{\hat{p}_1 - \hat{p}_0}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_0}\right)}} \quad \text{donde } p_1 = IA_1,\ p_0 = IA_0,\ p = \text{IA calculadas}$$

El valor calculado del estadístico de prueba se obtiene sustituyendo las incidencias acumuladas calculadas y los tamaños de muestra correspondientes en la fórmula del estadístico de prueba, como se indica a continuación. El valor resultante es 2,65.

$$T^* = \frac{(27/75) - (14/81)}{\sqrt{\frac{41}{156}\left(1-\frac{41}{156}\right)\left(\frac{1}{75}+\frac{1}{81}\right)}} = 2.65$$

El valor p para esta prueba se obtiene, entonces, encontrando el área en la cola derecha de la distribución normal estándar por arriba del valor calculado de 2,65. El valor p exacto resulta ser 0,0040. Como el valor p de 0,0040 está muy por debajo del nivel de significación convencional de 0,05, rechazamos la hipótesis nula y llegamos a la conclusión de que la razón de riesgos es significativamente mayor que el valor nulo de 1. En otras palabras, hemos descubierto que entre los pacientes que fuman y sufrieron un ataque cardíaco, los que continúan fumando tienen un riesgo significativamente más alto de morir que quienes dejaron de fumar después de su ataque cardíaco.

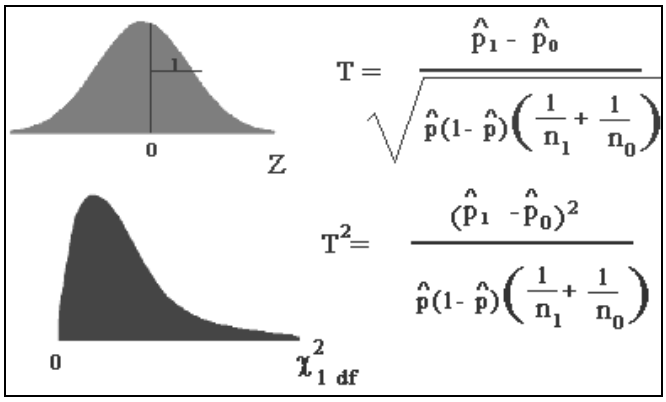
Resumen

- Cuando se prueba la hipótesis sobre una razón de riesgos (RR) en un estudio de cohorte con incidencias acumuladas, la hipótesis nula se puede expresar en forma equivalente como RR = 1, $IA_1 - IA_2 = 0$ ó ROR = 1, donde IA_1 e IA_2 son las incidencias acumuladas para los grupos expuesto y no expuesto, respectivamente.
- La hipótesis alternativa puede expresarse en términos de RR como RR ≠ 1, RR > 1 o RR < 1, si la hipótesis alternativa es de dos colas, de una cola superior o de una cola inferior, respectivamente.
- Para probar la hipótesis nula de que RR = 1, el estadístico de prueba es el mismo que el que se usa para comparar la diferencia entre dos proporciones.
- Suponiendo muestras grandes, el estadístico de prueba tiene aproximadamente una distribución N(0,1) bajo la H_0.

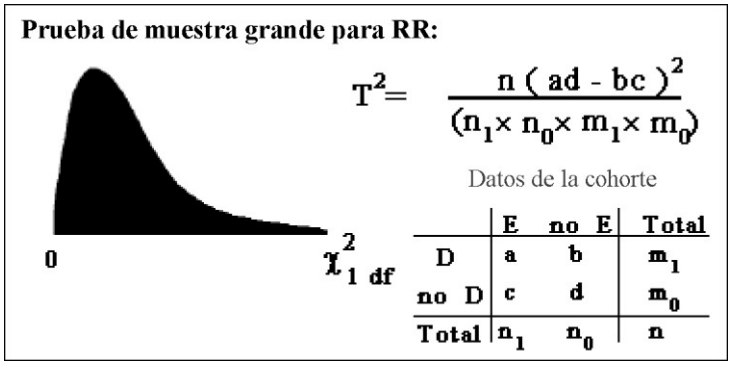
La versión chi cuadrado de la prueba de muestra grande

La prueba de muestra grande para una razón de riesgos puede realizarse mediante la **distribución normal** o la **distribución chi cuadrado**. La razón para esta equivalencia es que si se eleva al cuadrado una variable normal estándar Z, entonces Z al cuadrado tiene una distribución chi cuadrado con 1 grado de libertad.

Más específicamente, para nuestro estudio de mortalidad de pacientes que sufrieron ataque cardíaco, presentamos aquí el estadístico de prueba que hemos descrito con anterioridad; éste sigue una distribución normal estándar bajo la hipótesis nula de que la razón de riesgos es igual a 1. El cuadrado de este estadístico se muestra al lado de su correspondiente distribución chi cuadrado:



Aplicando un poco de álgebra, podemos reescribir el estadístico en términos de las frecuencias de celda a, b, c y d del cuadro general de 2 x 2 que resume la información sobre exposición y enfermedad en un estudio de cohorte que calcula la incidencia acumulada.



Para el estudio de mortalidad de pacientes que sufrieron un ataque cardíaco, los valores de las frecuencias de celda se muestran después de este párrafo. Al sustituir estos valores en la fórmula del estadístico chi cuadrado, obtenemos el valor 7,04, que es el cuadrado del valor del estadístico de prueba que calculamos anteriormente ($2{,}65^2 \approx 7{,}04$).

$$\frac{156\,[(27)(67)-(48)(14)]^2}{(75)\,(81)\,(41)\,(115)} = 7.04$$

Estudio de Mortalidad

	E	no E	Total
D	27	14	41
no D	48	67	115
Total	75	81	156

¿Es significativa la versión chi cuadrado de nuestra prueba? La versión de la distribución normal fue significativa al nivel de significación de 0,01; por lo tanto, la versión chi cuadrado debería ser significativa.

Pero al comparar estas dos versiones de la prueba, debemos abordar un pequeño problema. Al ser el cuadrado de un estadístico Z, un estadístico chi cuadrado nunca puede ser negativo, pero un estadístico normal estándar puede ser positivo o negativo. Valores grandes de un estadístico chi cuadrado que pueden indicar un resultado significativo podrían ocurrir a partir de valores positivos o negativos grandes del estadístico normal. En otras palabras, si utilizamos un estadístico chi cuadrado para determinar la significación, automáticamente estamos realizando una prueba de una hipótesis alternativa de **dos colas**, aunque sólo estemos buscando valores grandes en la cola derecha de la distribución.

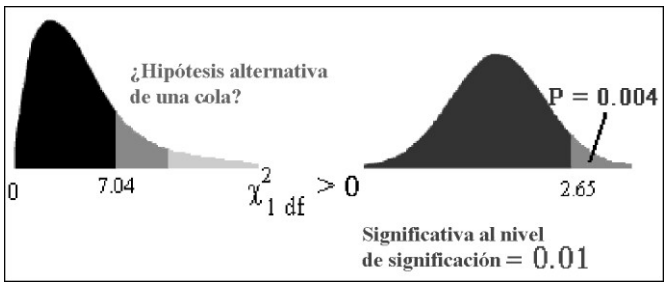
Preguntas de estudio (P12.8)

Los puntos porcentuales 0,99 y 0,995 de la distribución chi cuadrado con 1 grado de libertad son los valores 6,635 y 7,789, respectivamente.

1. ¿El valor chi cuadrado calculado de 7,04 cae en el 1% superior de la distribución chi cuadrado?
2. ¿El valor chi cuadrado calculado de 7,04 cae en el 0,5% superior de la distribución chi cuadrado?
3. ¿Una prueba de una hipótesis alternativa de dos colas para la RR sería significativa al nivel de significación de 0,01? (Nota: el estadístico de prueba calculado es 7,04).
4. ¿Una prueba de una hipótesis alternativa de dos colas para la RR sería significativa al nivel de significación de 0,005?

Entonces, si nuestro estadístico chi cuadrado nos permite evaluar una hipótesis alternativa de dos colas sobre la razón de riesgos, ¿cómo podemos evaluar una hipótesis alternativa de una cola? Una forma es simplemente realizar la versión de la distribución normal de la prueba como se ejemplificó anteriormente. La otra manera es dividir el área en la cola derecha de la curva chi cuadrado por la mitad.

Preguntas de estudio (P12.8) (continuación)



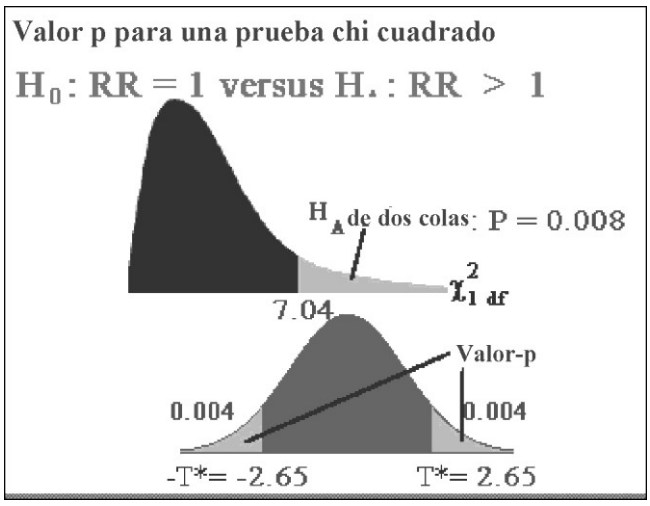
5. Utilizando los resultados anteriores, si RR > 1 en la prueba de una hipótesis alternativa de una cola superior, ¿dicha prueba sería significativa al nivel de significación de 0,01?
6. Si RR > 1 en la prueba de una hipótesis alternativa de una cola superior, ¿dicha prueba sería significativa al nivel de significación de 0,005?

Resumen

- Una forma alternativa pero equivalente de probar la significación de una razón de riesgos es utilizar una prueba de chi cuadrado.
- El cuadrado de una variable normal estándar tiene la distribución chi cuadrado con un grado de libertad.
- Podemos calcular el estadístico chi cuadrado en forma directa mediante una fórmula que involucra las frecuencias de celda de un cuadro de 2 x 2 para datos de cohorte que permiten el cálculo del riesgo.
- Cuando utilizamos un estadístico chi cuadrado a fin de determinar la significación, en realidad, estamos realizando la prueba de una hipótesis alternativa de dos colas.
- Podemos evaluar una hipótesis alternativa de una cola mediante el empleo de la versión de la prueba de distribución normal o valiéndonos de la distribución chi cuadrado para calcular un valor p.

El valor p para una prueba chi cuadrado unilateral

Ahora describimos cómo calcular el valor p cuando se utiliza una prueba chi cuadrado que involucra una hipótesis alternativa unilateral (de una cola) sobre una razón de riesgos. Ilustraremos este cálculo nuevamente mediante datos del estudio de mortalidad sobre el comportamiento con respecto al tabaquismo de los pacientes que sufrieron un ataque cardíaco. Para determinar el valor p, primero debemos calcular el área sombreada de rosa (sombreado más suave en la cola derecha de la distribución) bajo la distribución chi cuadrado por arriba del valor del estadístico chi cuadrado calculado de 7,04. En realidad, esta área muestra el valor p para una hipótesis alternativa bilateral (de dos colas). De hecho, es equivalente al área combinada en las dos colas de la distribución normal correspondiente, definida por el valor calculado de 2,65.



Para la hipótesis alternativa de una cola superior que indica que la razón de riesgos de la población es mayor que 1, anteriormente descubrimos que el valor p es solamente el área en la cola derecha de la curva normal por arriba de 2,65. Esta área de la cola derecha es sólo la mitad del área correspondiente en la cola derecha de la curva chi cuadrado por arriba de 7,04. En consecuencia, el área en la cola derecha de la curva chi cuadrado brinda el doble del valor p para una hipótesis alternativa de una cola. Por lo tanto, esta área debe dividirse por 2 para poder obtener un valor p de una cola.

$$P = \frac{.008}{2} = .004$$

Preguntas de estudio (P12.9)

1. Suponga que la hipótesis alternativa en nuestro ejemplo anterior era una hipótesis alternativa de una cola **inferior** y que el valor calculado del estadístico chi cuadrado es 7,04, correspondiente a un estadístico Z de +2,65. ¿Cuál es el valor p correspondiente?
2. En función de la respuesta a la pregunta anterior, ¿qué conclusión sacaría sobre la hipótesis nula de que la RR es igual a 1?
3. Suponga que la hipótesis alternativa era una hipótesis alternativa de una cola **inferior** y que el valor calculado del estadístico chi cuadrado es 7,04, pero esta vez correspondiente a un estadístico Z de -2,65. ¿Cuál es el valor p correspondiente?
4. En función de la respuesta a la pregunta anterior, ¿qué conclusión sacaría sobre la hipótesis nula que indica que la RR es igual a 1?

Resumen

- Cuando se utiliza una prueba chi cuadrado de una hipótesis alternativa de dos colas sobre una razón de riesgos, el valor p se obtiene como el área debajo de la curva chi cuadrado a la derecha del estadístico de prueba calculado.
- Si la hipótesis alternativa es de una cola, por ejemplo, RR > 1, entonces el valor p es la mitad del área en la cola derecha de la curva chi cuadrado.

Realización de pruebas cuando los tamaños de las muestras son pequeños

Este cuadro muestra los resultados hipotéticos de un ensayo clínico aleatorizado de cinco meses que comparaba una nueva droga antiviral para el herpes zoster con la droga antiviral convencional "Valtrex". Sólo 13 pacientes se incluyeron en el ensayo, que fue un estudio piloto para un ensayo clínico más grande. La razón de riesgos calculada para estos datos es 3,2, lo cual indica que, en la muestra de 13 pacientes, la nueva droga tuvo 3,2 veces más éxito que la droga convencional.

Ensayo clínico piloto en pacientes con herpes zoster

	Droga		Total
	Nueva	Convencional	
Éxito	4	2	6
Fracaso	1	6	7
Total	5	8	13

droga nueva: $4/5 = 0.80$

droga convencional: $2/8 = 0.25$

$\Rightarrow \hat{RR} = 3.2$

¿Esta razón de riesgos es significativamente distinta del valor nulo de 1? Aquí consideramos un tamaño pequeño de muestra, por lo que no es adecuada una prueba de hipótesis de muestra grande. (Nota: es debatible cuál es el alcance del término "grande", pero normalmente se requiere que el tamaño de la muestra para una proporción sea lo suficientemente grande [por ejemplo, >25] como para que la muestra tenga una distribución normal). Sin embargo, hay una prueba estadística para datos escasos llamada **prueba exacta de Fisher** que podemos utilizar en este caso. La prueba exacta de Fisher es adecuada para un cuadro de 2 x 2 que relaciona una variable de exposición dicotómica con una variable de enfermedad dicotómica. Para utilizar la prueba exacta de Fisher, debemos suponer que los valores en los márgenes de los cuadros son valores fijos antes del comienzo del estudio. Si tomamos este "supuesto de totales marginales fijos", podemos ver que una vez que identificamos el número en cualquiera de las celdas del cuadro tales como los casos expuestos o la celda **a**, los números en las otras tres celdas se pueden determinar utilizando sólo las frecuencias en los totales marginales.

Preguntas de estudio (P12.10)

1. Determine las fórmulas para calcular los valores de **b**, **c** y **d** en términos de **a** y los valores marginales fijos.

Disposición general	E	no E	Total
D	4		6
no D			7
Total	5	8	13

2. Calcule los valores en las otras celdas del cuadro anterior conociendo sólo los valores marginales y el valor de una celda.

Para probar nuestra hipótesis sobre la razón de riesgos, debemos considerar sólo el resultado para una de las cuatro celdas del cuadro. Por ejemplo, nos gustaría conocer qué valores podríamos obtener en la celda **a** que tengan menor probabilidad de presentarse según la hipótesis nula de que la razón de riesgos o, en realidad, el *odds ratio* de riesgo, es igual a 1. En especial, quisiéramos determinar si el valor obtenido en la celda **a** para nuestro estudio, que resultó ser 4, es un evento lo suficientemente poco frecuente conforme a la hipótesis nula para que la rechacemos y lleguemos a la conclusión de que la razón de riesgos observada es estadísticamente significativa. Podemos responder esta pregunta calculando el valor p para nuestra prueba.

Preguntas de estudio (P12.10) (continuación)

Disposición general	E	no E	Total
D	5		6
no D			7
Total	5	8	13

3. Suponga que el valor de la celda **a** fuera 5 en lugar de 4, asumiendo que los totales marginales son fijos. ¿Cuáles serían los valores corregidos de **b**, **c** y **d** correspondientes?
4. ¿Cuál sería la razón de riesgos para el cuadro corregido?
5. ¿La razón de riesgos corregida se aleja más de la razón de riesgos nula que la razón de riesgos de 3,2 que en verdad se observó?
6. ¿Existen otros valores posibles para la celda **a** que también se alejarían más de la razón de riesgos nula que 3,2?
7. Sobre la base de las preguntas anteriores, ¿por qué quisiéramos calcular la probabilidad de obtener un valor de la celda **a** de 4 ó 5 bajo la hipótesis nula?

Para calcular el valor p de la prueba exacta de Fisher, debemos determinar la distribución de probabilidad de la frecuencia de la celda **a** según la hipótesis nula. Suponiendo totales marginales fijos, esta distribución se denomina **distribución hipergeométrica**. Las fórmulas para la distribución hipergeométrica y el valor p correspondiente de la prueba exacta de Fisher se describen en el recuadro que se encuentra después de esta actividad.

Preguntas de estudio (P12.10) (continuación)

Utilizando la distribución hipergeométrica, el valor p de una cola (prueba exacta de Fisher) para el estudio descrito en esta actividad es P(a = 4 ó 5|Rr = 1) = P(a = 4|RR = 1) + P(a = 5|RR = 1) = 0,0816 + 0,0046 = 0,0863.

8. ¿Cuáles son sus conclusiones sobre la hipótesis nula de que RR = 1?

Resumen

- La prueba exacta de Fisher brinda una prueba de significación para una razón de riesgos o un *odds ratio* cuando los datos son escasos.
- Para utilizar la prueba exacta de Fisher, debemos suponer que los valores en los márgenes de un cuadro de 2 x 2 son fijos antes del comienzo del estudio.
- Las frecuencias marginales de un cuadro de 2 x 2 no brindan ninguna información sobre la fuerza de la asociación.
- Para calcular el valor p de la prueba exacta de Fisher, debemos determinar la distribución de probabilidad de la frecuencia de la celda **a** bajo la hipótesis nula.
- Suponiendo totales marginales fijos, la celda **a** tiene **distribución hipergeométrica**.

Para calcular el valor p para esta prueba se encuentran disponibles programas informáticos.

Cómo realizar la prueba exacta de Fisher mediante la distribución hipergeométrica

Para calcular el valor p de la prueba exacta de Fisher debemos emplear la distribución de probabilidad de la frecuencia de celda "a" (es decir, la cantidad de casos expuestos) bajo la hipótesis nula de que la exposición no tiene efecto sobre la enfermedad.

Disposición de los datos con totales marginales fijos y con celda "a" = "j"

	E	**no E**	
D	j	$m_1 - j$	m_1
no D	$n_1 - j$	$n_0 - m_1 + j$	m_0
	n_1	n_0	

Si consideramos totales marginales fijos, esta distribución se denomina distribución hipergeométrica, la cual es la siguiente fórmula de probabilidad para una variable aleatoria A:

$$pr(A = j \mid H_0) = \frac{C_j^{n_1} C_{m_1 - j}^{n_0}}{C_{m_1}^{n_1 + n_0}} \text{ donde } C_s^r = \frac{r!}{s!(r-s)!}$$

denota la cantidad de combinaciones de *r* elementos tomados *s* por vez, y

r! = *r*(*r*-l)(*r*-2)...(3)(2)(1) denota el factorial de *r*,

s! y (*r*-*s*)! denotan el factorial de *s* y el factorial de (*r*-*s*) , respectivamente.

Surge entonces que el valor p exacto para una prueba de H_0 versus H_A mediante la fórmula hipergeométrica anterior está dado por:

$$P = pr(A \geq a \mid H_0) = \sum_{j=a}^{\min(n_1, m_1)} pr(A = j \mid H_0)$$

Ilustramos el cálculo de este valor p empleando los datos del ensayo clínico piloto en un nuevo tratamiento para pacientes con herpes zoster. Aquí están los datos nuevamente:

Datos del ensayo clínico sobre pacientes con herpes zoster (continuación).

	Droga nueva	Droga convencional	
Éxito	4	2	6
Fracaso	1	6	7
	5	8	13

Por lo tanto, el valor p se calcula como la probabilidad de que la celda **a** sea 4 ó 5 según la hipótesis nula de que RR = l:

$$P = pr(A \geq 4 \mid H_0) = \sum_{j=4}^{5} pr(A = j \mid H_0)$$

$$= \frac{C_4^5 C_2^8}{C_6^{13}} + \frac{C_5^5 C_1^8}{C_6^{13}} = 0{,}0816 + 0{,}0047 = 0{,}0863$$

Ya que p es mayor que 0,05, no podríamos rechazar la hipótesis nula de que RR = 1 y llegaríamos a la conclusión de que no hay un efecto significativo de la nueva droga al nivel de significación de 0,05.

12-5 Análisis simples (continuación)

ESTUDIOS DE COHORTE QUE INVOLUCRAN RAZONES DE RIESGOS (continuación)

Versión de muestra grande de la prueba exacta de Fisher - La prueba de Mantel-Haenszel

Consideremos nuevamente los datos del estudio de mortalidad sobre el comportamiento con respecto al tabaquismo de los pacientes que sufrieron un ataque cardíaco. Para estos datos hemos descrito anteriormente un estadístico chi cuadrado de muestra grande a fin de probar la hipótesis nula de que la razón de riesgos es igual a 1.

Pacientes ataque cardíaco	Fuma	No Fuma	Total
Mortalidad	27	14	41
Supervivencia	48	67	115
Total	75	81	156

$$\chi^2 = \frac{n(ad - bc)^2}{(n_1 \times n_0 \times m_1 \times m_0)}$$

$$H_0: RR = 1$$

Debido a que los datos de mortalidad involucran una muestra grande, no es necesario utilizar la prueba exacta de Fisher para estos datos. Sin embargo, aún podríamos calcular el estadístico de la prueba exacta de Fisher. Para realizar la prueba exacta de Fisher, suponemos que las frecuencias en los márgenes del cuadro están fijas y luego calculamos la probabilidad bajo la hipótesis nula de obtener un valor de la celda **a** al menos tan grande como el valor de 27 que se obtuvo en realidad. Por lo tanto, computaríamos y sumaríamos 15 valores de probabilidad, desde a = 27 hasta a = 41, para obtener el valor p de la prueba exacta de Fisher.

Prueba exacta de Fisher

calcular $P(a \geq 27|H_0)$

$P(27|RR=1) + P(28|RR=1) + P(29|RR=1) +$
$P(30|RR=1) + P(31|RR=1) + P(32|RR=1) +$
$P(33|RR=1) + P(34|RR=1) + P(35|RR=1) +$
$P(36|RR=1) + P(37|RR=1) + P(38|RR=1) +$
$P(39|RR=1) + P(40|RR=1) + P(41|RR=1)$

Aunque dicho cálculo puede lograrse con un programa informático adecuado, a menudo se utiliza una aproximación a la prueba exacta de Fisher de muestra grande más conveniente, cuando las frecuencias de las celdas en el cuadro de 2 x 2 son moderadamente grandes. Este enfoque de muestra grande se denomina prueba de **Mantel-Haenszel** (**MH**) para el análisis simple. A continuación se incluye el estadístico de Mantel-Haenszel. Este estadístico tiene, aproximadamente, una distribución chi cuadrado con un grado de libertad según la hipótesis nula de que la razón de riesgos es igual a 1. En consecuencia, el valor p para una hipótesis alternativa de una cola mediante este estadístico se obtiene de la forma normal, encontrando el área bajo la distribución chi cuadrado por arriba del estadístico de prueba calculado y luego dividiendo esta área por 2.

Prueba de Mantel-Haenszel **(MH)**

$$\chi^2(MH) = \frac{(n-1)(ad - bc)^2}{(n_1 \times n_0 \times m_1 \times m_0)} \qquad \text{approx. } \chi^2_{1\ df}$$

según H_0: RR=1

Si comparamos el estadístico chi cuadrado de muestra grande de la prueba exacta de Fisher con el estadístico chi cuadrado de muestra grande descrita anteriormente para comparar dos proporciones, vemos que estes dos estadísticos de prueba son asombrosamente similares. De hecho, difieren sólo en que el estadístico para aproximar la prueba exacta de Fisher contiene n-1 en el numerador, mientras que la versión anterior de muestra grande contiene n en el numerador. Cuando **n** es grande, utilizar **n** o **n**-1 en el numerador tendrá un efecto pequeño en el estadístico chi cuadrado calculado.

Prueba de Mantel-Haenszel **(MH)**

$$\chi^2(MH) = \frac{(n-1)(ad - bc)^2}{(n_1 \times n_0 \times m_1 \times m_0)}$$

$$\chi^2 = \frac{n(ad - bc)^2}{(n_1 \times n_0 \times m_1 \times m_0)}$$

En nuestro supuesto de los datos de mortalidad, por ejemplo, el valor de la aproximación chi cuadrado a la prueba exacta de Fisher es igual a 6,99. Anteriormente se mostró que el estadístico chi cuadrado de muestra grande para comparar dos proporciones era 7,04. Claramente, los dos estadísticos chi cuadrado son muy cercanos, aunque no exactamente iguales. Los correspondientes valores p de una cola son 0,0040 y 0,0041, respectivamente, es decir, básicamente iguales.

Pacientes ataque cardiaco	Fuma	No Fuma	Total
Mortalidad	27	14	41
Supervivencia	48	67	115
Total	75	81	156

Prueba de Mantel-Haenszel (MH)

$$\chi^2(MH) = \frac{(n-1)(ad - bc)^2}{(n_1 \times n_0 \times m_1 \times m_0)} = \frac{(156-1)(27x67-14x48)^2}{(75x81x41x115)}$$

$$= 6.99 \ (P = .0040)$$

$$\chi^2 = \frac{n(ad - bc)^2}{(n_1 \times n_0 \times m_1 \times m_0)} = \frac{(156)(27x67-14x48)^2}{(75x81x41x115)}$$

$$= 7.04 \ (P = .0041)$$

*Nota: en el recuadro anterior se brindan valores p de una cola.

Este ejemplo ilustra que en las muestras grandes, estas dos versiones de pruebas de chi cuadrado son básicamente equivalentes y conducen a las mismas conclusiones sobre la significación.

Preguntas de estudio (P12.11)

	Droga		
	Nueva	Convencional	Total
Éxito	4	2	6
Fracaso	1	6	7
Total	5	8	13

Para los datos anteriores, mostramos previamente que el valor p de una cola de la prueba exacta de Fisher es 0,0863. El estadístico MH calculado a partir de estos datos resulta ser 3,46.

1. El valor p para una prueba de MH de dos colas es 0,0630. ¿Cuál es el valor p para una prueba de Mantel-Haenszel de una cola?
2. ¿Por qué el valor p de la prueba de Mantel-Haenszel es diferente del valor p de la prueba exacta de Fisher?
3. El valor calculado para el estadístico chi cuadrado de muestra grande para comparar dos proporciones es 3,75. ¿Por qué este último valor es distinto del estadístico de la prueba de Mantel-Haenszel calculado de 3,46?

Resumen

- Una aproximación de muestra grande a la prueba exacta de Fisher es el estadístico chi cuadrado de Mantel-Haenszel (MH) para el análisis simple.
- La fórmula del estadístico MH es :

$$\chi^2 = \frac{(n-1)(ad-bc)^2}{n_1 n_0 m_1 m_0}$$

- El estadístico chi cuadrado de MH contiene *n*-1 en el numerador, mientras que la versión chi cuadrado de muestra grande para comparar dos proporciones contiene *n* en el numerador.
- En las muestras grandes, cualquiera de estas dos versiones de prueba de chi cuadrado es básicamente equivalente y conducirá a las mismas conclusiones sobre la significación.

Intervalo de confianza de muestra grande para una razón de riesgos

Nuevamente utilizamos los datos del estudio de mortalidad sobre el comportamiento con respecto al tabaquismo de pacientes que sufrieron un ataque cardíaco, pero esta vez describiremos cómo obtener un intervalo de confianza de muestra grande para una razón de riesgos.

Pacientes ataque cardíaco	Fuma	No Fuma	**Total**
Mortalidad	27	14	41
Supervivencia	48	67	115
Total	75	81	156

La razón de riesgos es una razón de dos proporciones, cada una de las cuales es una medida de incidencia acumulada. Si estuviésemos interesados en la **diferencia**, en lugar de la **razón**, entre dos incidencias acumuladas, el intervalo de confianza de muestra grande estaría dado por la siguiente fórmula de intervalo de confianza empleada comúnmente para dos proporciones:

Intervalo de confianza de muestra grande para $p_1 - p_2$

$$(\hat{p}_1 - \hat{p}_2) \pm 1.96 \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

Esta fórmula indica que, a la diferencia entre las dos proporciones calculadas, debemos sumar y restar 1,96 y multiplicarlo por el error estándar de esta diferencia. La fórmula correspondiente para el intervalo de confianza de 95% para una razón de riesgos es apenas más complicada. A diferencia de la fórmula de diferencia de riesgos, la fórmula de razón de riesgos tiene la siguiente forma:

Intervalo de confianza de muestra grande para RR $\left(= \frac{p_1}{p_2}\right)$?

$$\hat{RR} \exp\left[\pm 1.96 \sqrt{\frac{1 - \hat{CI}_1}{n_1 \hat{CI}_1} + \frac{1 - \hat{CI}_0}{n_0 \hat{CI}_0}}\right]$$

Nota: $\exp[\,a\,] = e^a$

A partir de esta fórmula, los límites inferior y superior son las siguientes expresiones:

$$L = \hat{RR} \exp\left[- 1.96 \sqrt{\frac{1 - \hat{IA}_1}{n_1 \hat{IA}_1} + \frac{1 - \hat{IA}_0}{n_0 \hat{IA}_0}}\right]$$

$$U = \hat{RR} \exp\left[+ 1.96 \sqrt{\frac{1 - \hat{IA}_1}{n_1 \hat{IA}_1} + \frac{1 - \hat{IA}_0}{n_0 \hat{IA}_0}}\right]$$

A continuación, se muestra una versión equivalente de la fórmula de la razón de riesgos, que ayuda a explicar de dónde proviene esta fórmula:

$$\frac{1 - \hat{IA}_1}{n_1 \hat{IA}_1} + \frac{1 - \hat{IA}_0}{n_0 \hat{IA}_0} \approx \hat{Var}\{\ln(\hat{RR})\}$$

El intervalo de confianza para una razón de riesgos se obtiene exponenciando un intervalo de confianza de muestra grande para el logaritmo natural de la razón de riesgos. En realidad, la fórmula utilizada para la varianza del logaritmo de la razón de riesgos es una fórmula aproximada, no una fórmula exacta.

$$\frac{1 - \hat{CI}_1}{n_1 \hat{CI}_1} + \frac{1 - \hat{CI}_0}{n_0 \hat{CI}_0} \approx \hat{Var}\{\ln(\hat{RR})\}$$

Hay dos razones que explican por qué la fórmula para la razón de riesgos es más complicada que para la diferencia de riesgos. En primer lugar, la diferencia calculada entre las dos proporciones tiene aproximadamente una distribución normal, pero la razón calculada de las dos proporciones tiene una distribución muy asimétrica. Por el contrario, el logaritmo de la razón de riesgos se acerca más a una distribución normal.

En segundo lugar, la varianza de la razón de dos proporciones es complicada desde el punto de vista matemático y no es igual a la razón de las varianzas de cada proporción. Sin embargo, como el logaritmo de una razón es la diferencia entre los logaritmos, es mucho más sencillo aproximar la varianza de una diferencia.

Ahora aplicamos la fórmula de la razón de riesgos al conjunto de datos del estudio de mortalidad. Al

sustituir los valores de la razón de riesgos, las incidencias acumuladas en cada grupo y los tamaños de muestras, obtenemos los límites de confianza inferior y superior que se muestran a continuación.

Pacientes ataque cardiaco	Fuma	No Fuma	Total
Mortalidad	27	14	41
Supervivencia	48	67	115
Total	75	81	156

$$2.08\exp\left[\pm 1.96\sqrt{\frac{1-27/75}{75*27/75}+\frac{1-14/81}{81*14/81}}\right]$$

$$1.185<RR<3.661$$

Preguntas de estudio (P12.12)

1. Interprete los resultados anteriores.

Resumen

- Desde el punto de vista matemático, la fórmula del intervalo de confianza para la razón de dos proporciones es más complicada que la fórmula para la diferencia entre dos proporciones.
- La fórmula de la razón de riesgos de 95% multiplica la razón de riesgos calculada por el exponencial de ±1,96 multiplicado por la raíz cuadrada de la varianza del logaritmo de la razón de riesgos.
- Esta fórmula de la razón de riesgos se obtiene exponenciando un intervalo de confianza de muestra grande para el logaritmo natural de la razón de riesgos.
- La razón de dos proporciones tiene una distribución muy asimétrica, mientras que el logaritmo de la razón de riesgos se acerca más a una distribución normal.
- Desde el punto de vista matemático, es más sencillo derivar la varianza del logaritmo de una razón de riesgos que la varianza de la razón de riesgos en sí misma, aunque aún se requiere una aproximación.

Fórmula de aproximación de muestra grande para un intervalo de confianza para la razón de riesgos

Para obtener un intervalo de confianza de 95% para una razón de riesgos (**RR**), podemos estar inclinados inicialmente a utilizar el intervalo:

$$\hat{\mathrm{R}}\mathrm{R} \pm 1{,}96\sqrt{\hat{\mathrm{V}}\mathrm{ar}(\hat{\mathrm{R}}\mathrm{R})}$$

Sin embargo, no se recomienda este intervalo porque hay asimetría en la distribución de la RR calculada, es decir, $\hat{RR}$ no tiene una distribución normal. El enfoque propuesto involucra trabajar con el cálculo de **lnRR**, en lugar de la **RR** en sí misma, ya que la transformación del logaritmo natural tiende a tornar una distribución asimétrica en una distribución aproximadamente normal. Se utiliza un procedimiento de dos pasos: en primer lugar, se obtiene un intervalo de confianza de 95% de muestra grande para **lnRR** de la forma:

$$l\mathrm{n}\hat{\mathrm{R}}\mathrm{R} \pm 1{,}96\sqrt{\hat{\mathrm{V}}\mathrm{ar}(\ln\hat{\mathrm{R}}\mathrm{R})}$$

Entonces, ya que **RR = exp(lnRR)**, los valores superior e inferior deseados del intervalo de confianza de 95% para **RR** se encuentran tomando los antilogaritmos de los valores superior e inferior del intervalo de confianza para **RR = exp(lnRR)**. La fórmula resultante adopta la forma general:

$$\hat{\mathrm{R}}\mathrm{R}\exp\left[\pm 1{,}96\sqrt{\hat{\mathrm{V}}\mathrm{ar}(\ln\hat{\mathrm{R}}\mathrm{R})}\right]$$

Sin embargo, para completar el cálculo debemos encontrar una expresión para $\hat{Var}(\ln \hat{RR})$. Aunque no haya una expresión matemática exacta inmediatamente disponible para este cálculo de la varianza, una buena aproximación a ella puede obtenerse mediante lo que se denomina **aproximación de la serie de Taylor** (para más detalles, véase Kleinbaum, Kupper y Morgenstern, *Epidemiologic Research*, p. 298-299, John Wiley and Sons, 1982). Mediante lo que se conoce como **aproximación de la serie de Taylor de primer orden**, la fórmula aproximada para esta varianza es:

$$\hat{\mathrm{V}}\mathrm{ar}(\ln\hat{\mathrm{R}}\mathrm{R}) \approx \frac{(1-\hat{\mathrm{I}}\mathrm{A}_1)}{\mathrm{n}_1\hat{\mathrm{I}}\mathrm{A}_1} + \frac{(1-\hat{\mathrm{I}}\mathrm{A}_0)}{\mathrm{n}_0\hat{\mathrm{I}}\mathrm{A}_0}$$

donde $\hat{IA}_1$ e $\hat{IA}_0$ denotan la incidencia acumulada para las personas expuestas y no expuestas, respectivamente. Por lo tanto, la fórmula general de aproximación para el intervalo de confianza de 95% de muestra grande puede escribirse como:

$$\hat{\mathrm{R}}\mathrm{R}\exp\left[\pm 1{,}96\sqrt{\frac{(1-\hat{\mathrm{I}}\mathrm{A}_1)}{\mathrm{n}_1\hat{\mathrm{I}}\mathrm{A}_1} + \frac{(1-\hat{\mathrm{I}}\mathrm{A}_0)}{\mathrm{n}_0\hat{\mathrm{I}}\mathrm{A}_0}}\right]$$

Auto-evaluación (12.13)

Para los datos del estudio de cohorte que se muestran a continuación, la probabilidad de muerte calculada para

quienes continúan fumando (CF) es 0,303 y para los que dejaron de fumar (DF) es 0,256.

Pacientes Ataque cardiaco	Fuma	No Fuma	Total
Mortalidad	277	243	520
Supervivencia	638	705	1343
Total	915	948	1863

$$T = \frac{\hat{P}_{CF} - \hat{P}_{DF}}{\sqrt{\hat{P}(1-\hat{P})\left(\frac{1}{n_{CF}} + \frac{1}{n_{DF}}\right)}}$$

1. ¿Cuál es el valor de la probabilidad de muerte en toda la muestra? **???**.

2. Sobre la base de los valores de probabilidad que se obtuvieron, utilice una calculadora o una computadora para cuantificar el valor del estadístico de prueba T para probar la hipótesis nula de que RR es igual a 1. Su respuesta es T* = **???**

Opciones

0,279 0,475 0,491 0,533 0,721 2,20 2,23 2,27 2,31 2,36

La herramienta de densidad normal que se describe en una actividad al final de la sección 12-2 del libro de estudio puede utilizarse para determinar los valores p.

3. Utilice la herramienta de densidad normal para encontrar el valor p basado en el estadístico T calculado (2,23) y suponiendo una hipótesis alternativa de una cola superior. Su respuesta es: valor p = **???**

4. En función del valor p, debería rechazar la hipótesis nula de que RR = 1 al nivel de significación de **???**, pero no al nivel de **???**

Opciones

0,003 0,013 0,021 0,120 1% 5%

El estadístico Z calculado para probar que RR = 1 para estos datos es 2,23.

5. Mediante el estadístico Z calculado, ¿cuál es el valor del estadístico chi cuadrado de Mantel-Haenszel (CHI^2 MH) para estos datos? CHI^2 MH = **???**
6. ¿Cuál es el valor p para una prueba de dos colas de la hipótesis nula de que RR = 1? Valor p = **???**
7. La prueba exacta de Fisher no es necesaria aquí, ya que el tamaño de la muestra para este estudio es **???**

Opciones

0,006 0,013 0,026 0,120 0,240 4,970 4,973 grande pequeño

Cálculo del tamaño de la muestra para ensayos clínicos y estudios de cohorte

Cuando se envía una propuesta de investigación de la prueba de la hipótesis etiológica a fin de procurar financiamiento, normalmente se requiere que la propuesta demuestre que la cantidad de individuos que se estudiarán es "lo suficientemente grande" como para rechazar la hipótesis nula si ésta no es verdadera. Esta cuestión se refiere al **tamaño de la muestra** recomendado para el estudio planteado. Quien presenta la propuesta normalmente elaborará una sección donde describirá los cálculos del tamaño de la muestra y las decisiones resultantes sobre los requisitos del tamaño de la muestra.

Como la mayoría de los estudios epidemiológicos implican tener en cuenta varias variables (**de control**) en el análisis, incluso si se considera una única variable de exposición dicotómica, la metodología para determinar el tamaño de la muestra puede ser muy complicada. En consecuencia, se han desarrollado muchos programas de software para incorporar tales complejidades de múltiples variables, por ejemplo, **Egret SIZ**, **PASS** y **Power and Precision** (paquete basado en la web). Probablemente, el uso de dichos programas sea la forma más rigurosa, desde el punto de vista matemático, y más precisa, desde el punto de vista del cálculo, para tomar las decisiones necesarias sobre el tamaño de la muestra.

Sin embargo, hay principios básicos de los cuales derivan todos los software sofisticados y dichos fundamentos se describen en forma práctica en el contexto de un cuadro de 2 x 2 que considera el análisis simple (es decir, crudo) de la relación exposición-enfermedad principal bajo estudio. Asimismo, a menudo es conveniente utilizar fórmulas del tamaño de muestra para el análisis simple y el enfoque sin caja negra para brindar un argumento razonable y comprensible sobre los requisitos relativos al tamaño de la muestra para un estudio dado. A continuación se brinda una descripción de dichas fórmulas.

Todas las fórmulas que corresponden a los requisitos del tamaño de la muestra de la prueba de hipótesis consideran dos tipos de error que pueden cometerse a partir de la prueba estadística de hipótesis. El **error de tipo I** se presenta si la prueba estadística rechaza (incorrectamente) una hipótesis nula verdadera y el **error de tipo II** ocurre si la prueba no rechaza (incorrectamente) una hipótesis nula falsa. La probabilidad de cometer un error de tipo I generalmente se llama **α**, el **nivel de significación de la prueba**. La probabilidad de cometer un error de tipo II generalmente se denomina **β** y a la **potencia de la prueba** se denota como **1-β**. Todas las fórmulas para el tamaño de muestra que se refieren a la prueba de hipótesis tienen el objetivo de determinar aquel tamaño de muestra para un estudio dado que logrará valores deseados (pequeños) de **α** y **β** y que detectará una desviación específica de la hipótesis nula, generalmente denotada como **Δ**. En consecuencia, el investigador debe especificar los valores de **α**, **β** y **Δ** en una fórmula adecuada para establecer el tamaño de muestra requerido.

Para los ensayos clínicos y los estudios de cohorte, la fórmula del tamaño de muestra para detectar una razón de riesgos (RR) que difiera del valor nulo de 1, por al menos la cantidad **Δ**, es decir (**Δ = RR –1**), es:

$$n = \frac{(Z_{1-\alpha/2} + Z_{1-\beta})^2 \, \bar{p}\bar{q}(r+1)}{\left[p_2(RR-1)\right]^2 r}$$

donde

$Z_{1-\alpha/2}$ = el punto porcentual 100(1 - α/2) de la distribución N(0,1).

$Z_{1-\beta}$ = el punto porcentual 100(1 – β) de la distribución N(0,1).

p_2 = riesgo esperado para individuos no expuestos.

$\bar{p} = p_2(RR + r) / (r+1)$

$\bar{q} = 1 - \bar{p}$

r = razón de individuos no expuestos sobre expuestos.

(Nota: si los tamaños de las muestras deben ser iguales en los grupos de individuos expuestos y no expuestos, entonces **r** = 1. Cuando **r** no es igual a 1, la fórmula anterior brinda el tamaño de la muestra para el grupo expuesto; para obtener el tamaño de la muestra correspondiente al grupo no expuesto, utilice **n x r**).

Para ejemplificar el cálculo de **n**, suponga **α** = 0,05, **β** = 0,20, **RR** = 2, $\mathbf{p_2}$ = 0,04 y **r** = 3. Entonces:

$$\bar{p} = (0{,}04)(2 + 3) / (3 + 1) = 0{,}05$$

y al sustituir estos valores en la fórmula de **n**, obtenemos:

$$n = \frac{(1{,}96 + 0{,}8416)^2 (0{,}05)(0{,}95)(3+1)}{[(0{,}04)(2-1)]^2 3} = 311{,}69$$

Por lo tanto, el **tamaño de muestra (n)** necesario para detectar una razón de riesgos (**RR**) de 2 con una **α** de 0,05 y una **β** de 0,20, cuando el riesgo esperado para los expuestos ($\mathbf{p_2}$) es 0,04 y la razón de individuos no expuestos sobre los expuestos (**r**) es 3, equivalente a 312 individuos expuestos y 312 x 3 = 936 individuos no expuestos. La fórmula del tamaño de muestra anterior también puede utilizarse para determinar el tamaño de muestra para calcular una razón de prevalencia (es decir, **RP**) en un estudio transversal; simplemente sustituya **RR** por **RP** en la fórmula.

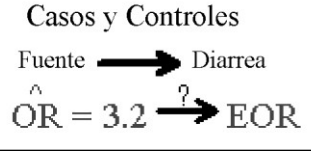
12-6 Análisis simples (continuación)

ESTUDIOS DE CASOS Y CONTROLES

Prueba de muestra grande (Z o chi cuadrado)

El cuadro que se muestra a continuación se describió en lecciones anteriores para ejemplificar datos de un estudio de casos y controles cuyo objetivo era determinar la fuente de un brote de diarrea en un club de vacaciones de Haití. El valor del *odds ratio* aquí es más grande que el valor nulo de 1 y, por lo tanto, indica que existe un efecto de magnitud moderada de la exposición. ¿Qué conclusiones podemos sacar sobre el ***odds ratio* de exposición de la población** basado en el ***odds ratio* calculado** y obtenido de la muestra?

Estudio del club	Hamburguesa Cruda: Comió	No comió	Total
Casos	17	20	37
Controles	7	26	33
Total	24	46	70



Podemos responder esta pregunta realizando una prueba de hipótesis que indique que el *odds ratio* de exposición de la población es igual a 1. La hipótesis alternativa lógica aquí es que el *odds ratio* es > 1, ya que, antes de observar los datos, los investigadores estaban interesados en saber si era más probable que comer hamburguesa cruda provocara diarrea en una persona que no comerla. La hipótesis nula que indica que el *odds*

ratio es igual a 1 (H_0: EOR = 1) puede expresarse en forma equivalente diciendo que la diferencia entre dos proporciones es igual a cero. Las proporciones, en este caso, son la proporción expuesta entre los casos y la proporción expuesta entre los controles. Estas **proporciones de exposición no son riesgos**, ya que se trata de un estudio de casos y controles. Sin embargo, podemos aún utilizar un **estadístico Z de muestra grande** para probar la diferencia entre estas dos proporciones.

$H_0 : p_1 - p_0 = 0$ $\quad H_0 : EOR = 1 \quad H_A : EOR > 1$

$\hat{p}_1 = \hat{p}($ expuesta entre casos $) = 17/37 = 0.46$

$\hat{p}_0 = \hat{p}($ expuesta entre controles $) = 7/33 = 0.21$

A continuación se incluye el estadístico de prueba, que da la diferencia entre las dos proporciones de exposición calculadas, dividida por el error estándar de esta diferencia según la hipótesis nula. Si los tamaños de las muestras en ambos grupos de casos y controles son razonablemente grandes, como vamos a suponer aquí, este estadístico de prueba tiene aproximadamente una distribución normal estándar según la hipótesis nula.

$$T = \frac{\hat{p}_1 - \hat{p}_0}{S_{\hat{p}_1 - \hat{p}_0}}$$

error estándar según H_0

Si n_{casos} , $n_{controles}$ son razonablemente grandes

aprox. **N(0,1)** según H_0

Debido a que el estadístico de muestra grande para probar la diferencia entre dos proporciones tiene una distribución normal, el cuadrado del estadístico de prueba tiene aproximadamente una distribución chi cuadrado con 1 grado de libertad. (Nota: aquí describiremos solamente la versión chi cuadrado de esta prueba. Sin embargo, es posible que desee ver el recuadro a continuación para ver el valor calculado del estadístico T y su correspondiente valor p basado en la distribución normal).

$$T^2 = \frac{\left(\hat{p}_1 - \hat{p}_0\right)^2}{\left(S_{\hat{p}_1 - \hat{p}_0}\right)^2}$$

Si n_{casos} , $n_{controles}$ son razonablemente grandes

aprox. **N(0,1)** según H_0

La disposición general de los datos correspondientes a esta prueba de chi cuadrado para los datos de casos y controles es básicamente la misma que la que se describió anteriormente para los datos de cohorte. La única diferencia es que para los datos de casos y controles, los márgenes de las filas están fijos antes del comienzo del estudio, mientras que para los datos de cohorte, los márgenes de las columnas se fijan con antelación.

Casos y Controles

	E	no E	Total
D	a	b	m_1
no D	c	d	m_0
Total	n_1	n_0	n

Cohorte

	E	no E	Total
D	a	b	m_1
no D	c	d	m_0
Total	n_1	n_0	n

En consecuencia, para los datos de casos y controles se puede usar exactamente la misma fórmula de chi cuadrado descrita anteriormente para los datos de cohorte.

$$X^2 = \frac{\left(\hat{p}_1 - \hat{p}_0\right)^2}{\left(S_{\hat{p}_1 - \hat{p}_0}\right)^2} = \frac{n(ad - bc)^2}{(n_1 n_0 m_1 m_0)}$$

Al sustituir las frecuencias de celda de los datos de nuestro estudio de casos y controles en la fórmula del estadístico chi cuadrado, obtenemos el valor 4,74. Si, en su lugar, se utiliza la fórmula chi cuadrado de Mantel-Haenszel, de forma que el numerador se reemplaza por n – 1, el estadístico chi cuadrado cambia sólo levemente a 4,67.

$$X^2 = \frac{70(17*26 - 20*7)^2}{(24*46*37*33)} = 4.74$$

$$X^2_{MH} = \frac{(70-1)(17*26 - 20*7)^2}{(24*46*37*33)} = 4.67$$

Preguntas de estudio (P12.14)

1. Utilizando el valor chi cuadrado de Mantel Haenszel de 4,67, describa cómo determinaría el valor p para la hipótesis alternativa de una cola que indica que el *odds ratio* de exposición es mayor que 1. (Esta pregunta pide un proceso, no un valor numérico).
2. Los puntos porcentuales 95 y 97,5 de la distribución chi cuadrado con 1 grado de libertad son 3,84 y 5,02, respectivamente. Sobre la base de esta información, ¿qué puede decir del valor p de la prueba de Mantel-Haenszel? (Se le pide un número o rango de números para el valor p).
3. ¿El *odds ratio* de exposición calculado a partir del estudio de casos y controles es estadísticamente significativo al nivel de significación de 0,05? ¿Y al nivel de significación de 0,01?

Resumen

- Cuando se prueba la hipótesis sobre un *odds ratio* (OR) en un estudio de casos y controles, la hipótesis nula puede expresarse en forma equivalente como EOR = 1 **o** $p_1 - p_2 = 0$, donde p_1 y p_2 son probabilidades de exposición calculadas para los casos y los no casos.

- La hipótesis alternativa puede expresarse en términos de EOR, ya sea como EOR ≠ 1, EOR > 1 o EOR < 1, si la hipótesis alternativa es de dos colas, de una cola superior o de una cola inferior, respectivamente.
- Una versión del estadístico de prueba es un estadístico N(0,1) de muestra grande, utilizado para comparar dos proporciones.
- Una versión alternativa es un estadístico chi cuadrado de muestra grande, que es el cuadrado de un estadístico N(0,1).
- En forma alternativa, de la prueba de chi cuadrado puede emplearse un estadístico chi cuadrado de Mantel-Haenszel de muestra grande.

El estadístico Z para una prueba de muestra grande de un *odds ratio*

El estadístico Z de muestra grande para probar la diferencia entre dos probabilidades de exposición de un estudio de casos y controles es la fórmula:

$$T = \frac{(\hat{p}_{CA} - \hat{p}_{CO})}{\sqrt{\hat{p}(1-\hat{p})(\frac{1}{n_{CA}} + \frac{1}{n_{CO}})}}$$

donde $\hat{p}_{CA} = \hat{Pr}(E \mid D)$, $\hat{p}_{CO} = \hat{Pr}(E \mid \text{no } D)$ y $\hat{p} = \hat{Pr}(E)$

Este estadístico muestra la diferencia entre las dos proporciones de exposición calculadas, dividida por el error estándar de esta diferencia bajo la hipótesis nula de que el *odds ratio* es igual a 1. Si los tamaños de las muestras tanto en el grupo de los casos como en el de los controles son razonablemente grandes, como vamos a suponer aquí, este estadístico de prueba tiene aproximadamente una distribución normal estándar bajo la hipótesis nula.

Utilizando los datos del brote de diarrea en un club de vacaciones de Haití, el valor calculado del estadístico de prueba, que llamamos **T***, se obtiene sustituyendo las incidencias acumuladas calculadas y los tamaños de muestra correspondientes en la fórmula del estadístico de prueba. Estos valores son:

$$\hat{p}_{CA} = \frac{17}{37} = 0{,}459,\ \hat{p}_{CO} = \frac{7}{33} \text{ y } \hat{p} = \frac{24}{70} = 0{,}343,$$

donde $n_{CA} = 37$ y $n_{CO} = 33$

El valor resultante del estadístico **Z** es **T* = 2,18**. El **valor p** para esta prueba se obtiene, entonces, encontrando el área en la cola derecha de la distribución normal estándar por arriba del valor calculado de **2,18**. Esta área es **0,0146**. Este valor p está por debajo del nivel de significación convencional de 0,05, pero por arriba de 0,01 y, por lo tanto, rechazaríamos la hipótesis nula al nivel de 5%, pero no al nivel de 1%. En otras palabras, hemos descubierto que el *odds ratio* para la asociación entre la ingesta de hamburguesa cruda y el desarrollo de diarrea tiene una significación estadística límite.

Realización de pruebas cuando los tamaños de las muestras son pequeños

Este cuadro contiene resultados hipotéticos de casos y controles del estudio sobre el brote de diarrea en el club de vacaciones de Haití basados en una muestra de 19 personas. Aquí, el tamaño total de la muestra es menor a un tercio del tamaño de la muestra encontrado anteriormente para los datos reales del brote que se describieron en la actividad anterior. El *odds ratio* calculado es 3,5.

Resultados hipotéticos sobre el brote de diarrea

Estudio del club	Hamburguesa Cruda: Comió	No comió	Total
Casos	5	5	10
Controles	2	7	9
Total	7	12	19

$\hat{OR} = 3.5$

¿Este *odds ratio* es significativamente distinto del valor nulo de 1? Aquí estamos considerando un tamaño pequeño de muestra, por lo que no es adecuada una prueba de hipótesis de muestra grande. Sin embargo, podemos utilizar la prueba exacta de Fisher para estos datos. Nuevamente, consideramos la disposición general de los datos para un cuadro de 2 x 2 con un resultado de exposición dicotómico y un resultado de enfermedad dicotómico. Para utilizar la prueba exacta de Fisher, debemos suponer que los valores en los márgenes del cuadro están fijos antes del comienzo del estudio.

Disposición general	**E**	**no E**	**Total**
D	a	b	m_1
no D	c	d	m_0
Total	n_1	n_0	n

Suponiendo márgenes fijos

Suponiendo "márgenes fijos", sólo necesitamos considerar el resultado para una de las cuatro celdas del cuadro, digamos la celda **a**. Por lo tanto, nos gustaría saber qué valores podríamos obtener en la celda **a** que sería improbable que sucedieran bajo la hipótesis nula. En especial, nos gustaría determinar si el valor obtenido en la celda **a** para nuestro estudio de casos y controles, que resulta ser 5, es un evento lo suficientemente poco frecuente según la hipótesis nula como para que la rechacemos y lleguemos a la conclusión de que el *odds ratio* observado es estadísticamente significativo. Podemos responder esta pregunta calculando el valor p para nuestra prueba.

Preguntas de estudio (P12.15)

Estudio del club	Hamburguesa Cruda: Comió	No comió	Total
Casos	4	b	10
Controles	c	d	9
Total	7	12	19

Suponga que el valor de la celda **a** fuera 4, en lugar de 5, asumiendo que los totales marginales son fijos.

1. ¿Cuáles serían los valores corregidos de **b**, **c** y **d** correspondientes?
2. ¿Cuál sería el *odds ratio* para el cuadro corregido?
3. ¿Cuál *odds ratio* se aleja más del *odds ratio* nulo? ¿El *odds ratio* corregido o el *odd ratio* de 3,5 que en verdad se observó?
4. ¿Existen otros valores posibles para la celda **a** que resultarían en un *odds ratio* que se aleje más del *odds ratio* nulo que el *odds ratio* de 3,5?
5. Sobre la base de las preguntas anteriores, ¿por qué quisiéramos calcular la probabilidad de obtener un valor de la celda **a** de 5, 6 ó 7 según la hipótesis nula?

Para calcular el valor p de la prueba exacta de Fisher, debemos determinar la distribución de probabilidad de la frecuencia de la celda **a** según la hipótesis nula. Suponiendo totales marginales fijos, esta distribución se denomina **distribución hipergeométrica**. La fórmula para esta distribución y el valor p correspondiente de la prueba exacta de Fisher se brindan en el suplemento correspondiente a esta actividad en el CD-ROM y en el recuadro que se encuentra después de la actividad en este libro de texto.

Preguntas de estudio (P12.15) (continuación)

Utilizando la distribución hipergeométrica, este valor p es:

$$\Pr(a = 5 \text{ ó } 6 \text{ ó } 7 | OR = 1) = \Pr(a = 5|OR = 1) + \Pr(a = 6|OR = 1) + \Pr(a = 7|OR = 1)$$
$$= 0{,}1800 + 0{,}0375 + 0{,}0024 = 0{,}2199$$

6. ¿A qué conclusión llega sobre la hipótesis nula?

Resumen

- La prueba exacta de Fisher brinda una prueba de significación para un *odds ratio* y una razón de riesgos cuando los datos son escasos.
- Para utilizar la prueba exacta de Fisher, debemos suponer que los valores en los márgenes de un cuadro de 2 x 2 son valores fijos antes del comienzo del estudio.
- Para calcular el valor p de la prueba exacta de Fisher, debemos determinar la distribución de probabilidad de la frecuencia de la celda **a** bajo la hipótesis nula.
- Suponiendo totales marginales fijos, la celda **a** tiene distribución hipergeométrica.

Para calcular el valor p para esta prueba se encuentran disponibles programas informáticos.

Cómo realizar la prueba exacta de Fisher mediante la distribución hipergeométrica en un estudio de casos y controles

El procedimiento de la **prueba exacta de Fisher** es el mismo, independientemente de si el diseño del estudio es de casos y controles, de cohorte o transversal, siempre y cuando se supongan **totales marginales fijos**. Ahora describiremos la derivación de este procedimiento (como hicimos previamente en un suplemento/recuadro anterior cuando analizamos el tema de muestras pequeñas en los estudios de cohorte o en los ensayos clínicos). Luego, se brinda un ejemplo de la prueba exacta de Fisher con datos de casos y controles. Para calcular el **valor p** de la prueba exacta de Fisher debemos utilizar la distribución de probabilidad de la frecuencia de la celda "a" (es decir, la cantidad de casos expuestos) según la hipótesis nula que indica que la exposición no tiene efecto sobre la enfermedad.

Disposición de los datos con totales marginales fijos y con celda "a" = "j"

	E	**no E**	
D	j	$m_1 - j$	m_1
no D	$n_1 - j$	$n_0 - m_1 + j$	m_0
	n_1	n_0	

Suponiendo totales marginales fijos, esta distribución se denomina **distribución hipergeométrica**, la cual es la siguiente fórmula de probabilidad para una variable aleatoria A:

$$pr(A = j \mid H_0) = \frac{C_j^{n_1} C_{m_1 - j}^{n_0}}{C_{m_1}^{n_1 + n_0}} \qquad \text{donde} \quad C_s^r = \frac{r!}{s!(r-s)!}$$

denota la cantidad de combinaciones de *r* elementos tomados *s* por vez, y

r! = *r*(*r*-l)(*r*-2)...(3)(2)(1) denota el factorial de *r*,

s! y (*r*-*s*)! denotan el factorial de *s* y el factorial de (*r*-*s*) , respectivamente.

Surge entonces que el **valor p exacto** de una prueba de H_0:OR = 1 versus H_A: OR > 1, mediante la fórmula hipergeométrica anterior, está dado por

$$P = pr(A \geq a \mid H_0) = \sum_{j=a}^{\min(n_1, m_1)} pr(A = j \mid H_0)$$

Ejemplificamos el cálculo de este valor p mediante los datos de un estudio de casos y controles hipotético sobre un brote de diarrea en un club de vacaciones de Haití, que involucró a 19 individuos (en lugar de los 70 individuos reales del brote verdadero). Aquí están los datos nuevamente:

Estudio de casos y controles hipotético de un brote en un club de vacaciones de Haití.

	Comieron hamburguesa cruda	**No comieron**	
Casos	5	5	10
Controles	2	7	9
	7	12	19

Por lo tanto, el valor p se calcula como la probabilidad de que la celda "a" sea > 5 (pero menor que los valores marginales mínimos de 7 y 10) según la hipótesis nula de que OR = l:

$$P = pr(A \geq 5 \mid H_0) = \sum_{j=5}^{7} pr(A = j \mid H_0)$$

$$= \frac{C_5^7 C_{10-5}^{12}}{C_{10}^{19}} + \frac{C_6^7 C_{10-6}^{12}}{C_{10}^{19}} + \frac{C_7^7 C_{10-7}^{12}}{C_9^{19}} = 0{,}1800 + 0{,}0375 + 0{,}0024 = 0{,}2199$$

Ya que este valor p es mucho más grande que 0,05, llegamos a la conclusión de que la hipótesis nula no podría rechazarse; es decir, no hay evidencia suficiente a partir de estos 19 individuos de que comer hamburguesa cruda tuvo un efecto significativo sobre el hecho de que un individuo se enfermara o no.

Intervalo de confianza de muestra grande

Nuevamente utilizamos los datos de casos y controles del brote de diarrea en un club de vacaciones de Haití, pero esta vez describiremos cómo obtener un intervalo de confianza de muestra grande para un *odds ratio*. El valor del *odds ratio* que describe la asociación entre comer hamburguesa cruda y desarrollar diarrea fue 3,2.

Estudio del club	Hamburguesa Cruda: Comió	No comió	Total
Casos	17	20	37
Controles	7	26	33
Total	24	46	70

$$\hat{OR} = 3.2$$

A continuación se presenta la fórmula general correspondiente al intervalo de confianza de 95% para cualquier *odds ratio* basado en la disposición general de 2 x 2 de los datos de un estudio de casos y controles. También se brindan los límites inferior y superior.

$$\hat{OR}e^{\pm 1.96\sqrt{\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}}}$$

Limite Inferior

$$\hat{OR}e^{-1.96\sqrt{\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}}}$$

Limite Superior

$$\hat{OR}e^{+1.96\sqrt{\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}}}$$

Disposición general	E	no E	Total
D	a	b	m_1
no D	c	d	m_0
Total	n_1	n_0	n

A continuación se muestra una versión equivalente de la fórmula del intervalo de confianza de un *odds ratio*, que ayuda a explicar de dónde proviene esta fórmula:

$$e^{\ln\hat{OR} \pm 1.96\sqrt{\hat{Var}(\ln\hat{OR})}}$$

Esta expresión dice que el intervalo de confianza para un *odds ratio* se obtiene exponenciando un intervalo de confianza de muestra grande para el logaritmo natural del *odds ratio*. En realidad, la fórmula utilizada para la estimación de la varianza del logaritmo del *odds ratio* es una fórmula aproximada, no una fórmula exacta.

Ahora aplicamos la fórmula del *odds ratio* al conjunto de datos de casos y controles del brote. Al sustituir en la fórmula los valores para el *odds ratio* calculado y las cuatro frecuencias de celda, obtenemos los límites de confianza inferior y superior que se indican a continuación.

$$3.16\, e^{\pm 1.96\sqrt{\frac{1}{17}+\frac{1}{20}+\frac{1}{7}+\frac{1}{26}}}$$

$$1.099 < OR < 9.074$$

Preguntas de estudio (P12.16)

1. ¿Qué interpretación puede hacer de este intervalo de confianza?

Resumen

- La fórmula para un intervalo de confianza de 95% para una fórmula del *odds ratio* multiplica el *odds ratio* calculado por el exponencial de ±1,96 multiplicado por la raíz cuadrada de la varianza del logaritmo del *odds ratio*.
- La fórmula del *odds ratio* se obtiene exponenciando un intervalo de confianza de muestra grande para el logaritmo natural del *odds ratio*.
- La varianza del logaritmo de un *odds ratio* es aproximadamente igual a la suma de los inversos de las cuatro frecuencias de celda dispuestas en el cuadro de 2 x 2.

Derivación de la fórmula de aproximación de muestra grande para un intervalo de confianza para el *odds ratio*

Al igual que con la razón de riesgos (RR), para obtener un intervalo de confianza de 95% para un *odds ratio* (**OR**), podemos estar inclinados inicialmente a utilizar el intervalo:

$$\hat{OR} \pm 1{,}96\sqrt{\hat{Var}(\hat{OR})}$$

Sin embargo, este intervalo **no** se recomienda porque hay asimetría en la distribución del OR calculado, es

decir, $\hat{OR}$ no tiene una distribución normal. El enfoque propuesto involucra trabajar con el cálculo de **lnOR**, en lugar del **OR** en sí mismo, ya que la transformación del logaritmo natural tiende a tornar una distribución asimétrica en una distribución aproximadamente normal. Se utiliza un procedimiento de dos pasos: en primer lugar, se obtiene un intervalo de confianza de 95% de muestra grande para el **lnOR** en la forma de:

$$\ln\hat{OR} \pm 1{,}96\sqrt{\hat{Var}(\ln\hat{OR})}$$

Entonces, ya que **OR = exp(lnOR)**, los valores superior e inferior deseados del intervalo de confianza de 95% para **OR** se encuentran tomando los antilogaritmos de los valores inferior y superior para el intervalo de confianza para **lnOR**. La fórmula resultante adopta la forma general:

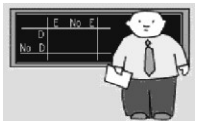
$$\hat{OR}\exp\left[\pm 1{,}96\sqrt{\hat{Var}(\ln\hat{OR})}\right]$$

Sin embargo, para completar el cálculo debemos encontrar una expresión para $\hat{Var}(\ln\hat{OR})$. Aunque no haya una expresión matemática inmediatamente disponible para este cálculo de la varianza, una buena aproximación a ella puede obtenerse mediante lo que se denomina **aproximación de la serie de Taylor** (para más detalles, véase Kleinbaum, Kupper y Morgenstern, *Epidemiologic Research*, p. 298-299, John Wiley and Sons, 1982). Mediante lo que se conoce como **aproximación de la serie de Taylor de primer orden**, la fórmula aproximada para esta varianza es:

$$\hat{Var}(\ln\hat{OR}) \approx \frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}$$

donde **a**, **b**, **c**, **d** denotan las frecuencias de celda en la disposición de datos de 2 x 2 para un estudio de casos y controles. Por lo tanto, la fórmula general de aproximación para el intervalo de confianza de 95% de muestra grande puede escribirse como:

$$\hat{OR}\exp\left[\pm 1{,}96\sqrt{\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}}\right]$$



Auto-evaluación (12.17)

1. En la disposición general de los datos para una prueba de chi cuadrado de datos de casos y controles, los **???** están fijos antes del comienzo del estudio.

2. La disposición general de los datos para una prueba de chi cuadrado de datos de cohorte tiene **???** fijos.
3. La fórmula de chi cuadrado de muestra grande para datos de casos y controles es la **???** la fórmula de chi cuadrado de muestra grande para datos de cohorte de incidencia acumulada.
4. La fórmula de cálculo para la prueba exacta de Fisher para datos de casos y controles es **???** la fórmula correspondiente para datos de cohorte de incidencia acumulada.

Opciones
frecuencias de celda **distribuciones chi cuadrado** **márgenes de columnas** **distinta de** **distribuciones normales** **márgenes de filas** **misma que**

Complete los espacios vacíos en el cuadro.

Estudio del club	Hamburguesa Cruda: Comió	No comió	Total
Casos	6	B	11
Controles	C	D	11
Total	10	12	22

5. B = **???**
6. C = **???**
7. D = **???**

Opciones
3 **4** **5** **6** **7** **8** **9**

8. La siguiente expresión indica que el **???** para un *odds ratio* se obtiene **???** un **???** de muestra grande para el logaritmo natural del **???**

$$e^{\ln\hat{OR}\pm 1.96\sqrt{\hat{Var}(\ln\hat{OR})}}$$

donde $\hat{Var}(\ln\hat{OR}) \approx 1/a + 1/b + 1/c + 1/d$

Opciones
intervalo de confianza **exponenciando** **prueba de hipótesis** **logaritmo natural** ***odds ratio*** **razón de riesgos**

Cálculo del tamaño de muestra para el *odds ratio* en los estudios de casos y controles y otros estudios

En un suplemento/recuadro en la sección anterior (12-5) describimos una fórmula para determinar el tamaño de muestra (**n**) a fin de calcular una **razón de riesgos** (**RR**) en ensayos clínicos y estudios de cohorte. Esta fórmula también se aplica para calcular tamaños de muestra para la **razón de prevalencia (RP)** en estudios transversales (es decir, se debe reemplazar la **RR** por la **RP** en la fórmula). Ahora describimos y ejemplificamos una variación de la fórmula anterior que podemos emplear para determinar el tamaño de la muestra para el cálculo de un ***odds ratio*** (**OR**) en un estudio de casos y controles, de cohorte, transversal o un ensayo clínico no equiparados.

Al igual que lo que sucede con la fórmula del tamaño de muestra para la **RR**, cuando se considera en su lugar el **OR**, el investigador debe especificar valores para el nivel de significación **α**, la probabilidad de un error de tipo II **β**, como también el grado de alejamiento de efecto nulo que se observa en el efecto del estudio, es decir, **Δ**, en una fórmula adecuada para determinar el tamaño requerido de la muestra. Para estos estudios, la fórmula del tamaño de la muestra para detectar un *odds ratio* (**OR**) que difiera del valor nulo de 1, por al menos **Δ**, es decir, (**Δ = OR - 1**) es:

$$n = \frac{(Z_{1-\alpha/2} + Z_{1-\beta})^2 \, \bar{p}\bar{q}(r+1)}{[p_1 - p_2]^2 \, r}$$

donde

$Z_{1-\alpha/2}$ = el punto porcentual 100(1 - α/2) de la distribución N(0,1).

$Z_{1-\beta}$ = el punto porcentual 100(1 – β) de la distribución N(0,1).

p_1 = proporción esperada de casos expuestos

p_2 = proporción esperada de controles expuestos

$$OR = \left(\frac{p_1}{1-p_1}\right) \Big/ \left(\frac{p_2}{1-p_2}\right)$$

$\bar{p} = (p_1 + rp_2) / (r+1)$

$\bar{q} = 1 - \bar{p}$

r = razón de la cantidad de controles sobre casos

(Nota: si los tamaños de las muestras deben ser iguales en los grupos de individuos expuestos y no expuestos, entonces **r = 1**. Cuando **r** no es igual a 1, la fórmula anterior brinda el tamaño de la muestra para el grupo de **casos**; para obtener el tamaño de la muestra correspondiente al grupo de controles, utilice **n x r**).

Para usar la fórmula anterior, normalmente se especifica (un valor estimado para) **p2** y el **OR** que se deben detectar; luego se resuelve la fórmula de **OR** para p1 en términos de OR y p2:

Para ejemplificar el cálculo de **n**, suponga **α** = 0,05, **β** = 0,20, **OR** = 2, $\mathbf{p_2}$ = 0,040 y **r** = 3. Entonces:

$$OR = 2 = \left(\frac{p_1}{1-p_1}\right) \Big/ \left(\frac{.04}{1-.04}\right)$$

Y reemplazando p_1 obtenemos:

$$p_1 = \frac{p_2 OR}{1 - p_2 + p_2 OR} = \frac{(0{,}04)(2)}{1 - 0{,}04 + (0{,}04)(2)} = 0{,}0769$$

$$\bar{p} = [0{,}0769 + 3(0{,}04)] / (3+1) = 0{,}0492$$

Y finalmente resolviendo **n** obtenemos:

$$n = \frac{(1{,}96 + 0{,}8416)^2 (0{,}0492)(0{,}9508)(3+1)}{[0{,}0769 - 0{,}04]^2 3} = 359{,}54$$

Por lo tanto, el **tamaño de muestra (n)** necesario para detectar un *odds ratio* (**OR**) de 2 con una **α** de 0,05 y una **β** de 0,20, cuando la proporción esperada de expuestos entre los controles (p_2) es 0,04 y la razón de controles sobre casos (**r**) es 3, equivale a **360 casos** y 360 x 3 = **1.080 controles**.

12-7 Análisis simples (continuación)

ESTUDIOS DE COHORTE QUE INVOLUCRAN RAZONES DE TASAS

Pruebas de razones de tasas

Este cuadro se describió en lecciones anteriores para ejemplificar los datos de un **estudio de cohorte con información de persona-tiempo** cuyo objetivo era evaluar la relación entre el nivel de colesterol en sangre y la mortalidad. La medida de efecto en este estudio era una **razón de tasas**, también denominada **razón de densidades de incidencia** (**RDI**), y su valor puntual fue 3,46.

Estudio de cohorte con información de persona-tiempo

Estudio de colesterol	Nivel de colesterol: Moderadamente alto	Normal
Muertes	26	14
P-años	36,581	68,239

$\hat{RDI} = 3.46$

¿Qué conclusión podemos sacar sobre la razón de tasas de la población basada en la razón de tasas obtenida de la muestra? Podemos realizar una prueba de la hipótesis de que la razón de tasas de la población es igual a 1. Esta hipótesis nula puede expresarse en forma equivalente diciendo que la diferencia entre dos tasas de

incidencia es igual a cero.

$$H_0: \text{RDI} = 1$$
o $$H_0: TI_1 - TI_0 = 0$$

Las tasas en este caso son: la tasa de mortalidad para las personas con colesterol moderadamente alto o alto ($\hat{T}I_1$) y la tasa de mortalidad para individuos con colesterol normal ($\hat{T}I_0$).

$$\hat{TI}_1 = \text{Tasa (moderadamente alto o alto)} = \frac{26}{36581}$$
$$= 7.11 \text{ por } 10{,}000 \text{ p-años}$$
$$\hat{TI}_0 = \text{Tasa (normal)} = \frac{14}{68239}$$
$$= 2.05 \text{ por } 10{,}000 \text{ p-años}$$

La hipótesis alternativa lógica aquí es que la razón de tasas es mayor que 1, ya que antes de observar los datos, los investigadores estaban interesados en saber si los individuos con colesterol moderadamente alto tenían una tasa de mortalidad mayor que las personas con colesterol normal. Como estamos tratando con tasas y no con cálculos de riesgos, no podemos utilizar una prueba Z de muestra grande que compara dos proporciones o su prueba equivalente de chi cuadrado. Pero podemos emplear un estadístico Z de muestra grande diferente o su estadístico chi cuadrado equivalente para comparar dos tasas. A continuación se muestra una disposición general para calcular una razón de tasas.

H_0: RDI = 1 H_A: RDI > 1

Disposición 2x2 para cohorte con información de persona-tiempo

	Expuestos	No expuestos	Total
Casos nuevos	I_1	I_0	I
Persona tiempo	PT_1	PT_0	PT

$\mathbf{I_1}$ e $\mathbf{I_0}$ denotan el número de casos nuevos en los grupos expuestos y no expuestos; además, $\mathbf{PT_1}$ y $\mathbf{PT_0}$ hacen referencia al valor persona-tiempo acumulado correspondiente a estos dos grupos. Aquí está el estadístico de prueba, **T**, donde $\mathbf{p_0}$ denota el valor de persona-tiempo del grupo expuesto, dividido por el valor total de persona-tiempo de los expuestos y no expuestos combinados:

$$\mathbf{T} = \frac{I_1 - I p_0}{\sqrt{I p_0 (1 - p_0)}} \quad \textbf{approx. } \mathbf{N(0, 1)} \text{ según } \mathbf{H_0}$$
$$p_0 = \frac{PT_1}{PT}$$

Este estadístico de prueba tendrá aproximadamente una distribución normal bajo la hipótesis nula, siempre que el número total de nuevos casos, I, sea lo suficientemente grande y que p_0 no esté cerca de 1 ó 0. Nos referiremos a este estadístico como un estadístico Z. Si elevamos al cuadrado este estadístico Z, obtenemos un estadístico chi cuadrado aproximado con un grado de libertad (1 gl). A continuación se incluye una fórmula simplificada para este estadístico chi cuadrado:

$$Z^2 = \frac{(I_1 PT_0 - I_0 PT_1)^2}{I\, PT_1 PT_0}$$

Ahora calculamos Z^2 para nuestro conjunto de datos sobre el colesterol y la mortalidad. El valor chi cuadrado calculado es 15,952.

$$Z^2 = \frac{(26 \times 68{,}239 - 14 \times 36{,}581)^2}{40 \times 36{,}581 \times 68{,}239} = 15.952$$

approx. chi cuadrado con 1 d.f. según H_0

Preguntas de estudio (12.18)

1. Utilizando el valor chi cuadrado de muestra grande de 15,592, describa cómo determinaría el valor p de esta hipótesis alternativa de una cola que indica que la razón de tasas (es decir, RDI) es mayor que 1. (Se le pide que describa un proceso, no que obtenga un valor numérico).
2. El punto porcentual 99,95 de la distribución chi cuadrado con 1 gl es 12,116. Sobre la base de esta información, ¿qué puede decir sobre el valor p de la prueba de chi cuadrado de muestra grande? (Ahora se le pide un número o rango de números).
3. En función del valor p anterior, ¿a qué conclusión llega a partir de esta prueba de hipótesis?
4. ¿Cuál es el valor calculado del estadístico Z que brinda una prueba que sea equivalente a la prueba de chi cuadrado anterior?

Resumen

- Cuando se prueba la hipótesis sobre una razón de tasas en un estudio de cohorte con datos de persona-tiempo, la hipótesis nula puede expresarse en forma equivalente como RDI = 1 *o* $TI_1 - TI_0 = 0$.
- RDI denota la razón de tasas (es decir, la razón de densidades de incidencia); TI_1 y TI_0 son las tasas de incidencia para los grupos de individuos expuestos y no expuestos.
- La hipótesis alternativa puede expresarse en términos de RDI ≠ 1, RDI > 1, o RDI < 1, si la hipótesis alternativa es de dos colas, de una cola superior o de una cola inferior, respectivamente.
- Una versión del estadístico de prueba es un estadístico N(0,1) o Z de muestra grande utilizado para comparar dos tasas de incidencia.
- Una versión alternativa es un estadístico chi cuadrado de muestra grande, que es el cuadrado del estadístico Z.

La lógica detrás de la prueba Z para una razón de tasas

A continuación describiremos cómo probar **H_0: RDI = 1** o, en forma equivalente, **H_0: TI_1 - TI_0 = 0**, donde RDI denota la **razón de tasas** (es decir, la razón de densidades de incidencia) en un análisis simple de un estudio de cohorte con datos de persona-tiempo, y **TI_1** y **TI_0** son las tasas de incidencia para los grupos de individuos expuestos y no expuestos. La disposición de datos es la siguiente:

Disposición de datos para el estudio de cohorte con información de persona-tiempo	**Expuestos**	**No expuestos**	**Total**
Casos nuevos (D)	**I_1**	**I_0**	**I**
Persona-tiempo	**PT_1**	**PT_0**	**PT**

Ya que los valores de $\hat{T}I_1$ y $\hat{T}I_0$ son proporciones no binomiales, no es adecuado aquí utilizar una prueba estándar que compare dos proporciones binomiales. Sin embargo, es posible un enfoque de prueba si suponemos que cada uno de los casos **I** representan un ensayo de **Bernoulli** independiente, con el "éxito" y "fracaso" definidos dentro de las categorías de expuestos y no expuestos, respectivamente.

Bajo este supuesto, podemos asumir que la cantidad de casos expuestos, I_1, tiene una distribución binomial para I ensayos y una probabilidad de éxito **p_0** según la hipótesis nula, donde **$p_0 = PT_1/PT$** es la proporción del total de persona-tiempo asociada al grupo expuesto.

Por lo tanto, si **A** denota la variable aleatoria para la cantidad de casos del total de casos **I** que están expuestos, de la distribución binomial surge que el **valor p exacto** de una prueba de **H_0: RDI = 1** versus **H_A: RDI > 1** mediante la fórmula binomial anterior está dada por

$$\Pr(A \geq I_1 \mid H_0) = \sum_{j=I_1}^{I} C_j^I p_0^j (1-p_0)^{I-j} \quad \text{donde} \quad C_s^r = \frac{r!}{s!(r-s)!}$$

denota la cantidad de combinaciones de **r** elementos tomados **s** por vez, y

r! = r(r-l)(r-2)...(3)(2)(1) denota el **factorial de r**,

s! y (r-s)! denotan el **factorial de s** y el **factorial de (r-s)**, respectivamente.

Ilustramos el cálculo de este **valor p** mediante los datos de un estudio de persona-tiempo que compara las tasas de mortalidad de los individuos con colesterol moderadamente alto con las de las personas con colesterol normal. Aquí están los datos nuevamente:

Estudio de cohorte sobre el colesterol con información de persona-tiempo	**Expuestos**	**No expuestos**	**Total**

Muertes (D)	**26**	**14**	**40**
Personas-años	**36.581**	**68.239**	**104.820**

Por lo tanto, el valor p se calcula como la probabilidad de que el número de casos expuestos sea 26 o mayor (hasta 40 casos totales) bajo la hipótesis nula de que RDI = l:

$$Pr(A \geq 26 \mid H_0) = \sum_{j=26}^{40} C_j^{40} p_0^j (1-p_0)^{40-j}$$

Esta expresión involucra la suma de 15 términos y es difícil de calcular sin un programa informático. Sin embargo, se puede obtener una buena aproximación a este **valor p exacto** a partir de una prueba **Z** de muestra grande. Por lo tanto, se recomienda la prueba **Z**, a menos que la cantidad total de casos sea pequeña, en cuyo caso se debería utilizar la fórmula del valor p exacto que se brindó anteriormente. El estadístico Z de muestra grande está dado por la siguiente fórmula:

$$Z = \frac{I_1 - Ip_0}{\sqrt{Ip_0(1-p_0)}}$$

Para los datos sobre el colesterol y la mortalidad que se presentaron con anterioridad, el estadístico Z se calcula de la siguiente manera:

$$Z = \frac{26 - 40(36.581/104.820)}{\sqrt{40(36.581/104.820)\{1-(36.581/104.820)\}}} = 3{,}994$$

El estadístico **Z** calculado es bastante grande; en especial, el **valor p** de una prueba de $\mathbf{H_0}$**: RDI = 1** versus $\mathbf{H_A}$**: RDI > 1** es extremadamente pequeño; de hecho, es más pequeño que 0,0001. En consecuencia, la hipótesis nula debería rechazarse y se puede llegar a la conclusión de que las personas con colesterol moderadamente alto tienen una tasa de mortalidad significativamente más alta que aquellas con colesterol normal.

Intervalo de confianza (IC) de muestra grande para una razón de tasas

Para describir la fórmula general correspondiente a un intervalo de confianza de 95% para una razón de tasas, consideramos nuevamente la disposición de datos para un estudio de cohorte con información de persona-tiempo. Esta fórmula del intervalo de confianza tiene la siguiente forma:

Intervalo de confianza del 95% de muestra grande para RDI

Disposición 2x2 para cohorte con información de persona-tiempo

	Expuestos	No expuestos	Total	
Casos nuevos	I_1	I_0	I	**RDI** = $\frac{\frac{I_1}{PT_1}}{\frac{I_0}{PT_0}}$
Persona tiempo	PT_1	PT_0	PT	

$$\hat{RDI}\ e^{\pm 1.96\sqrt{\frac{1}{I_1}+\frac{1}{I_0}}}$$

Luego, las expresiones para los límites inferior (L) y superior (U) son las siguientes:

$$L=\hat{RDI}\ e^{-1.96\sqrt{\frac{1}{I_1}+\frac{1}{I_0}}}\quad U=\hat{RDI}\ e^{+1.96\sqrt{\frac{1}{I_1}+\frac{1}{I_0}}}$$

Una versión equivalente de la fórmula de la razón de tasas es la siguiente:

$$e^{\ln(\hat{RDI})\pm 1.96\sqrt{\hat{Var}\{\ln(\hat{RDI})\}}}$$

La fórmula utilizada anteriormente para la varianza calculada del logaritmo de la razón de tasas es una fórmula aproximada, no una fórmula exacta, de la varianza.

Apliquemos la fórmula de la razón de tasas al conjunto de datos del estudio de mortalidad. Al sustituir en la fórmula los valores para la razón de tasas calculada y las frecuencias de celda para los grupos de individuos expuestos y no expuestos, obtenemos los límites de confianza inferior y superior: 1,81 < RDI < 6,63

$$3.46\,e^{\pm 1.96\sqrt{\frac{1}{26}+\frac{1}{14}}}$$

$$=1.81 < RDI < 6.63$$

Preguntas de estudio (12.19)

1. Interprete los resultados anteriores. ¿Qué significan?

Resumen

- La fórmula del IC del 95% para una razón de tasas multiplica la razón de tasas calculada por el exponencial de ±1,96 multiplicado por la raíz cuadrada de la varianza del logaritmo de la razón de tasas calculada.
- Esta fórmula de la razón de tasas se obtiene exponenciando un intervalo de confianza de muestra grande para el logaritmo natural de la razón de tasas.
- La varianza del logaritmo de una razón de tasas es aproximadamente igual a la suma de los inversos de los casos expuestos y no expuestos.

Derivación de la fórmula de aproximación de muestra grande para un intervalo de confianza para la razón de tasas

Al igual que con la razón de riesgos (**RR**), para obtener un intervalo de confianza de 95% para una razón de tasas (**RDI**), podemos estar inclinados inicialmente a utilizar el intervalo:

$$\hat{RDI} \pm 1{,}96\sqrt{\hat{Var}(\hat{RDI})}$$

Sin embargo, este intervalo **no** se recomienda porque la RDI calculada no presenta una distribución normal. El enfoque propuesto involucra trabajar con el cálculo de **lnRDI**, en lugar de la **RDI** en sí misma, ya que la transformación del logaritmo natural tiende a tornar una distribución asimétrica en una distribución aproximadamente normal. Se utiliza un procedimiento de dos pasos: en primer lugar, se obtiene un intervalo de confianza de 95% de muestra grande para **lnRDI** de la forma:

$$\ln\hat{RDI} \pm 1{,}96\sqrt{\hat{Var}(\ln(\hat{RDI})}$$

Entonces, ya que **RDI = exp(lnRDI)**, los valores superior e inferior deseados del intervalo de confianza de 95% para **RDI** se encuentran tomando los antilogaritmos de los valores inferior y superior para el intervalo de confianza para **lnRDI**. La fórmula resultante adopta la forma general:

$$\hat{RDI}\exp\left[\pm 1{,}96\sqrt{\hat{Var}(\ln\hat{RDI})}\right]$$

Sin embargo, para completar el cálculo debemos encontrar una expresión para $\hat{Var}(\ln\hat{RDI})$. Mediante lo que se denomina **aproximación de la serie de Taylor** de primer orden, la fórmula aproximada para esta varianza está dada por:

$$\hat{Var}(\ln\hat{RDI}) \approx \frac{1}{I_1} + \frac{1}{I_0}$$

donde $\mathbf{I_1}$ e $\mathbf{I_0}$ denotan la cantidad de casos expuestos y no expuestos, respectivamente. Por lo tanto, la fórmula general de aproximación para el intervalo de confianza del 95% de muestra grande puede escribirse como:

$$\hat{RDI}\exp\left[\pm 1{,}96\sqrt{\frac{1}{I_1} + \frac{1}{I_0}}\right]$$

Fórmula de aproximación de muestra grande para un IC para la diferencia de tasas

Se puede utilizar la siguiente fórmula para obtener un intervalo de confianza de 95% para una diferencia de tasas (es decir **DDI = TI_1 - TI_0**), donde **TI_1** y **TI_0** denotan las tasas de incidencia para las personas expuestas y no expuestas, respectivamente:

$$\hat{DDI} \pm 1{,}96\sqrt{\hat{Var}(\hat{DDI})}$$

donde $\hat{DDI} = \hat{TI}_1 - \hat{TI}_0$ y una fórmula de muestra grande aproximada para $\hat{Var}(\hat{DDI})$ está dada por:

$$\hat{Var}(\hat{DDI}) \approx \frac{I_1}{PT_1^2} + \frac{I_0}{PT_0^2}$$

donde **I_1**, **I_0**, **PT_1**, **PT_0** son nuevos casos y antecedentes de persona-tiempo en la siguiente disposición de datos para un estudio de cohorte con información de persona-tiempo:

Disposición de datos para el estudio de cohorte con información de persona-tiempo	**Expuestos**	**No expuestos**	Total
Casos nuevos (D)	I_1	I_0	I
Persona-tiempo	PT_1	PT_0	PT

La fórmula de aproximación de la varianza anterior se deriva de suponer que la cantidad de casos expuestos y no expuestos, es decir, **I_1** e **I_0**, tienen, cada uno, una **distribución de Poisson** de donde pueden calcularse las varianzas para cada grupo mediante **$I_1/(PT_1)^2$** e **$I_0/(PT_0)^2$** para los grupos de individuos expuestos y no expuestos, respectivamente.

Ejemplificamos el cálculo del IC de muestra grande para la DDI mediante los datos de un estudio de cohorte con información de persona-tiempo que compara a las personas con colesterol moderadamente alto con quienes tienen colesterol normal:

Estudio de cohorte sobre el colesterol con información de persona-tiempo	**Expuestos**	**No expuestos**	Total
Muertes (D)	26	14	40
Personas-años	36.581	68.239	104.820

Al sustituir los valores en este cuadro para **I_1**, **I_0**, **PT_1**, **PT_0** en la fórmula del intervalo de confianza se obtiene el siguiente IC de 95% para la diferencia de tasas **DDI**:

$$\frac{26}{36581}-\frac{14}{68239}\pm 1{,}96\sqrt{\frac{26}{36581^2}+\frac{14}{68239^2}}=\frac{7{,}1075}{10^4}-\frac{2{,}0516}{10^4}\pm 1{,}96\sqrt{\frac{1{,}9430}{10^8}+\frac{0{,}30065}{10^8}}$$

$$=\frac{5{,}0559}{10^4}\pm 1{,}96\sqrt{\frac{2{,}2437}{10^8}}=\frac{5{,}0559\pm 2{,}9359}{10^4}$$

Por lo tanto, los superior e inferior del IC para la diferencia de tasas son **L = 2,1200** por 10.000 personas-años y **U = 7,9918** por 10.000 personas-años, donde $\hat{DDI}$ = 5,0559 por 10.000 personas-años.

Auto-evaluación (12.20)

1. La siguiente expresión indica que el intervalo de confianza para una **???** se obtiene **???** un intervalo de confianza de muestra grande para el **???** de la razón de tasas.

2. La fórmula utilizada para la varianza **???** del logaritmo de la razón de tasas es una fórmula **???** , no una fórmula **???** , para esta varianza.

Opciones
aproximada **calculada** **exacta** **exponenciando** **exponenciación** **logaritmo natural**
odds ratio **razón de tasas** **prueba**

$$e^{\ln(\hat{RDI})\pm 1.96\sqrt{\hat{Var}\{\ln(\hat{RDI})\}}}$$

$$\hat{Var}\{\ln(\hat{RDI})\}\approx 1/I_1+1/I_0$$

Nomenclatura

Organización en forma de cuadro para los estudios de cohorte, casos y controles y prevalencia:

	Expuestos	No expuestos	Total
Enfermedad/casos	a	b	m_1
Sin enfermedad/controles	c	d	m_0
Total	n_1	n_0	n

Organización en forma de cuadro correspondiente a datos de cohorte con información de persona-tiempo:

	Expuestos	No expuestos	Total

Enfermedad (casos nuevos)	I_1	I_0	I
Sin enfermedad	-	-	-
Persona-tiempo total sin enfermedad	PT_1	PT_0	PT

χ^2	Chi cuadrado
χ^2_{MH}	Chi cuadrado de Mantel-Haenszel (CHI^2 MH)
IC	Intervalo de confianza
$\hat{I}A_0$	Incidencia acumulada o "riesgo" en los no expuestos (b/m_0)
$\hat{I}A_1$	Incidencia acumulada o "riesgo" en los expuestos (a/m_1)
DIA	Diferencia de incidencias acumuladas o diferencia de riesgos, $IA_1 - IA_0$; igual a diferencia de riesgos (DR)
E	Exponenciación o "antilogaritmo"; e≈2,71828
H_0:	Hipótesis nula
H_A	Hipótesis alternativa
I	Número de casos nuevos
I_0	Número de casos nuevos en los no expuestos
I_1	Número de casos nuevos en los expuestos
$D\hat{I}$	Densidad de incidencia (o "tasa") en la población (I/PT)
$D\hat{I}_0$	Densidad de incidencia (o "tasa") en los no expuestos (I_0/PT_0)
$D\hat{I}_1$	Densidad de incidencia (o "tasa") en los expuestos (I_1/PT_1)
DDI	Diferencia de densidades de incidencia o diferencia de tasas, $DI_1 - DI_0$
RDI	Razón de densidades de incidencia o razón de tasas: DI_1 / DI_0; igual a la razón de tasas de incidencia (RTI); IDR, por sus siglas en inglés
RTI	Razón de tasas de incidencia o razón de tasas: DI_1 / DI_0; igual a la razón de densidades de incidencia (RDI)
L	Límite inferior del intervalo de confianza
Ln	Logaritmo natural
N	Tamaño de la población en estudio
MH	Mantel-Haenszel
OR	*Odds ratio*: ad/bc
P	Proporción poblacional
$\hat{p}$	Proporción muestral *o* promedio ponderado de dos proporciones
P o Pr	Probabilidad
P_1	Proporción con resultado en expuestos
P_0	Proporción con resultado en no expuestos
PT	Persona-tiempo sin enfermedad
PT_0	Persona-tiempo sin enfermedad en los no expuestos
PT_1	Persona-tiempo sin enfermedad en los expuestos
DR	Diferencia de riesgos: el riesgo en los expuestos menos el riesgo en los no expuestos; igual a la diferencia de incidencia acumulada (DIA)
RR	Razón de riesgos: el riesgo en los expuestos dividido por el riesgo en los no expuestos; igual a la razón de incidencia acumulada (RIA)
ROR	*Odds ratio* de riesgo

S	Error estándar
T	Estadístico de prueba T
T*	Estadístico de prueba T calculada
U	Límite superior del intervalo de confianza
Var	Varianza
Z	Estadístico de prueba Z

Fórmulas

Pruebas estadísticas

Estadístico T para la diferencia entre dos proporciones muestrales

$$T = \frac{\hat{p}_1 - \hat{p}_0}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_0}\right)}} \qquad \textit{donde} \qquad \hat{p} = \frac{n_1\hat{p}_1 + n_0\hat{p}_0}{n_1 + n_0}$$

Estadístico chi cuadrado para un cuadro de 2 x 2

$$\chi^2 = \frac{n(ad-bc)^2}{n_1 n_0 m_1 m_0}$$

Estadístico chi cuadrado de Mantel-Haenszel para un cuadro de 2 x 2

$$\chi^2_{MH} = \frac{(n-1)(ad-bc)^2}{n_1 n_0 m_1 m_0}$$

Estadístico T para la diferencia entre dos tasas muestrales

$$T = \frac{I_1 - Ip_0}{\sqrt{Ip_0(1-p_0)}} \qquad \textit{donde} \qquad p_0 = \frac{PT_1}{PT}$$

Estadístico chi cuadrado para comparar dos tasas

$$\chi^2 = \frac{(I_1 PT_0 - I_0 PT_1)^2}{IPT_1 PT_0}$$

Intervalos de confianza

Intervalo de confianza de 95% de muestra grande para la diferencia entre dos proporciones (diferencia acumulada o diferencia de riesgos)

$$\hat{p}_1 - \hat{p}_0 \pm 1{,}96\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_0(1-\hat{p}_0)}{n_0}}$$

Intervalo de confianza de 95% de muestra grande para la razón de riesgos (razón de dos proporciones)

$$R\hat{R}\ \exp\left[\pm 1{,}96\sqrt{\frac{1-\hat{I}A_1}{n_1\hat{I}A_1} + \frac{1-\hat{I}A_0}{n_0\hat{I}A_0}}\right]$$

Intervalo de confianza de 95% de muestra grande para el odds ratio

$$\hat{OR}\ \exp\left[\pm 1{,}96\sqrt{\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}}\right]$$

Intervalo de confianza de 95% de muestra grande para la razón de densidades de incidencia

$$\hat{RDI}\ \exp\left[\pm 1{,}96\sqrt{\frac{1}{I_1}+\frac{1}{I_0}}\right]$$

Intervalo de confianza de 95% de muestra grande para la diferencia de densidades de incidencia

$$\hat{DDI} \pm 1{,}96\sqrt{\frac{I_1}{PT_1^2}+\frac{I_0}{PT_0^2}}$$

Referencias

Referencias sobre estadística básica

Existen muchos textos para principiantes que abarcan los conceptos y métodos fundamentales de la estadística o bioestadística. (Nota: la bioestadística se refiere a las aplicaciones de la estadística en las ciencias biológicas y de la salud). En consecuencia, no es nuestra intención aquí brindar una lista exhaustiva de la totalidad de dichos textos, sino sugerir algunas referencias que este autor ha utilizado o recomienda. Entre ellas, proponemos primero que considere el texto del CD ROM ActivStats de Velleman, que tiene el mismo formato que ActivEpi. Las referencias sugeridas incluyen:

Kleinbaum DG, Kupper LL, Muller KA, Nizam A *Applied Regression Analysis and Other Multivariable Methods*, 3era edición. Duxbury Press, 1998. (El capítulo 3 contiene una revisión compacta de los conceptos y métodos de estadística básica)

Moore D. *The Active Practice of Statistics*. WH Freeman Publishers, 1997. (Este libro está diseñado específicamente para ser utilizado con ActivStats y ambos tienen amplias coincidencias).

Remington RD y Schork MA. *Statistics with Applications to the Biological and Health Sciences*. Prentice Hall Publishers, 1970.

Velleman P. ActivStats - *A Multimedia Statistics Resource* (CDROM), Addison-Wesley Publishers, 1998.

Weiss N. *Introductory Statistics*, 6ª edición. Addison-Wesley Publishing Company, Boston, 2002.

Referencias sobre análisis simple

Fleiss JL. *Confidence Intervals for the odds ratio in case-control studies: the state of the art*. J Chronic Dis 1979; 32(1-2): 69-77.

Goodman SN. *Toward evidence-based medical statistics. 1: The P value fallacy*. Ann Intern Med 1999; 130(12): 995-1004.

Goodman SN. *Toward evidence-based medical statistics. 2: The Bayes factor*. Ann Intern Med 1999; 130(2): 1005-13.

Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic Research: Principles and Quantitative Methods*, Capítulo 15, John Wiley and Sons, 1982.

Miettinen O. *Estimability and estimation in case-referent studies.* Am J Epidemiol 1976; 103(2): 226-235.

Mantel N, Haenszel W. *Statistical aspects of the analysis of data from retrospective studies of disease*. J Natl Cancer Inst 1959; 22(4): 719-48.

Thomas DG. *Exact confidence limits for an odds ratio in a 2x2 table*. Appl Stat 1971; 20: 105-10.

Ejercicios

EPI-1. *Odds ratio* de exposición

Los siguientes datos provienen de un estudio de casos y controles diseñado para investigar la teoría que sostiene que cierto factor de estudio (E) es el determinante de cierta enfermedad (D) poco frecuente. Se identificó un grupo representativo de casos incidentes de la enfermedad que surgieron en una población dada durante un período de cinco años. Luego se compararon estos casos con una muestra aleatoria de una cantidad igual de no casos de la misma población.

	Expuestos	**No expuestos**	Total
Casos	27	23	50
Controles	18	32	50
Total	45	55	100

a. Calcule el *odds ratio* de exposición (EOR) para los datos. ¿Podemos decir (utilizando sólo esta estimación puntual) que hay evidencia estadística de una asociación entre el factor de estudio y la enfermedad? Desarrolle su respuesta.
b. Evalúe la significación estadística del valor observado de EOR, empleando una prueba chi cuadrado de Mantel-Haenszel. Determine el valor p de una cola e interprete el resultado.
c. Calcule el intervalo de confianza del 92% (!) de muestra grande para el EOR.
d. ¿A qué conclusión llega sobre la relación E-D en función de los resultados encontrados en los ítems a, b y c?

EPI-2. Medida en diferencias vs. medida en razones

Los siguientes datos derivan de un estudio sobre cáncer de mama entre mujeres con tuberculosis. Las mujeres que fueron expuestas en repetidas ocasiones a múltiples fluoroscopías de rayos x fueron comparadas con mujeres no tan expuestas:

	Exposición a la radiación		
	Sí	No	Total
Cáncer de mama	41	15	56
Personas-años	28.010	19.017	47.027

a. Calcule una medida de asociación adecuada en forma de diferencia para estos datos e interprete el resultado.
b. Calcule un intervalo de confianza de 90% para la medida en forma de diferencia del ítem a. ¿A qué conclusión llega sobre la base de este resultado?
c. Calcule una medida de asociación adecuada en forma de razón para estos datos e interprete el resultado.
d. Calcule un intervalo de confianza de 90% para la medida en forma de razón del ítem c.
e. ¿Su conclusión en el ítem d sobre la asociación entre la exposición y el resultado es la misma que antes (es decir, igual al ítem b)?

EPI-3. Prueba de hipótesis

Helicobacter pylori (HP) es una bacteria que afecta las células que recubren el estómago, causando una inflamación aguda y crónica. El organismo se considera un factor causal de la enfermedad de úlcera péptica y se le ha vinculado epidemiológicamente al desarrollo de adenocarcinoma gástrico. Un grupo de investigadores tenía interés en evaluar la relación entre el consumo de alcohol y la infección por HP. Identificaron a 300 individuos cuya sangre contenía anticuerpos contra el HP, 63 de los cuales consumían alcohol. De las 267 personas que no tenían anticuerpos contra el HP, 91 consumían alcohol.

a. Utilice esta información para completar el siguiente cuadro de 2 x 2:

	Consumía alcohol	No consumía alcohol
HP		
Sin HP		

b. Calcule e interprete la medida de asociación adecuada en forma de razón para estos datos.
c. Utilice estos datos para realizar una prueba de hipótesis de la asociación entre el consumo de alcohol y la infección por HP. Asegúrese de expresar las hipótesis nula (H_0) y alternativa (H_A) y brinde un valor p adecuado. ¿A qué conclusión llega?
d. Calcule un intervalo de confianza de 95% para la medida en el ítem b anterior.

EPI-4. Prueba exacta de Fisher

5. El siguiente cuadro contiene los datos de un estudio de casos y controles que relaciona una exposición (E) con una enfermedad (D).

	Expuestos	**No expuestos**	
Enfermos	5	6	11
Sanos	3	2	5
	8	8	16

Si la significación de la asociación entre la exposición y la enfermedad se tuviera que evaluar mediante la prueba exacta de Fisher, ¿cuántos términos se sumarían para calcular el valor p?

i. 11
ii. 8
iii. 4
iv. 7

a. Realice los cálculos de la prueba exacta de Fisher. ¿Qué conclusión puede sacar de esta prueba?

EPI-5. Razón de densidades de incidencia

Suponga que desea analizar los datos de cohorte de persona-tiempo dados en el siguiente cuadro:

	Exposición		
	Sí	No	Total
Enfermedad	50	150	200
Personas-años	10.000	40.000	50.000

a. ¿Cuál es la razón de densidades de incidencia?
b. Realice una prueba de muestra grande de la hipótesis (H_0: RDI = 1).
c. Brinde una fórmula para calcular un intervalo de confianza de 95% para la RDI. Al plantear esta fórmula, ponga valores numéricos en el lugar de los símbolos que describen la fórmula general del intervalo de confianza en esta situación.
d. Calcule el intervalo de confianza de 95% descrito en el ítem c.
e. En función de su respuesta a la pregunta anterior, ¿a qué conclusión llega sobre la relación E-D?

EPI-6. Tamaño de muestra: ensayo clínico

Se planea llevar a cabo un ensayo clínico para evaluar la eficacia de una nueva terapia para la osteoporosis, la cual se comparará con la terapia convencional, que ciertamente produce una respuesta positiva (mayor densidad ósea) en 20% de los pacientes. Los datos preliminares sugieren que la nueva terapia puede generar una respuesta favorable en el doble de la proporción de pacientes, en comparación con la convencional. Si esto es verdad, ¿cuántos individuos deberán participar en el ensayo clínico para mostrar una respuesta estadísticamente significativa? Suponga que $\alpha = 0{,}05$ (2 colas) y $\beta = 0{,}10$ (1 cola). Incluya el desarrollo de sus cálculos. (Sugerencia: utilice la siguiente fórmula del tamaño de muestra).

$$\mathbf{n} = \frac{(\mathbf{Z}_{1-\alpha/2} + \mathbf{Z}_{1-\beta})^2\, \overline{\mathbf{pq}}(\mathbf{r}+1)}{[\mathbf{p}_2(\mathbf{RR}-1)]^2\mathbf{r}}$$

EPI-7. Tamaño de muestra: estudio de casos y controles

Se utilizaron datos del estudio de seroconversores de tres continentes para un análisis de casos y controles de la asociación potencial entre la condición de VIH y el uso de sustancias. Se compararon recientes seroconversores al VIH en sujetos con resultado negativo en la prueba del VIH; se les preguntó a todos los individuos sobre el uso de sustancias durante el año anterior al ingreso en el estudio. El cuadro siguiente resume los datos sobre el uso de anfetaminas:

	Uso de anfetaminas		
	Sí	**No**	**Total**
VIH+	78	267	345
VIH-	35	310	345

a. Calcule el *odds ratio* de exposición (EOR) para estos datos.
b. ¿Cuántos casos y controles habrían necesitado los investigadores para poder detectar la misma EOR si hubiesen decidido seleccionar **r** = 3 controles por caso? Suponga que **α = 0,05** (2 colas) y **potencia = 90%**.

Sugerencia: puede utilizar la siguiente fórmula del tamaño de muestra:

$$\mathbf{n} = \frac{(\mathbf{Z}_{1-\alpha/2} + \mathbf{Z}_{1-\beta})^2 \, \overline{\mathbf{pq}}(\mathbf{r}+1)}{[\mathbf{p}_1 - \mathbf{p}_2)]^2 \mathbf{r}}$$

EPI-8. Factores que afectan el tamaño de la muestra

Un epidemiólogo planea efectuar un estudio para evaluar la asociación entre una exposición (E) y una enfermedad (D). Para cada una de las siguientes situaciones, indique si el tamaño de muestra requerido aumentaría, disminuiría o se mantendría igual.

a. Un aumento de α, de 0,05 a 0,10.
b. Una disminución del impacto esperado de una terapia nueva (en comparación con la convencional).
c. Un mayor deseo de arriesgar un error de tipo I.

EPI-9. Tamaño de muestra y potencia

A los autores de un estudio epidemiológico les preocupaba que, debido al tamaño pequeño de la muestra, su estudio tuviese baja potencia. Describa brevemente el fundamento de esa preocupación, es decir, ¿cuál es el problema potencial asociado a la baja potencia en un estudio epidemiológico?

Respuestas a las preguntas de estudio y auto-evaluaciones

P12.1

5. La RR, la razón de dos proporciones o el OR, la razón de dos probabilidades, cada una con la forma p/(1-p).
6. La hipótesis nula normal para una razón de riesgos es RR = 1, donde 1 es el valor nulo de la razón de riesgos.
7. La hipótesis nula normal para un *odds ratio* es OR = 1, donde 1 es el valor nulo del *odds ratio*.

P12.2

1. El cálculo combinado de la proporción común está dado por {(0,7972 x 1667) + (0,2565 x 425)} / {1667 + 425} = 0,6874.

P12.3

1. La hipótesis alternativa de dos colas A1, en la esquina abajo a la derecha.
2. Ya que la curva normal es simétrica, las colas izquierda y derecha tienen igual área. Por lo tanto, calcule el área de una cola y luego multiplique por dos.
3. Para una hipótesis alternativa de una cola superior, el valor p es el área bajo la curva normal a la derecha de T* = 21,46. Del cuadro, encontramos que P<0,0001. Por lo tanto, si la hipótesis nula fuese verdadera y nuestra hipótesis alternativa hubiese sido de una cola, nuestros resultados tendrían una probabilidad de ocurrir menor al 0,01%.
4. El valor p es el doble del área más allá del valor de 21,46 bajo la curva, por lo tanto, P < 0,0002. Esto es así porque un valor de T* menor que –21,46 en la cola izquierda de la distribución normal también representaría un peor valor bajo la hipótesis nula que el valor

T* = 21,46 que, en verdad, se observó. Por lo tanto, si la hipótesis nula fuese verdadera y nuestra hipótesis alternativa hubiese tenido dos colas, nuestros resultados tendrían una probabilidad de ocurrir menor al 0,02%.

5. Sí. El valor p, que representa la probabilidad de que ocurrieran nuestros resultados si la hipótesis nula fuese verdadera, es extremadamente pequeño.
6. Debería llegar a la conclusión de que la prueba es significativa, es decir, rechazar la hipótesis nula y concluir que las proporciones para los hombres y las mujeres son significativamente distintas.
7. a) $P > 0,01$: no rechace la H_0. b) $P = 0,023$: rechace la H_0.
8. c) $P > 0,001$: rechace la H_0. d) $P = 0,54$: no rechace la H_0. b) $P = 0,0002$: rechace la H_0.
9. a) $P > 0,01$: no rechace la H_0. b) $P = 0,023$: no rechace la H_0.
10. c) $P > 0,001$: rechace la H_0. d) $0,01 < P < 0,05$: no rechace la H_0. b) $P = 0,0002$: rechace la H_0.

P12.4

1. la prueba de hipótesis
2. la estimación de intervalos
3. la prueba de hipótesis
4. la estimación de intervalos
5. nivel de significación
6. hipótesis alternativa
7. región de rechazo, hipótesis nula, estadísticamente
8. valor p, hipótesis nula
9. no rechazar
10. 0,01; 0,001

P12.5

1. El error estándar utilizado aquí no supone que la hipótesis nula sea verdadera. Por lo tanto, la varianza para cada proporción debe calcularse en forma separada mediante este valor de proporción muestral.

P12.6

1. 100% de confianza.
2. El intervalo de confianza de 99% sería el más amplio y el intervalo de confianza del 90% sería el más acotado.
3. El conjunto de datos A brinda el cálculo más preciso porque su intervalo de confianza es más acotado que el del conjunto de datos B.

P12.7

1. error estándar calculado
2. diferencia
3. parámetro poblacional
4. límite superior
5. 95%
6. 5%

P12.8

1. Sí, porque 7,04 es más grande que 6,635, que es el punto porcentual 0,99 de la distribución chi cuadrado con 1 grado de libertad.
2. No, porque 7,04 es menor que 7,879, que es el punto porcentual 0,995 de la distribución chi cuadrado con 1 grado de libertad.
3. Sí, porque el valor p para una prueba de dos colas es el área por arriba de 7,04, la cual es menor que el área por arriba de 6,635, que es 0,01.
4. No, porque el valor p es el área por arriba de 7,04, que es mayor que el área por arriba de 7,879, que es 0,005.
5. Sí. Porque la prueba de una cola superior mediante la curva normal fue significativa al nivel de 1%.
6. Sí. El valor p de una cola mediante la distribución normal fue 0,0040, que es menor que 0,005.

P12.9

1. El valor p correcto es el área bajo la curva normal por debajo de +2,65, que es 1 menos el área por arriba de 2,65, cuyo cálculo es 1 – 0,004 = 0,996. En forma equivalente, el valor 0,996 puede obtenerse de la curva chi cuadrado, tomando la mitad de 0,008 y restándolo de 1, es decir, 1 – (0,008 / 2) = 0,996. Por lo tanto, sería incorrecto tomar la mitad del área de 0,008 bajo la curva chi cuadrado y también sería incorrecto tomar la mitad de 1 menos 0,008.

2. No hay que rechazar la hipótesis nula porque el valor p es muy alto.
3. El valor p correcto es el área bajo la curva normal debajo de –2,65, que es 0,004. Esto también puede obtenerse tomando la mitad del área por arriba del valor chi cuadrado de 7,04, que es 0,008/2 = 0,004.
4. Rechazar la hipótesis nula porque el valor p es muy pequeño, y mucho más pequeño que 0,05 y 0,01.

P12.10

1. $b = m_1 – a$; $c = n_1 – a$; $d = m_0 – n_1 + a$
2. b = 2; c = 1; d = 6
3. b = 1; c = 0; d = 7
4. RR = (5/5)/(1/8) = 8
5. Sí
6. No. Un valor de la celda **a** mayor que 5 no es posible porque entonces se excedería el total marginal de columna fijo supuesto de 5.
7. P(a = 4 ó 5|RR = 1) = P(a = 4|RR = 1) + P(a = 5|RR = 1). Esto nos dice qué poco probables son nuestros resultados observados bajo la hipótesis nula. Es el valor p para probar esta hipótesis.
8. Ya que p es mayor a 0,05, no podemos rechazar la hipótesis nula de que RR = 1 al nivel de significación de 0,05.

P12.11

1. El valor p de una cola es 0,0630 / 2 = 0,0315.
2. El supuesto de muestra grande no se cumple.
3. El tamaño de muestra n = 13 no es grande, de forma que la diferencia entre n = 13 y n – 1 = 12 tiene un efecto más fuerte sobre el cálculo de cada estadístico de prueba.

P12.12

1. Este intervalo de confianza tiene una probabilidad de 9% de abarcar la verdadera razón de riesgos que compara a *los que continuaron fumando* con *los que dejaron de hacerlo*. Aunque el intervalo de confianza contiene el valor nulo de 1, es lo suficientemente amplio como para sugerir que la verdadera razón de riesgos podría ser cercana a 1 o tan grande como 3,6. En otras palabras, la estimación puntual de 2,1 es un poco imprecisa; además, el verdadero efecto de dejar de fumar después de un ataque cardíaco puede ser muy débil o muy fuerte.

P12.13

1. 520/1863 = 0,279
2. 2,23: T* = (0,303 – 0,256) / [sqrt{0,279 * (1-0,279) * [(1/915) + (1/948)]}] = 2,23
3. 0,013
4. 5%, 1%
5. 4,970: chi cuadrado MH = (1862 * $2{,}23^2$)/1863 = 4,970
6. 0,026
7. grande

P12.14

1. Determine el área bajo la distribución chi cuadrado con 1 gl por arriba del valor 4,67. El valor p es la mitad de esta área.
2. El área bajo la curva chi cuadrado por arriba de 4,67 está entre 0,025 y 0,05. En consecuencia, después de dividir cada límite por 2, el valor p para esta prueba se encuentra entre 0,0125 y 0,025.
3. Al nivel de 0,05, sí (significativo). Al nivel de 0,01, no.

P12.15

1. b = 6, c = 3, d = 6
2. OR = (4*6)/(6*3) = 1,3
3. No. Un OR de 1,3 es más cercano a 1 que un OR de 3,5.
4. Sí; los valores de la celda **a** de 6 ó 7 darían *odds ratios* mayores que 3,5.
5. Porque Pr(a = 5, 6 ó 7|OR = 1) = Pr(a = 5|OR = 1) + Pr(a = 6|OR = 1) + Pr(a = 7|OR = 1) nos dice con qué poca frecuencia es nuestro resultado observado según la hipótesis nula.
6. Ya que p es mucho más grande que 0,05, no rechazamos la hipótesis nula de que OR = 1 al nivel de significación de 0,05. En otras palabras, no hay evidencia estadísticamente

significativa en este conjunto de datos hipotéticos para decir que la ingesta de hamburguesa cruda fue la fuente del brote.

P12.16

1. El intervalo de confianza, que va desde 1,1 a 9,1, es muy amplio, lo que indica una falta de precisión en el *odds ratio* calculado de 3,2. Por lo tanto, aunque el *odds ratio* calculado es estadísticamente significativo al nivel de 0,05, el verdadero *odds ratio* podría ser muy débil (es decir, cercano a 1) o muy fuerte (o sea, apenas mayor que 9).

P12.17

1. los márgenes de las filas
2. los márgenes de las columnas
3. misma que
4. misma que
5. 5
6. 4
7. 7
8. intervalo de confianza, exponenciación, intervalo de confianza, *odds ratio*

P12.18

1. Hay que determinar el área bajo la distribución chi cuadrado con 1 gl por arriba del valor 15,952. El valor p es la mitad de esta área.
2. El área bajo la curva chi cuadrado por arriba de 15,952 es menor que 0,0005. En consecuencia, después de dividir por 2, el valor p para esta prueba es menor que 0,00025, el cual es extremadamente pequeño.
3. Los individuos con colesterol moderadamente alto tienen una tasa de mortalidad significativamente mayor que las personas con colesterol normal.
4. Z al cuadrado = 15,952; por lo tanto, Z = sqrt(125,952) = 3,994. Este Z calculado debe ser un número positivo porque las personas con colesterol moderadamente alto tienen una tasa de mortalidad mayor (7,11 por 10.000 personas-años) que aquellas con colesterol normal (2,05 por 10.000 personas-años).

P12.19

1. Los límites de confianza, que oscilan entre 1,81 y 6,63, son bastante amplios, lo que indica imprecisión en el valor de la razón de tasas de 3,46. La verdadera razón de tasas puede ser cercana a 2 o mayor que 6. Sin embargo, estamos 95% seguros de que la verdadera razón de tasas es mucho mayor que 1.

P12.20

1. razón de tasas, exponenciando, logaritmo natural
2. calculada, aproximada, exacta

LECCIÓN 13

CONTROL DE FACTORES EXTERNOS

*En lecciones anteriores se analizaron y ejemplificaron varios conceptos importantes sobre el **control** de variables adicionales (**externas**) cuando se evalúa la relación entre una variable de exposición y una variable de resultado de salud. En esta lección revisamos brevemente estos conceptos y luego brindamos una descripción general de varias opciones para el proceso de control, las que se encuentran disponibles tanto en la etapa de diseño como en la fase de análisis de un estudio.*

13-1 Control de factores externos

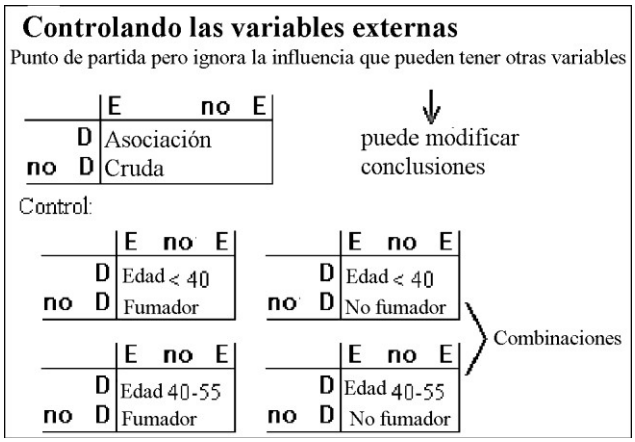
¿Qué queremos decir con el término "control"?

Suponga que estudiamos la existencia de una asociación entre la exposición a un químico tóxico y el desarrollo de cáncer de pulmón en una industria química. Para responder a esta pregunta adecuadamente, desearíamos aislar el efecto del químico de la posible influencia de otras variables, en especial, la edad y la condición de tabaquismo, dos factores de riesgo conocidos del cáncer de pulmón. Es decir, nuestro objetivo es determinar si la exposición al químico aporta algo más a los efectos de la edad y el tabaquismo en el desarrollo de cáncer de pulmón.

Comúnmente, las variables como la edad y el tabaquismo en este ejemplo reciben el nombre de **variables de control**. Cuando evaluamos la influencia de dichas variables de control sobre la relación **E→D**, decimos que estamos **controlando las variables externas**. Por "externas" simplemente nos referimos a que estamos considerando variables distintas de **E** y **D** que no son de interés primordial, pero que, sin embargo, podrían influir en nuestras conclusiones sobre la relación **E→D**.

En general, realizamos un análisis simple de la relación entre la exposición y la enfermedad como punto de partida para análisis más complicados que, probablemente, tendremos que llevar a cabo. Un análisis simple nos permite ver la **asociación cruda** entre la exposición y la enfermedad y, por lo tanto, nos da la posibilidad de obtener conocimientos preliminares sobre la relación entre la exposición y la enfermedad. Lamentablemente, el análisis simple, por definición, ignora la influencia que pueden tener otras variables, distintas de la exposición, sobre la enfermedad. Si existen otras variables que con seguridad predicen la enfermedad, entonces probablemente sea necesario modificar las conclusiones sugeridas por el análisis simple cuando se consideran dichos factores de riesgo.

En consecuencia, cuando controlamos variables externas, evaluamos el efecto de la exposición **E** sobre la enfermedad **D** en diferentes combinaciones de valores de las variables que examinamos. Cuando corresponde, evaluamos la relación **E→D** general mediante la fusión de la información en las distintas combinaciones de valores de control.



Preguntas de estudio (P13.1)

Considere un estudio de casos y controles que busca evaluar si existe una asociación entre cierto químico tóxico (**E**) y el desarrollo de cáncer de pulmón (**D**) en una industria química. Suponga que deseamos tener en cuenta el control de la edad y la condición de tabaquismo. Presuma que clasificamos la edad en tres grupos: menos de 40 años, entre 40 y 55 años y más de 55 años. También categorizamos la condición de tabaquismo como "fumó alguna vez" y "nunca fumó".

1. ¿Cuántas combinaciones de categorías de edad y tabaquismo existen?

Dos tipos de análisis combinados con estos datos son:

- agrupar en cuadros de 2 x 2 para estas combinaciones en un cuadro "combinado" general y calcular el *odds ratio* para este cuadro combinado; y
- calcular un *odds ratio* para cada cuadro de 2 x 2 correspondiente a cada combinación y luego sacar el promedio de estos *odds ratios* separados de alguna manera.

2. ¿Cuál de estos dos análisis controla la edad y el tabaquismo?

Cuando se considera el control de variables externas, surgen varias preguntas: en primer lugar, ¿por qué queremos hacer un control? Es decir, ¿qué logramos con dicho control? ¿Cuáles son las diferentes opciones que están disponibles para realizarlo? ¿Qué opción de control deberíamos elegir en nuestro estudio? ¿Cuáles de las variables que se consideran deberían controlarse en realidad? ¿Qué deberíamos hacer si tenemos tantas variables para controlar que nos quedamos sin datos suficientes? Estas preguntas se abordarán en las siguientes actividades.

Preguntas sobre el control:

¿Por qué controlar?

¿Cuáles son las opciones para el control?

¿Qué opción de control deberíamos elegir?

¿Cuáles de las variables deberían controlarse en realidad?

¿Qué deberíamos hacer sin datos suficientes?

Resumen

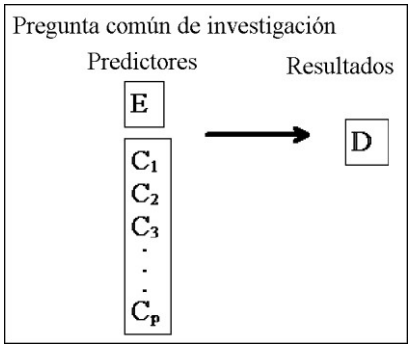
- Cuando evaluamos una relación E→D, determinamos si **E** aporta algo más a los efectos de otros predictores conocidos (es decir, **variables de control**) de **D**.
- Cuando evaluamos la influencia de las variables de control, decimos que estamos **controlando las variables externas**.
- El análisis simple ignora el control de variables externas.
- El control evalúa la relación **E→D** en las combinaciones de valores de las variables de control.
- Cuando corresponde, el control examina la relación **E→D general** después de considerar las variables de control.

Razones para el control

Una pregunta típica de una investigación epidemiológica evalúa la relación entre una o más variables de resultado de salud, **D**, y una o más variables de exposición, **E**, teniendo en cuenta los efectos de otras variables, **C**, que, según ya se sabe, pueden predecir el resultado.

D = resultado de salud
E = variables de exposición
C = variables de control

Cuando hay sólo una **D** y una **E** y existen diversas variables de control, la pregunta más común de investigación puede expresarse según se muestra a continuación, donde la flecha indica que la variable **E** y las variables **C** que se encuentran a la izquierda son las que se evaluarán como predictores del resultado **D**, a la derecha.



¿Por qué las variables **C** están aquí? Es decir, ¿cuáles son las razones por las cuales queremos controlar las **C**? Uno de los motivos para controlarlas es garantizar que cualquier efecto que encontremos de la variable de exposición no pueda explicarse por las variables que con seguridad tienen algún efecto sobre el resultado de salud. En otras palabras, queremos cerciorarnos de que hemos explicado la posible confusión de la relación **E→D** en razón de la influencia de factores de riesgo conocidos para el resultado de salud.

Una segunda causa para efectuar el control es garantizar que eliminamos cualquier variabilidad en el cálculo de efecto en E→D que aportan otros predictores conocidos. Como resultado del control, podríamos ganar **precisión** en nuestro cálculo de efecto, por ejemplo, un intervalo de confianza acotado. En algunas situaciones puede haber una pérdida de precisión cuando se controlan los factores de confusión.

Una tercera razón para el control es permitirnos evaluar si el efecto de la exposición puede variar según las características de otros predictores. Por ejemplo, puede haber un efecto fuerte de la exposición para los fumadores, pero ningún efecto de exposiciones para los no fumadores. Este tema atañe a la **interacción** o

modificación de efecto.

Éstos son los tres motivos principales del control: 1) controlar la confusión; 2) aumentar la precisión; y 3) explicar la posibilidad de **modificación de efecto**.

1. Confusión
2. Precisión
3. Modificación de efecto

Estas tres razones son importantes; sin embargo, hay un orden establecido para el momento en que se deben considerar en el curso de un análisis.

¿Orden de un análisis?

1. Modificación de efecto
2. Confusión (validez)
3. Precisión (error aleatorio)

La posibilidad de **modificación de efecto** debe contemplarse en primer término, porque si hay una modificación de efecto, un único cálculo ajustado que controla la confusión puede ocultar el hecho de que la relación **E→D** es distinta para diferentes categorías de una variable de control.

Una vez que se aborda la modificación de efecto o se descubre que no existe ninguna, debe tenerse en cuenta la **confusión**, especialmente en términos de aquellas variables de control que **no** se consideran modificadores de efecto. La confusión debe abordarse antes que la precisión porque la confusión atañe a la **validez** de un cálculo. La **precisión** sólo se refiere al **error aleatorio**. Preferiríamos tener un cálculo válido a tener un intervalo de confianza acotado alrededor de un cálculo sesgado.

Preguntas de estudio (P13.2)

Considere nuevamente un estudio de casos y controles que busca evaluar si existe una asociación entre cierto químico tóxico (**E**) y el desarrollo de cáncer de pulmón (**D**) en una industria química, donde se desea controlar la edad y la condición de tabaquismo. Asimismo, suponga que clasificamos la edad en tres grupos: menos de 40 años, entre 40 y 55 años y más de 55 años; además, asuma que categorizamos la condición de tabaquismo en "fumó alguna vez" y "nunca fumó".

1. Verdadero o falso. Podemos evaluar la confusión de la edad o la condición de tabaquismo determinando si la relación **E→D** difiere dentro de distintas categorías de edad o tabaquismo o de ambas combinadas.
2. Verdadero o falso. Se obtendrá un cálculo más preciso del *odds ratio* (OR) para la asociación **E→D** si controlamos tanto la edad como la condición de tabaquismo.

Suponga que cuando controlamos la edad y la condición de tabaquismo, el OR para la asociación **E→D** es 3,5, pero cuando ignoramos tanto la edad como la condición de tabaquismo, el OR crudo correspondiente es 1,3.

3. ¿Esto indica confusión, precisión o interacción?

Presuma que cuando controlamos la edad y la condición de tabaquismo, el OR ajustado es 3,5, como antes, con un intervalo de confianza del 95% oscilando entre 2,7 y 4,5, pero que el OR crudo de 1,3 tiene un intervalo de confianza del 95% que va de 1,1 a 1,5.

4. ¿Cuál de estos dos OR es más apropiado?

Suponga, además de la información anterior, que sabe que el OR que relaciona **E** con **D** es 5,7 para los fumadores, pero sólo 1,4 para los no fumadores.

5. ¿Desearía controlar confusión por edad y por tabaquismo?

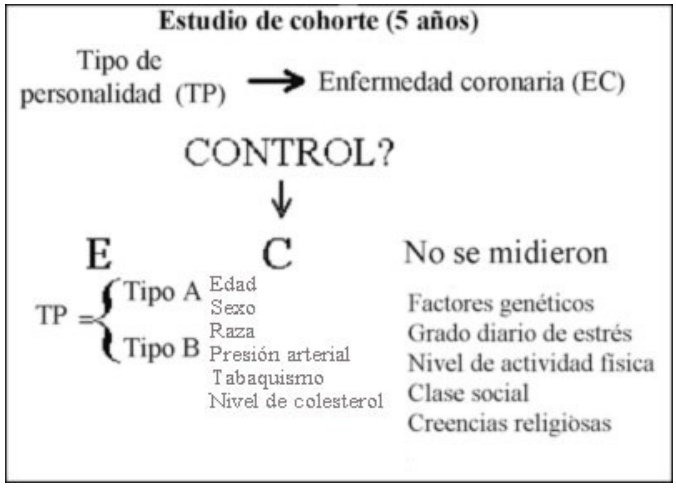
Resumen

- La pregunta típica de una investigación epidemiológica evalúa la relación entre una o más variables **E** y una o más variables **D** controlando varias variables **C**.
- Las tres razones para efectuar el control son la confusión, la precisión y la modificación de efecto.
- La posibilidad de modificación de efecto debe considerarse en primer lugar, seguida de la confusión y, luego, de la precisión.

Opciones para el control

Opciones de diseño

Suponga que desea evaluar la posible asociación del tipo de personalidad con la enfermedad coronaria. Usted: (1) decide realizar un estudio de cohorte para comparar el riesgo de enfermedad coronaria (EC) en un grupo de individuos con patrón de personalidad tipo A al riesgo correspondiente en individuos que poseen personalidad tipo B; (2) planea hacer un seguimiento de ambos grupos durante la misma cantidad de tiempo, digamos, 5 años. Su variable de exposición, **E**, es dicotómica; (3) reconoce que edad, sexo, raza, presión arterial, tabaquismo y nivel de colesterol son factores de riesgo de EC importantes que debe observar o medir para controlar en su estudio; (4) finalmente, admite que hay otros factores como, por ejemplo, factores genéticos, grado diario de estrés, nivel de actividad física, clase social y creencias religiosas, y le gustaría considerarlos, pero no posee los recursos para medirlos.



¿Cómo efectúa el estudio para controlar cualquiera o todas las variables que acabamos de mencionar? Es decir, ¿cuáles son sus opciones para el control? Algunas de ellas deben considerarse en la etapa de diseño del estudio, **antes** de recabar los datos. Otras se desarrollan **durante la etapa de análisis**, después de que se hayan recabado los datos. Es posible elegir más de una opción en el mismo estudio.

Opciones de diseño - antes de recabar los datos

Opciones para análisis- después de recabar los datos

Esta actividad se centra sólo en las **opciones de diseño**. Una de ellas es la **aleatorización**. Tal vez desee asignar en forma aleatoria un individuo sin enfermedad inicial al tipo de personalidad A o al tipo B. Si pudiese hacerlo, entonces las variables que desea controlar, incluso las que no mide realmente, podrían distribuirse en forma similar para ambos grupos de exposición. Sin embargo, esta opción no está disponible aquí. No se puede forzar a que la gente tenga un tipo de personalidad o el otro; cada cual es como es.

La segunda opción de diseño se denomina **restricción**. Esto significa especificar un rango acotado de valores para una o más variables de control. Por ejemplo, puede decidir restringir el estudio a los afroamericanos, a las mujeres, a las personas de más de 50 años o a las tres categorías, sin restringir ninguna de las otras variables de la lista.

Preguntas de estudio (P13.3)

1. ¿Cuál es una de las desventajas de limitar el estudio sólo a las mujeres?
2. Además de la generalización a otros grupos etarios, ¿cuál es otra de las desventajas de limitar el rango etario a personas mayores de 50 años?

La tercera opción de diseño se conoce como **equiparamiento**, que impone una "restricción parcial" sobre la variable de control que se equipara. Por ejemplo, si equiparamos con respecto al tabaquismo mediante lo que se denomina **formación de pares**, o emparejamiento, nuestra cohorte estaría compuesta por pares de individuos, donde cada par tendría un individuo de personalidad tipo A y otro de tipo B y donde ambos serían fumadores o no fumadores. Por lo tanto, la condición de tabaquismo no estaría restringida a solamente fumadores o solamente no fumadores. No obstante, la distribución de la condición de tabaquismo estaría restringida, y sería la misma para ambos grupos con personalidad tipo A y B. A menudo, el equiparamiento no es práctico en estudios de cohorte, como el estudio descrito aquí. Por lo consiguiente, el equiparamiento se usa en raras ocasiones en estudios de cohorte. Más bien se emplea con mayor frecuencia en los estudios de casos y controles o en los ensayos clínicos.

Preguntas de estudio (P13.3) (continuación)

Suponga que realiza un estudio de casos y controles para comparar los riesgos de EC que tienen los individuos de personalidad tipo A con aquellos de personalidad tipo B. Usted decide formar pares, tanto con respecto a la condición de tabaquismo ("alguna vez" versus "nunca") como del sexo. Sus casos son pacientes con EC identificados en un registro de enfermedad cardiovascular y sus controles son individuos sanos sin EC basados en la comunidad.

3. ¿Cuál es la condición de tabaquismo y el sexo de un individuo de control que se equipara con un caso masculino no fumador?
4. Si hay 110 casos en el estudio, ¿cuál es la cantidad total de individuos del estudio? (Suponga que se forman pares (equiparamiento), como se describió anteriormente).
5. ¿Existe alguna restricción sobre la condición de tabaquismo o el sexo de los casos?
6. ¿Existe alguna restricción sobre la condición de tabaquismo o el sexo de los controles?

Resumen

Las tres opciones de diseño para el control de variables externas son las siguientes:

- **Aleatorización**: asignación aleatorizada de individuos a grupos de comparación.
- **Restricción**: especificación de un rango acotado de valores posibles de una variable de control.
- **Equiparamiento**: restricción parcial de la distribución del grupo de comparación.

Opciones de análisis

¿Qué opciones de control están disponibles en la etapa de análisis de un estudio? Continuaremos con el ejemplo del estudio de cohorte a cinco años, cuyo objetivo es evaluar la posible asociación entre el tipo de personalidad y la enfermedad coronaria. Una vez que se recaban los datos, la opción de análisis más directa y lógica es un **análisis estratificado**.

En nuestro estudio de cohorte, si no hemos utilizado la restricción o el equiparamiento en la etapa de diseño, se puede llevar a cabo la **estratificación** mediante la categorización de todas las variables de control y la formación de combinaciones de categorías denominadas **estratos**. Por ejemplo, podemos clasificar de la siguiente manera: edad en tres grupos, digamos, menos de 50 años, 50 a 60 años y 60 años o más; raza, en no blancos versus blancos; presión arterial diastólica, < 95 mm Hg y $\geq$ 95 mm Hg; nivel de HDL $\leq$ 35 mg/dl versus > 35mg/dl. Como hay 6 variables que se clasifican en 3 categorías de edad y 2 categorías para las otras 5 variables, la cantidad total de combinaciones de categorías es 3 multiplicado por 2^5, es decir, 96 estratos.

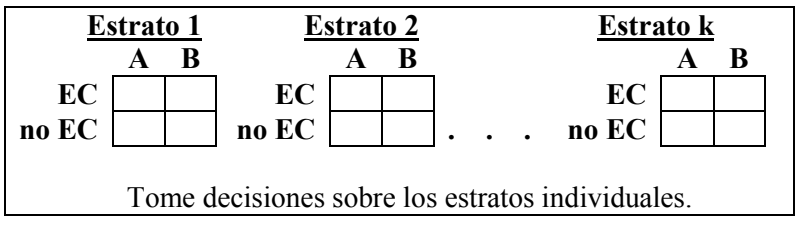
Control: EDAD, SEXO, RAZA, DBP, TAB, HDL

Categorización de todas las variables de control	Combinaciones de categorías
EDAD - menos de 50, 50 a 60 y 60 o más	6 Variables
RAZA - no blancos versus blancos	$3 \times 2^5 = 96$ (Estratos)
DBP - menos de 95 y 95 o más	
HDL - 35 o menos versus más de 35	
SEXO - Hombre, Mujer	
TAB - Sí, No	

Un ejemplo de estrato consiste en: mujeres de más de 60 años, no blancas, con presión arterial diastólica $\geq$ 95 mm Hg, fumadoras y con nivel de HDL $\leq$ 35 mg/dl. Para cada estrato formamos el cuadro de 2 x 2 que relaciona la exposición (el tipo de personalidad), con la enfermedad, en nuestro caso, el estatus de EC.

	A	**B**
EC		
no EC		

Luego, se realiza un análisis estratificado mediante la adopción de decisiones sobre la relación **E→D** para los estratos individuales y, si correspondiera, la combinación de la información de todos los estratos para brindar un cálculo ajustado general que controle todas las variables juntas.



Obtenga el cálculo ajustado general (si corresponde).

Preguntas de estudio (P13.4)

1. ¿Cuándo no sería apropiado sacar el cálculo ajustado general que combine la información de todos los estratos?
2. ¿Cómo evaluaría la relación E→D dentro de cualquier estrato dado?
3. Para el estudio de cohorte que se ejemplifica aquí, ¿cuál es el mayor obstáculo de realizar análisis específicos de cada estrato?
4. ¿Cómo podría realizar análisis estratificados que eviten manejar una cantidad grande de estratos que contengan celdas con ceros?
5. ¿Cuál sería la ventaja y la desventaja de hacer varios análisis estratificados con una variable por vez?

Uno de los problemas importantes de efectuar análisis estratificados cuando hay muchas variables para controlar es que la cantidad suficiente de individuos se agota con rapidez. **Una opción alternativa que soluciona este inconveniente es utilizar un *modelo* matemático**. El modelo matemático es una expresión o fórmula matemática que describe la manera en que una variable de resultado, como la condición de EC en nuestro ejemplo, se puede predecir a partir de otras variables (en nuestro supuesto, la variable de exposición y las variables de control que hemos medido u observado). En otras palabras, tenemos una variable que **predecir**, la cual generalmente se denomina **variable dependiente** y a la que normalmente se le denota con **Y**; además, hay **predictores**, que, a menudo, se llaman **variables independientes** y a los que comúnmente se les denota con **X**.

Modelo Matemático

Una fórmula que describe la manera en que una variable de resultado se puede predecir a partir de otras variables

$$\mathbf{Y} = \mathbf{f(X_1, X_2, X_3, ..., X_p)}$$

(variable dependiente) (variables independientes)

f denota "formula"

Cuando empleamos modelos matemáticos, no tenemos que dividir los datos en estratos. En su lugar, obtenemos una fórmula para predecir la variable dependiente a partir de las variables independientes. También podemos recurrir a la fórmula para obtener cálculos de medidas de efecto tales como las razones de riesgo o los *odds ratios*. Pero el uso de modelos matemáticos también plantea dificultades propias, entre las que se pueden mencionar la elección de la especificación que se usará, las variables que se incluirán en el modelo inicial y en el final, como también los supuestos que se requieren para realizar inferencias estadísticas.

Dificultades:

Elección del modelo
Las variables que se introducirán
Supuestos para realizar inferencias

Para las variables dependientes dicotómicas, como la EC en nuestro ejemplo, el modelo matemático más conocido se denomina **modelo logístico**.

Preguntas de estudio (P13.4) (continuación)

Suponga que queremos utilizar un modelo matemático para evaluar la relación entre el tipo de personalidad (**E**) y la condición de EC (**D**), controlando la edad, el sexo,la raza, la presión arterial diastólica (PAD), el tabaquismo (TAB) y la lipoproteína de alta densidad (HDL, por sus siglas en inglés).

6. Verdadero o falso. Las únicas alternativas posibles para las variables independientes en este modelo son E y las 6 variables de control anteriores.
7. Verdadero o falso. En un modelo matemático se deben categorizar las variables continuas.
8. Verdadero o falso. En un estudio de casos y controles, la variable dependiente es la condición de exposición.

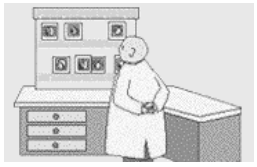
Suponga que $f(X_1, X_2, X_3, X_4, X_5, X_6, X_7)$ representa una fórmula matemática que brinda una buena predicción de la condición de EC, donde de X_1 a X_7 denotan E y las 6 variables de control descritas anteriormente.

9. Verdadero o falso. Si sustituimos los valores específicos correspondientes a una persona de X_1 a X_7 en la fórmula, determinaremos la correcta condición de EC de dicho individuo.

Resumen

En la etapa de análisis, hay dos opciones para el control:

- **Análisis estratificado**: categorización de las variables de control y formación de combinaciones de categorías o estratos.
- **Modelo matemático**: utilización de una expresión matemática para predecir el resultado a partir de la exposición y las variables que se controlan.
- El análisis estratificado tiene la desventaja de quedarse sin cantidades suficientes cuando el número de estratos es grande.
- El modelo matemático tiene sus desventajas, lo que incluye la elección del modelo y las variables que se introducirán en el modelo inicial y el final.



Auto-evaluación (P13.5)

Las tres razones principales para el control en un estudio epidemiológico, enumeradas en el orden en que deben evaluarse, son:

1. **???**
2. **???**
3. **???**

Opciones
confusión **modificación de efecto** **equiparamiento** **modelación matemática** **precisión** **aleatorización** **restricción** **estratificación**

Hay tres opciones para el control que pueden implementarse en la etapa de diseño.

4. **???** es una técnica para equilibrar la manera en que se distribuyen las variables no medidas entre los grupos de exposición.
5. **???** limita los individuos del estudio a un rango acotado o valores para una o más variables de control.
6. En un estudio de casos y controles, **???** garantiza que algunas o posiblemente todas las variables de control tengan la misma distribución o una similar entre los grupos de casos y controles.

Opciones
Equiparamiento **Modelación matemática** **Optimización** **Aleatorización** **Restricción** **Estratificación**

7. La estratificación es una técnica de análisis que comienza dividiendo a los individuos en diferentes **???** en función de las categorías de las variables **???**
8. Un problema importante con los análisis estratificados es tener demasiadas variables **???**, que pueden generar datos **???** en algunos estratos.

Opciones
de control **enfermedad** **exposición** **grandes cantidades** **muestras aleatorias** **escasos** **estratos** **tratamiento**

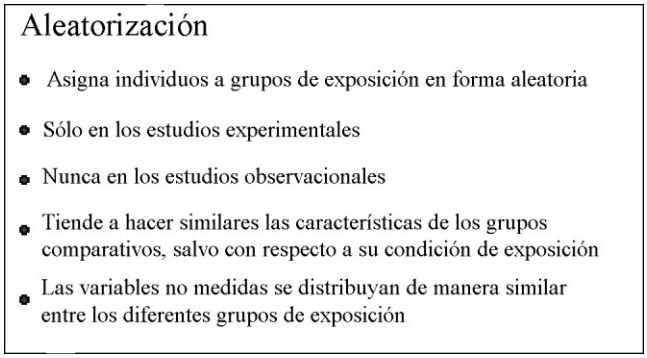
9. En la modelación matemática, usamos una fórmula para predecir una variable **???** a partir de una o más variables **???**
10. Los problemas de utilizar un modelo matemático incluyen la elección de la **???** del modelo, decidiendo qué **???** incluir en él y los **???** que se requieren para hacer inferencias estadísticas a partir del modelo.

Opciones
supuestos **casos** **complejidad** **de control** **dependientes** **forma** **independientes** **individuos** **tratamientos** **variables**

Aleatorización

La **aleatorización** asigna individuos a grupos de exposición en forma aleatoria. En la investigación epidemiológica, la aleatorización se emplea sólo en los **estudios experimentales** como, por ejemplo, los **ensayos clínicos** o **comunitarios**, y nunca se usa en los estudios observacionales.

¿Qué tiene que ver la aleatorización con el control de variables externas? El objetivo de la aleatorización es la **comparabilidad**. La aleatorización tiende a hacer similares las características demográficas, conductuales, genéticas, etc. de los grupos comparativos, **salvo** con respecto a su condición de exposición. Por consiguiente, el investigador espera que si el estudio arroja alguna diferencia en el resultado de salud entre los grupos comparativos, dicha diferencia sólo pueda atribuirse a su diferencia en estatus de exposición.



Por ejemplo, si a los individuos se les asigna aleatoriamente un medicamento nuevo o uno convencional para el tratamiento de la hipertensión, se espera que otros factores tales como edad y sexo, tengan aproximadamente la misma distribución en el caso de los individuos que reciben el medicamento nuevo y en el caso de los que toman el medicamento convencional.

En realidad, con la aleatorización, no existe ninguna garantía de que la distribución de edades, por ejemplo, sea equitativa en los dos grupos de tratamiento. El investigador siempre tiene la posibilidad de consultar los datos para ver qué sucedió en relación con una de estas características, siempre y cuando dicha particularidad se haya medido u observado en el estudio. Si, por ejemplo, se determina que la distribución de edades ha sido distinta entre los dos grupos de tratamiento, el investigador puede tenerlo en cuenta en el análisis mediante la estratificación por edad.

Una ventaja importante de la aleatorización es lo que ofrece en relación con las variables que **no** se miden en el estudio. Por supuesto, las variables que no se miden no pueden considerarse en el análisis. La aleatorización hace posible, aunque no garantiza, que estas variables no medidas se distribuyan de manera similar entre los diferentes grupos de exposición. Por otro lado, en los estudios observacionales, el investigador puede dar cuenta únicamente de las variables que se miden, en cuyo caso es más probable que se obtengan conclusiones falsas debido a los efectos desconocidos de importantes variables no medidas.

Preguntas de estudio (P13.6)

Suponga que planea realizar un estudio de casos y controles para evaluar si el tipo de personalidad es un factor de riesgo para el cáncer de colon.

1. ¿Puede asignar en forma aleatoria a los individuos del estudio a diferentes grupos de exposición?

Suponga que planea un ensayo clínico para comparar dos drogas antihipertensivas. Desea controlar la edad, la raza y el sexo, pero también quiere considerar los posibles factores genéticos que no puede medir.

2. ¿Puede controlar factores genéticos específicos en su análisis?
3. ¿Los dos grupos de drogas tendrán las mismas distribuciones de edad, raza y sexo?
4. ¿Qué espera que logre la aleatorización en relación con los factores genéticos que no ha medido?

Resumen

- Los **estudios experimentales** utilizan la **aleatorización**, mientras que los estudios observacionales **no** la emplean.
- El objetivo de la aleatorización es la comparabilidad.
- La aleatorización tiende a hacer que los grupos comparativos sean similares con respecto a otros factores que se van a controlar.
- Una ventaja importante de la aleatorización es que tiende a hacer que las variables no medidas se distribuyan en forma similar entre los grupos comparativos.
- No se garantiza que la aleatorización haga que los grupos comparativos sean similares, automáticamente, en relación con otros factores.

13-2 Control de factores externos (continuación)

Restricción

La **restricción** es otra opción de diseño para el control en la cual se acota la elegibilidad de los potenciales individuos del estudio, restringiendo las categorías de una o más variables de control. La restricción puede aplicarse tanto a las variables continuas como categóricas. Para una variable categórica, como el sexo la restricción significa simplemente que el estudio se limita a una o más categorías. Para una variable continua, la restricción requiere limitar el rango de valores, como, por ejemplo, usar un rango de edad acotado, digamos de 40 a 50 años.

La restricción normalmente brinda un control completo de una variable. Es práctica, poco costosa y requiere un análisis simple para lograr el control. Por ejemplo, si un estudio se limita sólo a las mujeres, el análisis no precisa la obtención de un efecto ajustado que promedie ambos sexos. La principal desventaja de la restricción es que no podemos generalizar nuestros resultados más allá de la categoría restringida. Para las variables continuas, otra desventaja es que el rango de valores que se restringe puede no ser lo suficientemente acotado y, por lo tanto, quizás aún exista la confusión de que haya que controlar dentro del rango seleccionado.

Restricción (Opción de diseño)
acota la elegibilidad

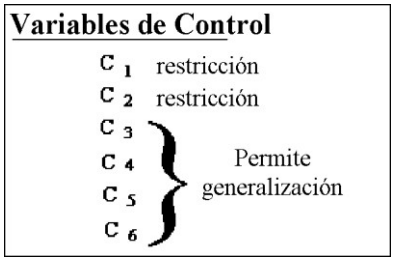
Continuas Requiere limitar el rango de valores

Categóricas Se limita a una o más categorías

Ventajas	Desventajas
Control completo	No se puede generalizar
Práctica	No es suficientemente
Poco costosa	acotado
Análisis simple	(variables continuas)

Dada una lista de diversas variables de control por medir, normalmente utilizamos la restricción en una cantidad pequeña de variables. Esto permite que se evalúen hipótesis sobre varias categorías de la mayoría de las

variables de control, permitiendo, de esta forma, mayores posibilidades de generalización de los resultados.

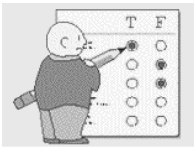


Por ejemplo, si queremos controlar la edad, el sexo la raza, la presión arterial diastólica, la condición de tabaquismo y el nivel de HDL, probablemente usaríamos la restricción para no más de dos o tres de estas variables, por ejemplo, la edad y el sexo.

La restricción puede emplearse en la etapa de análisis, incluso si no se utiliza en la etapa de diseño. Por ejemplo, aunque la muestra de estudio puede incluir varios grupos étnicos, podemos decidir solamente analizar los datos para uno de estos grupos, especialmente si otros grupos étnicos tienen relativamente pocos individuos. Sin embargo, es más ventajoso optar por la restricción en la etapa de diseño para ganar precisión en el efecto calculado o para reducir los costos del estudio. Por ejemplo, para un tamaño o costo fijo del estudio, la restricción de la raza a afroamericanos en la etapa de diseño ofrecerá más individuos afroamericanos y, por lo tanto, más precisión en las medidas de efecto para este grupo, que lo que se obtendría si el diseño permitiera que se eligieran varios grupos étnicos.

Resumen

- La restricción es una opción de diseño que acota la elegibilidad de los potenciales individuos del estudio, restringiendo las categorías de una o más variables de control.
- La restricción puede aplicarse tanto a las variables categóricas como a las continuas.
- La restricción brinda normalmente un control completo, es práctica, poco costosa y requiere un análisis simple para lograr el control.
- La principal desventaja de la restricción es que no permite generalizar los resultados más allá de la categoría restringida.
- Para las variables continuas, otra desventaja es la posibilidad de confusión residual dentro del rango de valores restringidos.
- La restricción puede emplearse en la etapa de análisis, pero si se utiliza en la etapa de diseño, puede lograrse precisión o se pueden reducir los costos del estudio.



Auto-evaluación (P13.7)

Coloque **verdadero** o **falso** en cada una de las siguientes afirmaciones:

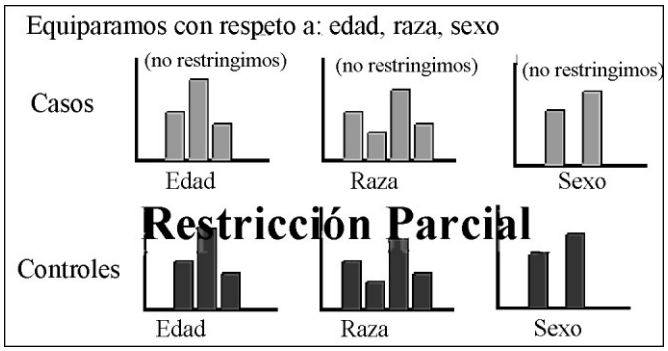
1. La restricción sólo puede emplearse con variables categóricas. . . . . **???**
2. Una ventaja de la restricción es que sólo requiere un análisis simple para lograr el control. **???**
3. Una desventaja de la restricción es que puede ser costosa de implementar. . . **???**
4. La restricción puede utilizarse tanto en la etapa de diseño como en la de análisis. . **???**

Equiparamiento

El **equiparamiento** es una opción de diseño a la que se puede recurrir en los **estudios experimentales** y en los **estudios de cohorte observacionales**, pero se utiliza más ampliamente en los **estudios de casos y controles**. En la lección 15 se incluye una definición general que permite el desarrollo de otros diseños.

En general, existen dos tipos de equiparamiento: **equiparamiento individual** y **equiparamiento por frecuencias**. Cuando se usa el equiparamiento individual en un estudio de casos y controles, se eligen uno o más controles para cada caso de forma que los controles tengan las mismas particularidades o características similares con respecto a cada una de las variables involucradas en el equiparamiento. Por ejemplo, si equiparamos con respecto a la edad, la raza y el sexo, y determinado caso tiene 40 años, es de raza negra y de sexo masculino, entonces el control o los controles que se equiparan con este caso también deben tener 40 años o aproximadamente esta edad, ser de raza negra y de sexo masculino. Para las variables continuas, como la edad, las categorías que se emplean para el equiparamiento se deben especificar antes de proceder a dicho proceso. Para la edad, por ejemplo, si se establece que las categorías de equiparamiento son bandas de 10 años, que incluyen el rango de edad de 35-45, entonces el equiparamiento de un control para un caso de 40 años debe provenir del rango de edad de 35-45 años.

Aquí, restringimos la distribución de la edad, la raza y el sexo en el grupo de los controles para que sea igual al grupo de los casos. Pero **no** estamos restringiendo los valores o la distribución de la edad, la raza y el sexo para los **casos**. Ésta es la razón por la cual decimos que el equiparamiento impone una **restricción parcial** sobre las variables de control que se están equiparando.



Si se recurre al **equiparamiento por frecuencias** en un estudio de casos y controles, entonces el equiparamiento se realiza en forma grupal, en lugar de individual. Por ejemplo, suponga que deseamos equiparar por frecuencias con respecto a la raza y el sexo en un estudio de casos y controles, donde los casos se desglosan según la raza y el sexo tal como se muestra a continuación:

	Hombre blanco	Mujer blanca	Hombre negro	Mujer negra
Casos	100 (33%)	100 (33%)	40 (13%)	60 (20%)

Entonces, los controles se deben elegir como un grupo que posea la misma distribución que tienen los casos en los cuatro estratos de raza por sexo. Si queremos tener el doble de controles que de casos, entonces el desglose de raza y sexo para los controles seguirá el mismo patrón de distribución que presenten los casos.

Cuando se realiza el equiparamiento, se deben tener en cuenta varias cuestiones.

✓ En primer lugar, ¿cuales son las **ventajas** y **desventajas** del equiparamiento? Una razón importante para

el equiparamiento es ganar eficiencia o precisión en el cálculo de los efectos, digamos, el *odds ratio*.

- ✓ ¿**Deberíamos equiparar** con respecto a todo, o deberíamos elegir los casos y controles sin equiparar? No hay una respuesta simple a esta pregunta, pero una pauta general es sólo equiparar con respecto a las variables que crea que pueden constituir factores fuertes de confusión en sus datos.
- ✓ ¿**Cuántos controles** deberíamos elegir para cada caso? Una regla general aquí indica que, comúnmente, no se necesitarán más de cuatro controles por cada caso para poder ganar precisión.
- ✓ ¿Cómo **analizamos los datos equiparados**? La respuesta depende en parte de si hay o no otras variables que se deban controlar en el análisis, además de aquellas que se equiparan. En especial, si las únicas variables que se controlan están involucradas en el equiparamiento, entonces el análisis adecuado es un tipo especial de **análisis estratificado**. Pero si además de las variables que se equiparan, hay otras que se tienen que controlar, entonces el análisis adecuado involucra la **modelación matemática**, normalmente mediante métodos de **regresión logística**.

Equiparamiento - Problemas

¿Cuales son las Ventajas y Desventajas
Ganar eficiencia o precisión

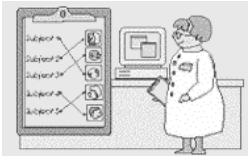
¿Deberíamos equiparar con respecto a todo?
Equiparar con respecto a las variables que pueden constituir factores de confusión fuertes

¿Cuántos controles deberíamos elegir para cada caso?
Cuatro controles por cada caso

¿Cómo analizamos los datos equiparados?
No hay otras variables: Análisis estratificado
Otras variables: Modelación matemática

Resumen

- ❖ El equiparamiento puede utilizarse tanto en los estudios experimentales como observacionales y se emplea más comúnmente en los estudios de casos y controles.
- ❖ Existen dos tipos de equiparamiento: **equiparamiento individual** y **equiparamiento por frecuencias**.
- ❖ Una razón importante para el equiparamiento es ganar eficiencia o precisión en el cálculo de efecto.
- ❖ Por lo general, no se requerirán más de cuatro controles por cada caso para poder ganar precisión.
- ❖ Si las únicas variables que se controlan están involucradas en el equiparamiento, entonces utilice el análisis estratificado.
- ❖ Si hay otras variables por controlar, entonces use la modelación matemática.



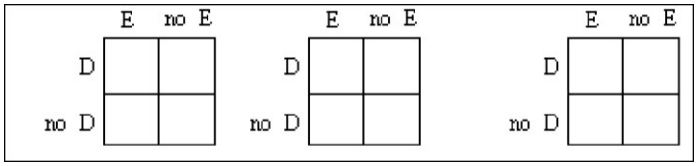
Auto-evaluación (P13.8)

Coloque **verdadero** o **falso** en cada una de las siguientes afirmaciones:

1. El equiparamiento se emplea principalmente con estudios de cohorte observacionales. . **???**
2. Para las variables continuas, los rangos utilizados para crear categorías de equiparamiento se deben especificar antes del proceso de equiparamiento. . . . . **???**
3. Para el equiparamiento por frecuencias, puede haber más controles que casos. . . **???**
4. El análisis estratificado se usa normalmente para datos equiparados a fin de controlar variables distintas de las involucradas en el equiparamiento. . . . . **???**

Análisis estratificado

El **análisis estratificado** es una opción de análisis para el control que comprende la categorización de todas las variables del estudio y la formación de combinaciones de categorías denominadas **estratos**. Si tanto las variables de exposición como de enfermedad son dicotómicas, entonces los estratos tienen la forma de varios cuadros de 2 x 2. La cantidad de estratos dependerá de cuántas variables se van a controlar y cuántas categorías se definen para cada variable.



Preguntas de estudio (P13.9)

1. Si se van a controlar tres variables y se dicotomiza cada una de ellas, ¿cuántos estratos se obtienen?
2. Si las variables de control son la edad, la raza y el sexo, dé un ejemplo de uno de los estratos.

Una vez que se definen los estratos, se realiza un análisis estratificado mediante análisis simples **específicos de cada estrato** y, si correspondiera, mediante una evaluación de resumen general de la relación **E→D** que represente todas las variables de control en forma simultánea. Tanto los análisis específicos de cada estrato como los análisis de resumen generales normalmente involucrarán el cálculo y la interpretación de una estimación puntual de efecto, por ejemplo, una razón de riesgos, un intervalo de confianza para la estimación puntual y una prueba de hipótesis para la significación de la estimación puntual.

Para una **evaluación de resumen** general, la estimación puntual es un cálculo ajustado que comprende alguna forma de promedio ponderado de los cálculos específicos para cada estrato. El intervalo de confianza es una estimación del intervalo alrededor de este promedio ponderado, y la prueba de hipótesis es una generalización de la prueba de chi cuadrado de Mantel-Haenszel que ahora considera varios estratos.

Resumen

El análisis estatificado involucra los siguientes pasos:

- Categorización de todas las variables.
- Formación de combinaciones de categorías (es decir, estratos).
- Realización de análisis específicos para cada estrato.
- Elaboración de una evaluación general de la relación E→D, si correspondiera.
- Tanto los análisis específicos para cada estrato como la evaluación general requieren estimaciones puntuales y de intervalos, como también una prueba de hipótesis.

En lo que respecta a la evaluación general:

- La estimación puntual es un cálculo ajustado que consiste en un promedio ponderado de los cálculos específicos para cada estrato.
- El intervalo de confianza es una estimación del intervalo alrededor del cálculo ajustado (ponderado).
- La prueba de hipótesis es una generalización de la prueba de chi cuadrado de Mantel-Haenszel.

Auto-evaluación (P13.10)



Coloque **verdadero** o **falso** en cada una de las siguientes afirmaciones:

1. El análisis estratificado es una opción de análisis para el control que comprende la categorización de todas las variables del estudio y la formación de combinaciones de categorías denominadas estratos. . . . . ???
2. Las pruebas de hipótesis no son adecuadas para los análisis estratificados. . . ???
3. Cuando se realizan análisis específicos para cada estrato, la estimación puntual se calcula normalmente como un promedio ponderado. . . . . ???
4. Cuando se efectúan análisis específicos para cada estrato, un estadístico de prueba adecuado para muestras grandes es el estadístico chi cuadrado de Mantel-Haenszel. . . . ???
5. Si corresponde realizar una evaluación general de todos los estratos, el estadístico de prueba adecuada para muestras grandes es el estadístico chi cuadrado de Mantel-Haenszel. . . ???
6. Cuando se hace una evaluación general de todos los estratos, el estadístico chi cuadrado de Mantel-Haenszel siempre es adecuada para muestras grandes. . . . . ???

13-3 Control de factores externos (continuación)

Modelación matemática

Introducción a la modelación matemática

El **modelo matemático** es una expresión o fórmula matemática que describe cómo puede predecirse una **variable de resultado** a partir de **variables explicativas** que afectan el resultado. A continuación se muestra una expresión general para un modelo matemático. En esta fórmula, **Y** denota la **variable de resultado**, a menudo denominada **variable dependiente**. Las variables **X** de la fórmula denotan los **predictores**, llamadas **variables independientes**, mientras que **f** denota una fórmula o función matemática que involucra las variables **X** de alguna manera. Finalmente, el término **e** denota el error que representa la variabilidad no explicada por las variables **X** en la función **f**.

Modelo Matemático:

$$Y = f(X_1, X_2, \ldots, X_p) + e$$

variable dependiente | variables independientes | **error**

En epidemiología, independientemente de la forma matemática de la función, la variable **Y** normalmente es la **variable del resultado de salud** o la **variable de enfermedad**, que comúnmente llamamos **D**. Al menos, una de las variables **X** es una **variable de exposición**, que hemos denominado **E**. Otras variables **X** pueden ser variables de control, que hemos denotado con la letra **C**. Asimismo, otras variables **X** pueden ser funciones de las variables **C**, como edad^2 si la edad es una variable **C** continua, o **términos de productos** como **E x edad**. Frecuentemente, para evaluar la **interacción** se consideran términos de productos que involucran la variable de exposición y una o más variables de control. El término del producto **E x edad**, por ejemplo, considera la interacción de la exposición con la edad, donde la edad es vista como un posible modificador de efecto de la

exposición.

Epidemiología: Y resultado de salud (D)
X_1 = Exposición (E)
X_2, X_3 variables de control (C)
$X_4 = EDAD^2$
X_5 = E x EDAD (interacción)

Preguntas de estudio (P13.11)

Suponga que queremos recurrir a la modelación matemática para evaluar la relación entre el tipo de personalidad (TP) y la condición de EC (Y), controlando el sexo, la raza y el tabaquismo (TAB), siendo todas variables (0,1).

1. Escriba una expresión general (es decir, una afirmación en términos de **f**) para un modelo matemático que contenga estas variables.

La función **f** puede adoptar una variedad de formas. Una de las más conocidas se llama **función lineal**, y el modelo correspondiente se denomina **modelo lineal**. Como ejemplo, suponga que las **X** en el modelo son las variables 0, 1 del tipo de personalidad, denotada como TP, el sexo, la raza y la condición de tabaquismo, denotada como TAB. Entonces, la función **f** para un modelo lineal que involucra estas variables podría tener la forma que se muestra a continuación:

Ejemplo: (0,1) variables
X_1 = TP X_3 = Raza
X_2 = Sexo X_4 = FUM
Lineal f (TP, Sexo , Raza, FUM)
$= b_0 + b_1(TP) + b_2(SEXO) + b_3(RAZA) + b_4(FUM)$

En esta función, las cantidades de $\mathbf{b_0}$ a $\mathbf{b_4}$ son parámetros desconocidos denominados **coeficientes de regresión** que deben calcularse a partir de los datos del estudio. Una forma más general de una función lineal se muestra a continuación, donde hemos reemplazado las cuatro variables en el ejemplo anterior por varias **X**.

Función lineal general:
$f(X_1, X_2, ..., X_p) = b_0 + b_1X_1 + b_2X_2 + ... + b_pX_p$

Las **X** en un modelo lineal pueden ser factores **categóricos** o **cuantitativos**. El modelo todavía se considera un modelo lineal porque es una suma que involucra los coeficientes de regresión.

Preguntas de estudio (P13.11) (continuación)

X_1 = TP X_3 = Raza
X_2 = Sexo X_4 = FUM
Función lineal general:
$f(X_1, X_2, ..., X_p) = b_0 + b_1X_1 + b_2X_2 + ... + b_pX_p$

Suponga que las **b** en el modelo lineal que involucra al TP, el sexo la raza y el TAB son, respectivamente:

$b_0 = 0{,}10$; $b_1 = 0{,}20$; $b_2 = 0{,}25$; $b_3 = 0{,}10$ y $b_4 = 0{,}15$

2. ¿Cuál es el valor predicho para EC (es decir, Y) para una persona con valores de variables TP = 1, sexo = 1, raza = 1 y TAB = 1?
3. ¿Cuál es el valor predicho para EC para una persona con valores de variables TP = 0, sexo = 1, raza = 1 y TAB = 1?

Suponga que además de las variables TP, sexo, raza y TAB, se agregan los siguientes términos de productos al modelo:

TP x sexo, TP x raza, TP x TAB

4. Escriba la especificación del modelo lineal que incluye estas variables nuevas, además de las cuatro variables originales.

Suponga en el nuevo modelo que las **b** son:

$b_0 = 0{,}08$; $b_1 = 0{,}18$; $b_2 = 0{,}35$; $b_3 = 0{,}07$; $b_4 = 0{,}12$; $b_5 = 0{,}03$; $b_6 = 0{,}02$
y $b_7 = 0{,}03$.

5. ¿Cuál es el valor predicho para EC (es decir, Y) para una persona con valores de variables TP = 1, sexo = 1, raza = 1 y TAB = 1?
6. ¿Cuál es el valor predicho para EC para una persona con valores de variables TP = 0, sexo = 1, raza = 1 y TAB = 1?

Resumen

- El modelo matemático es una expresión o fórmula matemática que describe cómo puede predecirse una variable de resultado a partir de variables explicativas.
- Una expresión general para un modelo matemático es aquella donde **Y** es la **variable dependiente**, las **X** son las **variables independientes**, **f** es una función matemática y **e** es el término de error.

  $\mathbf{Y = f(X_1, X_2,\ldots, X_p) + e}$
- En epidemiología, la **Y** normalmente es el **resultado de salud** y las **X** pueden incluir las variables de exposición, las variables de control y variables más complicadas, como los términos de productos.
- Una función muy conocida es la función lineal, que generalmente adopta la forma $\mathbf{f(X_1, X_2, \ldots, X_p) = b_0 + b_1X_1 + b_2X_2 + \ldots + b_pX_p}$.
- Dicho tipo de función es lineal en los coeficientes de regresión, aunque puede incluir términos de productos como las **X**.

Valor esperado

Otra manera de escribir el modelo matemático es en términos del **valor esperado** de **Y**, dados los valores de las **X**. Este es el valor promedio de **Y** para todas las personas de nuestra población de estudio con los mismos valores de **X** para cada **X** en el modelo:

$$\underbrace{\text{Exp}(Y \mid X_1, X_2, \ldots, X_p)}_{\text{valor esperado}} = f(X_1, X_2, \ldots, X_p)$$

(Nota: en esta lección, el valor esperado se puede representar como **Exp(Y|…)** o **E(Y|…)**).

Por ejemplo, si las **X** son edad, raza y sexo, entonces el valor esperado de **Y** para una persona que tiene 30 años, es de raza negra y de sexo femenino es la **Y** promedio para todas las personas en la población de estudio que tengan 30 años, sean de raza negra y de sexo femenino. De forma similar, el valor esperado para una persona de 50 años, de raza blanca y de sexo masculino es la **Y** promedio para todas las personas que tienen 50 años, son de raza blanca y de sexo masculino.

El valor esperado de **Y** dadas las **X** es igual al valor de la función **f** (sin término de error) porque el valor esperado del término de error siempre es igual a cero. Es decir, se supone que el error promedio para todos los individuos en la población de estudio es cero.

Cuando calculamos el modelo matemático empleando los datos de nuestro estudio, escribimos el modelo estimado como se muestra a continuación:

$$\text{Modelo estimado (modelo adaptado)} \quad \underbrace{\hat{Y}}_{\text{valor predicho de Y}} = \hat{f}(X_1, X_2, \ldots, X_p)$$

La $\hat{Y}$ en el lado izquierdo de esta fórmula denota el valor predicho de **Y**, que surge como resultado de especificar los valores para las **X** en el modelo estimado para un individuo específico. El modelo estimado se denota como $\hat{f}$. A menudo nos referimos a éste como el **modelo adaptado**, ya que es el que obtenemos cuando ajustamos el modelo a los datos de nuestro estudio.

Podemos predecir una **Y** mediante un modelo lineal que involucra las **X**. Las fórmulas para el valor esperado y el modelo adaptado tendrán la siguiente forma:

Modelo Lineal

$$\text{Exp}(Y|X_1, X_2, \ldots, X_p) = b_0 + b_1(X_1) + b_2(X_2) + \ldots + b_p(X_p)$$

$$\hat{Y} = \hat{b}_0 + \hat{b}_1(X_1) + \hat{b}_2(X_2) + \ldots + \hat{b}_p(X_p)$$

Hemos colocado "acentos circunflejos" para los coeficientes de regresión en el modelo adaptado, ya que los coeficientes representan aquí los cálculos que se obtuvieron al **adaptar el modelo** a los datos del estudio.

Preguntas de estudio (P13.12)

Suponga que deseamos emplear un modelo lineal para predecir la presión arterial sistólica (PAS) de una persona a partir de su edad (como variable continua) y la condición de tabaquismo (TAB = 1 si fumó alguna vez y TAB = 0 si nunca fumó).

1. Escriba un modelo lineal simple que involucre los predictores anteriores mediante la forma de valor esperado del modelo.

Suponga que adapta el modelo lineal a los datos de su estudio mediante las variables edad, TAB y edad x TAB como los predictores.

2. Escriba la especificación del modelo predicho (es decir, estimado) que involucra estas variables.

Suponga que adapta el siguiente modelo lineal, $\text{Exp(PAS|edad, TAB)} = b_0 + b_1(\text{edad}) + b_2(\text{TAB})$ y que sus coeficientes de regresión estimados son $\hat{b}_0 = 50$, $\hat{b}_1 = 2$, $\hat{b}_2 = 10$.

3. ¿Cuál es el valor predicho de PAS para un fumador de 30 años?
4. ¿Cuál es el valor predicho de PAS para un no fumador de 30 años?

Resumen

- Una forma alternativa de escribir el modelo matemático es en términos del **valor esperado** de **Y**, dados los valores de **X**.
- El valor esperado de **Y** dadas las **X** puede expresarse como:
  $$\mathbf{Exp(Y|X_1, X_2, \ldots, X_p) = f(Y|X_1, X_2, \ldots, X_p)}$$
- El valor esperado del término de error es siempre cero.
- El valor predicho de **Y** es la expresión: $\hat{Y} = \hat{f}(X_1, X_2, \ldots, X_p)$.
- El valor esperado para un modelo lineal puede escribirse como:
  $$\mathbf{Exp(Y|X_1, X_2, \ldots, X_p) = b_0 + b_1X_1 + b_2X_2 + \ldots + b_pX_p}$$
- El valor predicho de **Y** para un modelo lineal es: $\hat{Y} = \hat{b}_0 + \hat{b}_1X_1 + \hat{b}_2X_2 + \ldots + \hat{b}_pX_p$, donde las $\hat{b}$ denotan los coeficientes de regresión estimados que se obtuvieron al adaptar el modelo a los datos del estudio.

El modelo logístico

En epidemiología, muchos estudios consideran variables de resultado de salud que son dicotómicas. Tal vez busquemos determinar si una persona desarrolla o no determinada enfermedad, vive o muere, tiene presión arterial alta o normal, se enferma o no a raíz de un brote, etc. La **variable dependiente**, **Y**, que se analiza en dichos estudios toma, por lo tanto, uno de dos valores posibles, generalmente definidos de manera práctica como 1 ó 0.

Cuando la variable dependiente es dicotómica, el modelo matemático más conocido es no lineal y se denomina **modelo logístico**. Aquí mostramos la formulación matemática general, tanto del **modelo lineal** como del **modelo logístico**:

Modelo Lineal

$$E(Y|\ X_1,X_2,\ldots,X_p) = b_0+b_1X_1+b_2X_2+\ldots+b_pX_p$$

Modelo Logístico - no lineal

$$E(Y|\ X_1,X_2,\ldots,X_p) = \frac{1}{1+e^{-(b_0+b_1X_1+b_2X_2+\ldots+b_pX_p)}}$$

Estos dos modelos lucen bastante diferentes. En particular, el modelo logístico es no lineal, porque no está formulado como la suma lineal de los coeficientes de regresión. Sin embargo, hay un componente lineal en el denominador del modelo logístico, que le otorga un significado especial para los epidemiólogos, como explicaremos más adelante. La **e** en este modelo es la **base de los logaritmos naturales**.

En general, dados valores específicos de las **b** y las **X**, estos dos modelos darán valores predichos diferentes.

Preguntas de estudio (P13.13)

Suponga que Y = EC (1 = sí, 0 = no), X_1 = sexo (1 = femenino, 0 = masculino) y X_2 = TAB (1 = fumador, 0 = no fumador). Presuma también que $b_0 = 0{,}10$; $b_1 = 0{,}25$ y $b_2 = 0{,}15$.

1. Calcule y compare los valores predichos mediante el modelo lineal y el modelo logístico para una mujer fumadora. (Aquí debería utilizar una calculadora, pero ayudará saber que exp(-0,50) = 0,6065).

El modelo logístico tiene varias características que son importantes para los epidemiólogos.

Modelo Logístico

- Cuando la variable dependiente Y toma los valores 1 ó 0, el valor esperado se simplifica a una probabilidad
- El modelo logístico siempre arrojará un valor en el rango entre 0 y 1
- El componente lineal en el modelo logístico puede expresarse a través de una transformación simple

En primer lugar, cuando la variable dependiente **Y** toma los valores 1 ó 0, ya sea que el resultado de salud se presente o no, la expresión del valor esperado se simplifica a una declaración de probabilidad. Esto significa que el uso del modelo logístico describe la probabilidad o el riesgo de que una persona desarrolle el resultado de salud como una función de variables predictoras de interés. Los predictores pueden incluir una variable de

exposición, variables de control y términos de productos de exposición con variables de control. Por ende, el modelo logístico permite en forma conveniente que se modele el riesgo a partir de datos epidemiológicos.

Otra característica del modelo logístico es que la función no lineal que define el modelo tiene la propiedad de que, más allá de qué valores numéricos se especifiquen para los coeficientes de regresión o las **X**, la función logística siempre arrojará un valor en el rango entre 0 y 1. Esta propiedad asegura que el valor predicho obtenido para cualquier individuo esté limitado al rango de valores para una probabilidad (es decir, $0 \leq Pr \leq 1$). Por el contrario, el modelo lineal puede arrojar valores predichos por debajo de cero o por arriba de 1 para valores especificados de los predictores.

En tercer lugar, el componente lineal en el modelo logístico puede expresarse a través de una transformación simple de la expresión de probabilidad correspondiente al modelo. En especial, por principios algebraicos, el logaritmo natural de la probabilidad dividido por 1 menos la probabilidad (es decir, **ln[P/(1 – P)]**) da este componente lineal.

$$\ln\left[\frac{P}{1-P}\right] = b_0 + b_1X_1 + b_2X_2 + \ldots + b_pX_p$$

$$P = \Pr(Y=1|X_1, X_2, \ldots, X_p)$$

Preguntas de estudio (P13.13) (continuación)

Hagamos una introducción y revisión rápida de los números exponenciales y los logaritmos naturales (es decir, los últimos normalmente se denotan con ln).
Aquí se incluyen algunas reglas:

e = 2,7183
ln(e) = 1,0000
exp(z) = e elevado a la potencia de z (es decir, e^z)
ln[exp(z)] = z
exp(z + w) = e^{z+w} = exp(z)exp(w)
ln(z * w) = ln(z) + ln(w)
ln(z/w) = ln(z) – ln(w)

2. Exprese ln(P/(1-P)] como una diferencia entre dos logaritmos
3. Resuelva ln(A) = Z (en términos de una expresión exponencial)
4. Resuelva ln[P/(1-P)] = Z para P/(1 – P)
5. Resuelva ln[P/(1 – P)] = Z para P (éste es difícil).
6. Si ln(OR) = 1, ¿cuál es el OR?
7. Si ln(OR) = 0, ¿cuál es el OR?

Esta transformación (es decir, **ln[P/(1 – P)]**) se denomina transformación logit y es importante para los epidemiólogos porque la cantidad **P/(1-P)** denota un *odds*. En consecuencia, la función logit nos permite expresar el logaritmo de los *odds*:

ln[P/(1 – P)] = ln(*Odds*)

y, por lo tanto, los *odds* para un individuo con valores X específicos como una función lineal de las **X**. Si comparamos los *odds* para un individuo dividido entre los *odds* para otro sujeto, obtenemos un *odds ratio*, que es la medida fundamental de efecto que puede calcularse a partir del modelo logístico.

$$\frac{\frac{P_1}{1 - P_1}}{\frac{P_2}{1 - P_2}} = \text{Odds Ratio}$$

Preguntas de estudio (P13.13) (continuación)

Transformación Logit

$$\ln \frac{P}{1-P} = b_0 + b_1X_1 + b_2X_2 + \ldots + b_pX_p$$

Considere un modelo logístico para el cual Y = EC (1 = sí, 0 = no), X_1 = sexo (1 = femenino, 0 = masculino) y X_2 = TAB (1 = fumador, 0 = no fumador). Suponga que para este modelo, los coeficientes de regresión son $b_0 = 0{,}10$; $b_1 = 0{,}25$ y $b_2 = 0{,}15$.

8. ¿Cuáles son los *odds* para una mujer fumadora?
9. ¿ Cuáles son los *odds* para un hombre fumador?
10. ¿Cuál es el *odds ratio* que compara a una mujer fumadora con un hombre fumador?
11. ¿Cuál es el *odds ratio* que compara a una mujer no fumadora con un hombre no fumador?

Resumen

- Cuando **Y** es dicotómica, el modelo matemático más conocido es no lineal y se denomina **modelo logístico**.
- El valor esperado del modelo logístico es:

$$\mathbf{E(Y|X_1,\ldots,X_p)} = \frac{1}{1 + e^{-(b_0 + b_1X_1 + b_2X_2 + \ldots + b_pX_p)}}$$

- Para Y(0, 1), el valor esperado se simplifica a $\mathbf{Pr(Y = 1|X_1,\ldots,X_p)}$, que describe el riesgo de desarrollar el resultado de salud.
- La función logística siempre arrojará valores predichos entre 0 y 1.
- Mediante la transformación logit, **ln[P/(1-P)]**, el modelo logístico puede calcular los *odds* para un individuo y el *odds ratio* que compara a dos individuos.

Función de riesgo: ejemplo

Generalmente a los pacientes les preocupa su prognosis. Por ejemplo, un hombre al que se le diagnosticó diabetes mellitus le pedirá a su médico que le dé algún indicio del riesgo de sufrir un ataque cardíaco en los próximos diez años. El médico basará su predicción del riesgo no sólo en el hecho de que este paciente es diabético, sino también en la presencia o ausencia de otros varios factores de riesgo cardiovascular.

Preguntas de estudio (P13.14)

1. ¿Qué información específica (factor de riesgo) será importante para que el doctor suministre una predicción precisa del riesgo que tiene el paciente de sufrir un ataque cardíaco en los próximos 10 años?

El médico debe combinar toda esta información para predecir el riesgo de enfermedad cardiovascular para este paciente en particular. Una forma de hacer esta predicción implica utilizar una **función de riesgo**. Las funciones de riesgo pueden calcularse mediante varios modelos matemáticos. Uno de ellos es el **modelo logístico**. Nuevamente, la fórmula general para este modelo es la siguiente:

$$\text{Riesgo de enfermedad} = \frac{1}{1 + e^{-(b_0 + b_1x_1 + b_2x_2 + \ldots + b_px_p)}}$$

Las **b** en la fórmula se calculan a partir de los datos. Las **x** son los valores de varios factores de riesgo cardiovascular.

En el siguiente ejemplo, el riesgo de morir en 10 años puede calcularse mediante los valores de las **b** de varios factores de riesgo que se dan en el cuadro que se muestra a continuación:

Nombres de las Variables	
	$b_0 = -9.194$
Edad años	$b_1 = 0.104$
Sexo ; hombre = 1; mujer = 0	$b_2 = 0.658$
Indice de masa corporal kg/m^2	$b_3 = -0.041$
Presión arterial sistólica mmHg	$b_4 = 0.007$
Niveles de colesterol mg/100ml	$b_5 = 0.001$
Diabetes mellitus sí = 1; no = 0	$b_6 = 1.251$
Fumar cigarrillos por día	$b_7 = 0.023$
Frecuencia cardíaca latidos/min	$b_8 = 0.0053$

Preguntas de estudio (P13.14) (continuación)

Usando una calculadora con la fórmula: $\dfrac{1}{1 + e^{-(b_0 + b_1x_1 + b_2x_2 + \ldots + b_px_p)}}$

2. Calcule el riesgo de morir en 10 años para un hombre no diabético de 62 años que fuma 20 cigarrillos por día, tiene un índice de masa corporal de 25 kg/m^2, presión arterial de 140 mmHg, nivel de colesterol de 245 mg/100 mL y frecuencia cardíaca de 80 latidos/min.
3. Calcule el riesgo de morir en 10 años para un hombre no diabético de 62 años que no fuma, tiene un índice de masa corporal de 25 kg/m^2, presión arterial de 140 mmHg, nivel de colesterol de 245 mg/100 mL y frecuencia cardíaca de 80 latidos/min.

Las respuestas a las dos preguntas anteriores fueron 26,8% para el riesgo de un hombre no diabético que fuma 20 cigarrillos por día y 18,8% para un hombre no diabético que no fuma, manteniendo valores idénticos de otros 5 factores de riesgo en ambos casos.

4. ¿Cuál es la razón de riesgos que compara a estos dos individuos? ¿Qué indica esto sobre el efecto del tabaquismo en el riesgo de sufrir un ataque cardíaco?

5. ¿Qué dicen los cálculos de riesgo anteriores sobre la razón de riesgos que compara a un hombre diabético con otro no diabético, si se controlan los demás factores de riesgo de la lista?
6. ¿Cómo podría utilizar el modelo adaptado para comparar a un individuo diabético con uno no diabético?

Resumen

- La función de riesgo puede utilizarse para determinar el riesgo de un individuo de desarrollar una enfermedad específica.
- Las funciones de riesgo pueden calcularse mediante varios modelos matemáticos como, por ejemplo, el modelo de regresión logística. La fórmula para este modelo es:

$$\text{riesgo de enfermedad} = \frac{1}{1 + e^{-(b_0 + b_1X_1 + b_2X_2 + \ldots + b_pX_p)}}$$

Características básicas del modelo logístico

La siguiente es una lista de algunas propiedades importantes del **modelo logístico**:

1. **Fórmula general:**

$$\Pr(D = 1 \mid X_1, ..., X_p) = \frac{1}{1 + e^{-(b_0 + b_1X_1 + b_2X_2 + ... + b_pX_p)}}$$

2. Una manera equivalente de escribir este modelo es la forma **logit**, que se simplifica a la siguiente función lineal:

$$\text{logit } P(X) = \ln \frac{P(X)}{1 - P(X)} = b_0 + b_1X_1 + b_2X_2 + ... + b_pX_p$$

donde los parámetros $b_0, b_1, b_2, ...b_p$ en este modelo representan coeficientes de regresión desconocidos que se deben calcular. (Nota: generalmente se utilizan símbolos griegos, en lugar de latinos, para denotar los coeficientes de regresión, por ejemplo, β_i en vez de $\mathbf{b_i}$, pero en esta descripción seguiremos con las **b**).

3. El método de cálculo usual se denomina **estimación de máxima verosimilitud** (**MV**); los estimadores de estos coeficientes de regresión se llaman **estimadores de máxima verosimilitud** (**EMV**) y normalmente se escriben con acentos circunflejos sobre los parámetros, de la siguiente manera: $\hat{b}_0, \hat{b}_1, \hat{b}_2, ...\hat{b}_p$

4. Si $\mathbf{X_l}$ es una variable (0,1) y <u>ninguna</u> de las otras **X** son **términos de productos** de la forma $\mathbf{X_1 \times W_j}$, entonces la fórmula del ***odds ratio*** **ajustado** para el efecto de $\mathbf{X_l}$, controlando las otras **X**, es:

$$\hat{OR}_{\text{ajustado}} = \exp(\hat{b}_1)$$

donde $\mathbf{b_1}$ es el coeficiente de $\mathbf{X_1}$ en el modelo.

5. Si $\mathbf{X_l}$ es una variable (0,1) y <u>algunas</u> de las otras **X** son términos de productos de la forma $\mathbf{X_1 \times W_j}$, entonces la fórmula del ***odds ratio*** **ajustado** para el efecto de $\mathbf{X_l}$, controlando las otras X, es:

$$\hat{OR}_{\text{ajustado}} = e^{[\hat{b}_1 + \sum \hat{b}_j W_j]}$$

donde $\mathbf{b_1}$ es el coeficiente de $\mathbf{X_1 \times W_j}$ en el modelo.

6. Si $\mathbf{X_l}$ es una variable (0,1) y ninguna de las otras **X** son **términos de productos** de la forma $\mathbf{X_1 \times W_j}$, entonces un **intervalo de confianza de 95%** para el ***odds ratio*** **ajustado** correspondiente al efecto de $\mathbf{X_l}$, controlando las otras **X**, es la fórmula:

$$\exp(\hat{b}_1 \pm 1{,}96 \times s_{\hat{b}_1})$$

donde $\mathbf{b_1}$ es el coeficiente de $\mathbf{X_1}$ y

$$s_{\hat{b}_1} = \sqrt{\hat{Var}(\hat{b}_1)}$$

es el **error estándar** del valor de $\mathbf{b_1}$.

7. La **prueba de Wald** es una forma de prueba de la hipótesis nula de que un *odds ratio* ajustado basado en un modelo logístico es igual al valor nulo de 1. Si, nuevamente, suponemos que nuestra variable de exposición de interés es la variable (0,1) $\mathbf{X_1}$ y que no hay términos de productos en el modelo, dos formas alternativas de expresar la hipótesis nula son:

$$H_0 : b_1 = 0 \text{ ó } H_0 : OR = e^{b_1} = 1$$

donde $\mathbf{b_1}$ es el coeficiente de $\mathbf{X_1}$. El estadístico de la prueba de Wald es un estadístico chi cuadrado cuya fórmula es:

$$Z^2 = \left(\frac{\hat{b}_1}{s_{\hat{b}_1}} \right)^2 = \frac{\hat{b}_1^{\,2}}{\hat{Var}(\hat{b}_1)}$$

Este estadístico tiene la **distribución chi cuadrado** con 1 gl según la hipótesis nula.

8. Un procedimiento de prueba alternativo, que se considera que posee mejores propiedades estadísticas que la prueba de Wald, se denomina prueba de la **razón de verosimilitudes** (**RV**). Aunque estas dos pruebas pueden arrojar distintos resultados computacionales, ambas son pruebas de muestra grande y, en la mayoría de los casos, ofrecerán conclusiones similares. Sin embargo, en caso de duda, utilice la prueba de **RV**.

Ejemplo:

Ahora ejemplificaremos varias de estas características (excepto los procedimientos de prueba) del modelo logístico para un estudio que considera el efecto que tienen los antecedentes de tabaquismo ($\mathbf{X_1}$ = **TAB**, 1 si alguna vez fumó, 0 si nunca fumó) en el desarrollo de enfermedades coronarias (**D = EC**, 1 equivale a "sí", 0 equivale a "no") en una cohorte de 609 hombres del condado Evans, Georgia, a quienes se les hizo el seguimiento desde 1960 hasta 1967. Se deben controlar dos factores de riesgo conocidos, $\mathbf{X_2}$ **= EDAD** (continua) y $\mathbf{X_3}$ **= SEXO** (1 indica "femenino", 0 indica "masculino"). El siguiente es un modelo logístico sin interacción para estos datos:

$$\text{logit } P(X) = b_0 + b_1 TAB + b_2 EDAD + b_3 SEXO$$

El anterior se denomina **modelo sin interacción** porque no contiene ningún predictor de la forma $\mathbf{X_1}$ **x** $\mathbf{W_j}$, donde $\mathbf{W_j}$ es un modificador potencial de efecto. Entonces, la fórmula para el ***odds ratio*** **ajustado** para el efecto del **TAB**, ajustado por

EDAD y **SEXO**, es igual a:

$$\hat{OR}_{\text{ajustado}} = \exp(\hat{b}_1)$$

donde $\mathbf{b_1}$ es el coeficiente de la variable de exposición $\mathbf{X_1}$ = **TAB**. (Nota: los coeficientes de las $\mathbf{X_2}$ y $\mathbf{X_3}$, que no

involucran la variable de exposición $\mathbf{X_1}$ = **TAB**, no aparecen en la fórmula del **OR**).

Un modelo logístico alternativo para estos datos que permite la posibilidad de **interacción** de la variable de exposición **TAB** con ambas variables de control **EDAD** y **SEXO** es:

$$\text{logit } P(X) = b_0 + b_1 TAB + b_2 EDAD + b_3 SEXO + b_4 TAB \times EDAD + b_5 TAB \times SEXO$$

donde

$$X_1 = TAB, X_2 = EDAD, X_3 = SEXO, X_4 = TAB \times EDAD, X_5 = TAB \times SEXO$$

Para este modelo con interacción, hay dos **W**, es decir $\mathbf{W_1}$ = EDAD y $\mathbf{W_2}$ = SEXO Por lo tanto, la fórmula para el *odds ratio* ajustado es:

$$\hat{OR}_{ajustado} = \exp(\hat{b}_1 + \hat{b}_4 EDAD + \hat{b}_5 SEXO)$$

de forma que el valor del *odds ratio* "con interacción" cambiará (es decir, se modificará) en función de los valores de los modificadores potenciales de efecto **EDAD** y SEXO.

Como ejemplo numérico (hipotético), suponga que los **EVM** para $\mathbf{b_1}$, $\mathbf{b_4}$ y $\mathbf{b_5}$ son:

$$\hat{b}_1 = 0{,}25; \hat{b}_4 = 0{,}01; \hat{b}_5 = 0{,}03$$

Entonces, el cálculo del *odds ratio* ajustado toma la siguiente forma:

$$\hat{OR}_{ajustado} = \exp(0{,}25 + 0{,}01\ EDAD + 0{,}03\ SEXO)$$

Por lo tanto, si **EDAD = 40** y **SEXO = 1** (femenino), el *odds ratio* calculado es:

$$\hat{OR}_{ajustada} = \exp[0{,}25 + 0{,}01(40) + 0{,}03(1)] = 1{,}97$$

mientras que si **EDAD = 60** y **SEXO = 1**, el *odds ratio* calculado es (un valor diferente):

$$\hat{OR}_{ajustada} = \exp[0{,}25 + 0{,}01(60) + 0{,}03(1)] = 2{,}41$$

Referencia: Kleinbaum DG y Klein M, *Logistic Regression - A Self-learning Text*, 2da edición, Springer-Verlag Publishers, 2002.

El modelo logístico EVW

Un caso especial importante del modelo logístico que tiene relevancia especial para el análisis de los datos epidemiológicos es la versión **EVW** del modelo, cuya fórmula es la siguiente:

$$\Pr(D=1 \mid X_1, ..., X_p) = P(X) = \frac{1}{1+e^{-(\alpha+\beta E+\sum_{i=1}^{p1}\gamma_i V_i+\sum_{j=1}^{p2}\delta_j EW_j)}}$$

Una manera equivalente de escribir este modelo es la forma logit, definida como **logit P(X)**, que se obtiene tomando el logaritmo natural de **P(X)/[1 - P(X)]**. Al aplicar principios algebraicos, esta forma logit se simplifica a la siguiente función lineal:

$$\text{logit } P(\mathbf{X}) = \alpha + \beta E + \sum_{i=1}^{p1}\gamma_i V_i + \sum_{j=1}^{p2}\delta_j EW_j$$

Las **α**, **γ**, **β** y **δ** en este modelo (previamente denotadas con **b** en la actividad sobre modelación logística correspondiente a este suplemento/recuadro) representan coeficientes de regresión desconocidos que se deben calcular. El método de cálculo usual se denomina **estimación de máxima verosimilitud** (**MV**); los estimadores de estos coeficientes de regresión se llaman **estimadores de máxima verosimilitud** (**EMV**) y normalmente se escriben con acentos circunflejos sobre los parámetros, de la siguiente manera: $\hat{\alpha}, \hat{\beta}, \hat{\gamma}$ y $\hat{\delta}$.

Las **E**, **V** y **EW** representan las variables predictoras (es decir, las X) en este modelo. Más específicamente, la variable predictora **E** denota una única variable de exposición de interés, las **V** denotan las variables de control (es decir, los **factores potenciales de confusión**) y las **W** denotan los **modificadores** potenciales **de efecto** que entran en el modelo como términos de productos con la variable de exposición **E**. (Nota: nos referimos a las **V** y **W** como factores **potenciales** de confusión y modificadores **potenciales** de efecto, porque cualquiera de estas variables se puede eliminar finalmente del modelo cuando se analizan los datos del estudio).

La fórmula general para el ***odds ratio*** **ajustado** para el efecto de una variable de exposición (0,1) **E** **dicotómica**, ajustada para las variables **V** y **W** que se están controlando como factores potenciales de confusión y modificadores potenciales de efecto, respectivamente, es la siguiente expresión:

$$\hat{OR}_{ajustado} = \exp(\hat{\beta} + \sum_{j=1}^{p_2}\hat{\delta}_j W_j)$$

donde $\hat{\beta}$ y las $\hat{\delta}$ son los valores de los **EVM** de los coeficientes de sólo aquellas variables del modelo que involucran la variable de exposición **E** (nota: los coeficientes de las variables **V**, que no involucran **E** , no aparecen en la fórmula anterior del **OR**).

Si el modelo logístico que se considera no contiene ningún modificador potencial de efecto (es decir, ninguna W en el modelo), entonces la expresión anterior del *odds ratio* ajustado (nuevamente suponiendo una variable de exposición (0,1)) se simplifica a:

$$\hat{OR}_{ajustado} = \exp(\hat{\beta})$$

que sólo involucra **β**, el coeficiente de la variable de exposición **E**.

El modelo logístico EVW (continuación)

Ahora ejemplificaremos el modelo logístico **EVW** para un estudio que considera el efecto que tienen los antecedentes del tabaquismo (**E = TAB**, 1 si alguna vez fumó, 0 si nunca fumó) en el desarrollo de enfermedad coronaria (**D = EC**, 1 equivale a "sí", 0 equivale a "no") en una cohorte de 609 hombres del condado Evans, Georgia, a quienes se les hizo el seguimiento desde 1960 hasta 1967. Se deben controlar dos factores de riesgo conocidos, **EDAD** (continua) y **SEXO** (1 indica "femenino", 0 indica "masculino").

El siguiente es un modelo logístico **sin interacción** para estos datos basado en la fórmula general anterior de **EVW**:

$$\text{logit } P(X) = \alpha + \beta E + \gamma_1 V_1 + \gamma_2 V_2$$

En este ejemplo, **E = TAB**, $\mathbf{V_1}$ **= EDAD** y $\mathbf{V_2}$ **= SEXO**. Se denomina **modelo sin interacción** porque no contiene ningún predictor de la forma **EW**, donde **W** es un modificador potencial de efecto. Entonces, la fórmula del *odds ratio* ajustado para el efecto del **TAB**, ajustado por **EDAD** y **SEXO** es igual a:

$$\hat{OR}_{\text{ajustado}} = \exp(\hat{\beta})$$

donde **β** es el coeficiente de la variable de exposición **E = TAB**. (Observe nuevamente que los coeficientes de las variables **V**, que no involucran **E**, no aparecen en la fórmula del **OR**).

Un modelo logístico alternativo para estos datos que permite la posibilidad de interacción de la variable de exposición **TAB** con ambas variables de control **EDAD** y **SEXO** es:

$$\text{logit } P(X) = \alpha + \beta E + \gamma_1 V_1 + \gamma_2 V_2 + \delta_1 E W_1 + \delta_2 E W_2$$

donde **E = TAB**, $\mathbf{V_1}$ **= EDAD =** $\mathbf{W_1}$ y $\mathbf{V_2}$ **= SEXO =** $\mathbf{W_2}$. Para este modelo con interacción, la fórmula del *odds ratio* ajustado es:

de forma que el valor del *odds ratio* "con interacción" cambiará (es decir, se modificará) en función de los valores de los modificadores potenciales de efecto **EDAD** y **SEXO**

Como ejemplo numérico (hipotético), suponga que los **EVM** para β, δ_1 y δ_2 son:

$$\hat{\beta}_1 = 0{,}25, \hat{\delta}_1 = 0{,}01, \hat{\delta}_2 = 0{,}03$$

Entonces, el cálculo del *odds ratio* ajustado adopta la siguiente forma:

$$\hat{OR}_{\text{ajustado}} = \exp(0{,}25 + 0{,}01\,\text{EDAD} + 0{,}03\,\text{SEXO})$$

Por lo tanto, si **EDAD = 40** y **SEXO = 1** (femenino), el *odds ratio* calculado es:

$$\hat{OR}_{\text{ajustado}} = \exp[0{,}25 + 0{,}01(40) + 0{,}03(1)] = 1{,}97$$

mientras que si **EDAD = 60** y **SEXO = 1**, el *odds ratio* calculado es (un valor diferente):

$$\hat{OR}_{ajustado} = \exp[0{,}25 + 0{,}01(60) + 0{,}03(1)] = 2{,}41$$

Referencia: Kleinbaum DG y Klein M, *Logistic Regression - A Self-learning Text*, 2da edición, Springer-Verlag Publishers, 2002.



Auto-evaluación (P13.15)

Verdadero o **falso**; las siguientes expresiones pueden incluirse como variable X en el modelo matemático:

1. Variables de control. . . . . **???**
2. Variables de enfermedad. . . . **???**
3. Variables dependientes. . . . . **???**
4. Términos de productos de dos o más variables X. **???**
5. Funciones de variables de control. . . **???**

Complete los espacios vacíos

6. Para variables dependientes dicotómicas, la forma más conocida de modelo matemático es el modelo **???**
7. El modelo matemático incluye **???**, que se calculan a partir de los datos para sacar un valor predicho para la variable Y.

Opciones
curvo **lineal** **logístico** **probabilidades** **coeficientes de regresión**

Se ha construido un modelo lineal a partir de un conjunto de datos de estudio. La fórmula es la siguiente:

$$Y = 0.15 + 0.10_x \text{TP} + 0.2_x \text{SEXO} + 0.15_x \text{Raza} + 0.25_x \text{FUM}$$

8. ¿Cuál es el valor predicho de EC para los siguientes valores de la variable X? TP = 1, SEXO = 0, RAZA = 0 y TAB = 0. **???**
9. ¿Cuál es el valor predicho de EC con los siguientes valores de la variable X? TP = 1, SEXO = 1, RAZA = 0 y TAB = 0. **???**
10. Si se emplease un modelo logístico que involucrara los mismos datos y las mismas variables, en lugar de un modelo lineal, las respuestas a los ítems a) y b) serían probablemente **???**

Opciones
0,15 **0,25** **0,35** **0,45** **0,50** **diferentes** **idénticas**

11. La forma del valor esperado de un modelo matemático brinda el valor **???** de Y para todas las personas en la población de estudio que poseen los mismos valores para las X correspondientes en el modelo.
12. El modelo de valor esperado no incluye un término de error porque se supone que el error promedio para todos los individuos de la población de estudio es **???**

Opciones
1 **promedio** **más grande** **más pequeño** **cero**

13. El modelo logístico predice la probabilidad de que una persona desarrolle el **???** que resulta de interés sobre la base de los valores correspondientes a esa persona para las variables **???** y de control que se utilizan en el modelo.
14. La transformación **???** permite que el modelo logístico brinde información sobre un *odds ratio*.

Opciones
arcoseno **dependiente** **valor esperado** **de exposición** **resultado de salud** **lineal** **logit** **raíz cuadrada**

Nomenclatura

b_x	En los modelos matemáticos, los coeficientes de regresión, normalmente numerados b_o-b_p.
C	Variable de control o covariable.
D	Variable de enfermedad o de resultado.
E	Variable de exposición, o en los modelos matemáticos, el valor esperado.
X_x	En los modelos matemáticos, las variables independientes o predictoras, comúnmente numeradas X_1-X_p.
Y	En los modelos matemáticos, la variable dependiente o de resultado.

Modelo lineal general

$$E(Y \mid X_1, X_2, ..., X_p) = b_o + b_1 X_1 + b_2 X_2 + ... + b_p X_p$$

Modelo logístico general (no lineal)

$$E(Y \mid X_1, X_2, ..., X_p) = \frac{1}{1 + e^{-(b_o + b_1 X_1 + b_2 X_2 + ... + b_p X_p)}}$$

Transformación logit

$$\ln \frac{P}{1 - P} = b_o + b_1 X_1 + b_2 X_2 + ... + b_p X_p$$

Referencias

Feinstein AR. *Clinical Biostatistics*. Mosby Publishers, St. Louis, 1977.

Hoes AW, Grobbee DE, Valkenburg HA, Lubsen J, Hofman A. *Cardiovascular risk and all-cause mortality: a 12 year follow-up study in the Netherlands*. Eur J EpidemioI 1993; 9(3):285-92.

Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic Research: Principles and Quantitative Methods*. John Wiley and Sons Publishers, Nueva York, 1982.

Kleinbaum DG, Klein M. *Logistic Regression: A Self-Learning Text*, 2ª edición, Springer Verlag Publishers, 2002.

Miettinen OS. *Matching and design efficiency in retrospective studies*. Am J Epidemiol 1970; 91(2):111-8.

Ejercicios

EPI-1. Control de variables externas

a. ¿Qué significa decir que "estamos controlando variables externas"?
b. Describa tres razones por las cuales "controlamos variables externas" y explique el orden en el cual deben considerarse estas tres causas cuando se analizan los datos.
c. Exprese, al menos, una ventaja y una desventaja de las siguientes opciones de control:

    Aleatorización
    Restricción
    Equiparamiento
    Análisis estratificado
    Modelación matemática

d. ¿Cuál es la razón principal de por qué la modelación matemática se utiliza en lugar del análisis estratificado o en forma suplementaria a éste?
e. Suponga Y = EC, X1 = EDAD, X2 = TAB y X3 = EDAD*TAB. Formule el modelo logístico en términos de estas variables. ¿Qué parte de este modelo es un "componente lineal"?
f. Indique la forma logit del modelo logístico descrito en el ítem e).
g. ¿Por qué la forma logit del modelo tiene especial importancia para los epidemiólogos?

EPI-2. Modelo logístico: calcule un *odds ratio*

Una regla importante relacionada con el modelo logístico indica lo siguiente:

"Si las variables independientes en el modelo son E, C_1, …, C_p, donde E es una variable de exposición (0,1) y **ninguna** de las C son términos de productos de la forma E x C (por ejemplo, E x EDAD), entonces el *odds ratio* ajustado para el efecto de E en el resultado D, controlando C_1, …, C_p, es la expresión:

$$OR(E, D \mid C_1, \ldots, C_p) = \mathbf{exp}(b)$$

donde b es el coeficiente de E en el modelo".

a. Considere el modelo logístico cuya expresión es

$$P(EC = 1 \mid TAB, PESO, HPT) = \frac{1}{1 + \exp\left[-(b_0 + b_1 TAB + b_2 PESO + b_3 HPT)\right]}$$

donde
TAB = 1 si fumó alguna vez y TAB = 0 si nunca fumó,
PESO = 1 si pesa mucho y PESO = 0 si el peso es normal,
HPT =1 si es hipertenso y HPT = 0 si es normotenso.

Para este modelo, ¿cuál es la **fórmula** del *odds ratio* para el efecto de TAB en EC, controlando PESO y HPT?

b. Suponga que al adaptar el modelo anterior a los datos del estudio, los valores de los coeficientes de regresión son:

$$\hat{b}_0 = -6.80\ ,\ \hat{b}_1 = 0.84\ ,\ \hat{b}_2 = 0.37\quad \hat{b}_3 = .44$$

¿Cuál es su **cálculo** del *odds ratio* para el efecto de TAB en EC, controlando PESO y HPT?

c. Mediante el mismo conjunto de valores dados en el ítem b, ¿cuál es su cálculo del *odds ratio* para el efecto de HPT en EC, controlando TAB y PESO?
d. ¿Cómo corregiría el modelo logístico que se definió en el ítem a) si deseara calcular el *odds ratio* para el efecto de TAB en EC controlando sólo HPT? ¿Cuál sería la fórmula del *odds ratio* ajustado en este modelo corregido? ¿Esperaría obtener la misma respuesta que obtuvo en el ítem b)?

EPI-3. Modelo logístico: calcule un intervalo de confianza

Suponga, como en la pregunta 2, que las variables independientes en el modelo son E, C_1, …, C_p, donde E es una variable de exposición (0,1) y **ninguna** de las C son términos de productos de la forma E x C (por ejemplo, E x EDAD). Entonces, un **intervalo de confianza del 95% de muestra grande** para el *odds ratio* ajustado para el efecto de E en el resultado D, controlando C_1, …, C_p, es la expresión:

$$\mathbf{exp}\ [\ \hat{b} \pm 1.96\ s_{\hat{b}}\ ]$$

donde $\hat{b}$ es el coeficiente de E en el modelo y $s_{\hat{b}}$ es el error estándar calculado de $\hat{b}$.

Considere, como en la pregunta 2, el modelo logístico que involucra variables predictoras (0,1) cuya expresión es:

$$P(EC = 1 \mid TAB, PESO, HPT) = \frac{1}{1 + \exp\left[-(b_0 + b_1 TAB + b_2 PESO + b_3 HPT)\right]}$$

A continuación se incluye el resultado obtenido por computadora al adaptar este modelo:

Variable	Coeficiente	Error estándar	Estadístico de Wald	Valor p
Constante	- 6,7727	1,1402	35,3	0,0000
TAB	0,8347	0,3052	7,5	0,0062
PESO	0,3695	0,2936	1,6	0,2083
HPT	0,4393	0,2908	2,3	0,1309

Calcule un intervalo de confianza del 95% para OR(TAB, EC | PESO, HPT).

EPI-4. Estadístico de Wald

El estadístico de Wald es una prueba de muestra grande de la hipótesis nula que indica que el coeficiente de una variable dada en un modelo logístico es igual a cero, es decir, H_0: b = 0. Ya que **exp**(b) brinda un *odds ratio*, surge que una hipótesis nula equivalente es H_0: OR = 1. La fórmula del estadístico de Wald es:

$$\mathbf{Z}^2 = \left(\frac{\hat{b}}{s_{\hat{b}}}\right)^2$$

y este estadístico tiene la distribución chi cuadrado con la hipótesis nula.

Mediante el resultado obtenido por computadora que se brindó en la pregunta 3, realice pruebas de Wald **separadas** para la significación de TAB, PESO y HPT, respectivamente. (Al responder a esta pregunta, formule

la hipótesis nula en términos de un *odds ratio* ajustado, exprese el valor del estadístico de Wald, su valor p y derive una conclusión sobre la significación estadística).

EPI-5. Modelo logístico: interacción

Otra regla importante relacionada con el modelo logístico indica lo siguiente:

"Si las variables independientes en el modelo son E, C_1, …, C_p, donde E es una variable de exposición (0,1) y **alguna** de las C son términos de productos de la forma E x W (por ejemplo, E x EDAD), entonces el *odds ratio* ajustado para el efecto de E en el resultado D, controlando C_1, …, C_p, es la expresión:

$$OR(E, D \mid C_1, \ldots, C_p) = \mathbf{exp}(b + d_1W_1 + d_2W_2 + \ldots + d_qW_q)$$

donde b es el coeficiente de E en el modelo, W_1 , W_2 , … , W_q son variables q que están en el modelo en la forma C = E x W y las d_i son los coeficientes de los términos de productos $C_i = E_i$ x W_i , respectivamente.

a. Considere el modelo logístico de la siguiente expresión:

$$P(EC = 1 \mid TAB, PESO, HPT) = \frac{1}{1 + \exp\left[-(b_0 + b_1TAB + b_2PESO + b_3HPT + d_1(TAB \times PESO) + d_2(TAB \times HPT))\right]}$$

donde TAB, PESO y HPT son variables (0,1), como se definió en la pregunta 2.

Para este modelo, ¿cuál es la **fórmula** del *odds ratio* para el efecto de TAB en EC, controlando PESO y HPT? Para obtener un valor numérico para este OR, ¿qué debe especificar además de los valores de ciertos coeficientes de regresión? ¿Qué fórmula tiene sentido con respecto al significado de la interacción/modificación de efecto?

Respuestas a las preguntas de estudio y auto-evaluaciones

P13.1

1. 3 x 2 = 6 combinaciones
2. El enfoque 2 controla la edad y el tabaquismo, ya que considera lo que sucede cuando se controlan variables. El enfoque 1 ignora el control de la edad y el tabaquismo.

P13.2

1. Falso. La afirmación aborda la cuestión de la interacción/modificación de efecto, no la confusión.
2. Falso. La precisión obtenida dependerá de los datos; no se garantiza que siempre se gane precisión controlando variables externas.
3. Confusión, ya que la afirmación se refiere a lo que sucede cuando comparamos un cálculo crudo de efecto con un cálculo ajustado de efecto.
4. Si la edad y el tabaquismo son factores de riesgo de desarrollar cáncer de pulmón (de hecho, lo son), entonces el OR de 3,5 es más apropiado porque controla la confusión, a pesar de ser menos preciso que el cálculo crudo; es decir, la validez es más importante que la precisión.
5. No, los OR de 5,7 para los fumadores y 1,4 para los no fumadores indican una interacción fuerte a raíz del tabaquismo (siempre que la interacción observada sea estadísticamente significativa). La evaluación de la confusión requeriría comparar un cálculo crudo con el cálculo ajustado, pero el uso del último no sería adecuado, ya que ocultaría la presencia de una interacción fuerte.

P13.3

1. No puede generalizar sus resultados para los hombres; es decir, proceder de esa manera plantea un problema de "validez externa".
2. El grupo etario de las personas mayores de 50 años no es necesariamente lo suficientemente acotado como para controlar la edad por completo. En particular, todavía puede haber confusión "residual" debido a la edad dentro del grupo etario mayor a 50 años.
3. El control tendrá la misma condición de tabaquismo y sexo que el caso, es decir, el control será un hombre no fumador.
4. 220, ya que habrá 110 casos y 110 controles correspondientes.
5. No, los casos pueden ser hombres o mujeres y fumadores o no fumadores; además, pueden tener cualquiera de las distribuciones posibles de cada una de estas variables.
6. Sí. Los controles están limitados a tener la misma distribución de tabaquismo y sexo que los casos.

P13.4

1. No sería adecuado sacar un cálculo ajustado general si hay evidencia contundente de interacción, por ejemplo, si las razones de riesgos calculadas en dos o más estratos son ambas estadística y significativamente diferentes.
2. Hay que realizar un análisis simple para el estrato dado calculando una estimación puntual de efecto (por ejemplo, una razón de riesgos), un intervalo de confianza alrededor de la estimación puntual, como también una prueba de hipótesis sobre la significación de este cálculo.
3. Existen 96 estratos en total, por lo que es muy probable que todo el conjunto de datos se reduzca considerablemente con la estratificación, incluso muchos estratos que contienen una o más celdas con cero.
4. Se debe llevar a cabo un análisis estratificado con una o dos variables por vez, en lugar de controlar simultáneamente todas las variables.
5. Una ventaja es que se puede llegar a comprender en forma preliminar la confusión y la interacción para cada variable de control. Una desventaja es que no se controlarán todas las variables en forma simultánea.
6. Falso. Si la edad es continua, entonces también podría utilizarse edad^2. Algo similar ocurriría para otras variables continuas. Asimismo, podrían emplearse términos de producto como E x edad, edad x sexo
7. Falso. Las variables continuas se pueden tomar como variables continuas o categóricas, según la opinión del investigador. Sin embargo, para

un análisis estratificado, las variables continuas deben ser categóricas.
8. Falso. Aunque los casos y los controles se seleccionen primero y luego se determine la exposición previa, la condición de caso-control es la variable dependiente en un modelo matemático, porque representa la variable de resultado de salud que se predice.
9. Falso. Los modelos matemáticos raramente predicen en forma perfecta la variable de resultado, si es que alguna vez lo hacen. Siempre hay algún grado de error que representa la diferencia entre el valor predicho y el observado.

P13.5

1. Modificación de efecto
2. Confusión
3. Precisión
4. Aleatorización
5. Restricción
6. Equiparamiento
7. Estratos, de control
8. De control, escasos
9. Dependientes, independientes
10. Forma, variables, supuestos

P13.6

1. No. En un estudio de casos y controles, la condición de exposición se determina sólo después de que se seleccionan los casos y los controles. Por lo tanto, la aleatorización de los grupos de exposición (es decir, tipos de personalidad) no es posible en los estudios de casos y controles.
2. No. No puede controlar los factores que no ha medido en su análisis.
3. No necesariamente. La aleatorización tendería a tornar la distribución de la edad, la raza y el sexo similar en los dos grupos de drogas, pero no se garantiza que serán iguales.
4. Se espera que la distribución de factores genéticos sea similar dentro de los dos grupos de drogas, aunque no se han medido estos factores. Asimismo, se espera que, para cualquier otro factor no medido, la aleatorización distribuya dichos factores en forma similar en todos los grupos que se consideren.

P13.7

1. Falso. La restricción también puede emplearse para limitar el rango de valores de una variable continua.
2. Verdadero. Una ventaja de la restricción es que es poco costosa de implementar.
3. Falso
4. Verdadero

P13.8

1. Falso. La equiparación se usa principalmente con estudios de casos y controles.
2. Verdadero
3. Verdadero
4. Falso. Si además de las variables que se equiparan, hay otras que se deben controlar, entonces el análisis adecuado involucra la modelación matemática mediante métodos de regresión logística.

P13.9

1. 2 x 2 x 2 = 8 estratos
2. Un estrato contendría a todos los individuos del estudio que están en un grupo etario categorizado y tienen la misma raza y sexo (por ejemplo, mujeres blancas de 30-40 años).

P13.10

1. Verdadero
2. Falso. Tanto el estadístico específico para cada estrato como el de resumen general normalmente implicarán el cálculo de una estimación puntual, un intervalo de confianza y una prueba de hipótesis.
3. Falso. Cuando se realizan análisis específicos para cada estrato, la estimación puntual es un valor puntual simple calculado para un estrato específico.
4. Verdadero
5. Verdadero
6. Falso. La prueba de Mantel-Haenszel no es adecuada si hay una interacción significativa importante entre los estratos.

P13.11

1. EC = f(TP, sexo raza, TAB) + e
2. 0,10 + 0,20(1) + 0,25(1) + 0,10(1) + 0,15(1) = 0,80
3. 0,10 + 0,20(0) + 0,25(1) + 0,10(1) + 0,15(1) = 0,60
4. f(TP, sexo raza, TAB) = $b_0 + b_1$(TP) + b_2(sexo) + b_3(raza) + b_4(TAB) + b_5(TP x sexo) + b_6(TP x raza) + b_7(TP x TAB)
5. 0,08 + 0,18(1) + 0,35(1) + 0,07(1) + 0,12(1) + 0,03(1x1) + 0,02(1x1) + 0,03(1x1) = 0,88
6. 0,08 + 0,18(0) + 0,35(1) + 0,07(1) + 0,12(1) + 0,03(1x0) + 0,02(1x0) + 0,03(1x0) = 0,62

P13.12

1. Exp(PAS|edad, TAB) = b_0 + b_1(edad) + b_2(TAB).
   Otros modelos posibles incluyen:
   Exp(PAS|edad, TAB) = b_0 + b_1(edad) + b_2(TAB) + b_3(edad x TAB)
   Exp(PAS|edad, TAB) = b_0 + b_1(edad) + b_2(TAB) + b_3(edad x TAB) + b_4(edad^2)
2. $\hat{PAS} = \hat{b}_0 + \hat{b}_1(\text{edad}) + \hat{b}_2(\text{TAB}) + \hat{b}_3(\text{edad} \times \text{TAB})$
3. 50 + 2(30) + 10(1) = 120
4. 50 + 2(30) + 10(0) = 110

P13.13

1. Modelo lineal:
   $\hat{Y} = 0{,}10 + 0{,}24(1) + 0{,}15(1) = 0{,}50$
   Modelo logístico:
   $$\hat{Y} = \frac{1}{1 + e^{-(0{,}10 + 0{,}25(1) + 0{,}15(1))}} 0{,}6225$$
   Los dos valores predichos son bastante diferentes.
2. ln[P/(1-P)] = ln(P) – ln(1 – P)
3. A = exp(Z)
4. P/(1 - P) = exp(Z)
5. P = 1/[1 + exp{-Z}], una función logística. Si Z = $b_0 + b_1Z_1 + \ldots + b_pX_p$, entonces P = modelo logístico.
6. OR = 2,7183 = e
7. OR = 1
8. ln[P/(1–P)] = 0,10 + 0,25(1) + 0,15(1) = 0,50. Entonces los *odds* = exp(0,50) = 1,65.
9. ln[P/(1–P)] = 0,10 + 0,25(0) + 0,15(1) = 0,25. Entonces el *odds* = exp(0,25) = 1,28.
10. *Odds ratio* (fumador de sexo femenino sobre fumador de sexo masculino) = 1,65 / 1,28 = 1,29.
11. *Odds ratio* (no fumador de sexo femenino sobre no fumador de sexo masculino) = exp(0,35)/exp(0,10) = 1,29.

P13.14

1. Para brindar una predicción precisa del riesgo para ese hombre, el doctor recabará información sobre el comportamiento con respecto al tabaquismo, el índice de masa corporal, la presión arterial sistólica, los niveles de colesterol, el uso de medicación y otros posibles factores de riesgo cardiovascular.
2. Riesgo de morir en diez años = 1/{1 + exp – (9,194 + 0,62*0,104 + 1*0,658 + 25*(-0,041) + 140*0,007 + 245*0,001 + 0*1,251 + 20*0,023 + 80*0,0053)} = 26,8%.
3. Riesgo de morir en diez años = 1/{1 + exp – (9,194 + 1*0,658 + 62*0,104 + 25*(-0,041) + 140*0,007 + 245*0,001 + 0*1,251 + 0*0,023 + 80*0,0053)} = 18,8%.
4. RR = 26,8 / 18,8 = 14,4. Sin considerar la significación estadística, el riesgo para un hombre no diabético que fuma 20 cigarrillos por día es aproximadamente 1,4 veces el riesgo para un hombre no diabético que no fuma.
5. Los dos cálculos del riesgo no ofrecen información que compare un individuo diabético con uno no diabético.
6. Mediante la comparación de dos cálculos de funciones de riesgo, uno en el que la variable de diabetes sea igual a 1 y el otro en el que la variable de diabetes sea igual a 0 y manteniendo los valores correspondientes de todas las otras variables en el mismo valor.

P13.15

1. Verdadero
2. Falso. En un estudio epidemiológico normal, tratamos de determinar los predictores de un resultado de salud. Por lo tanto, la variable dependiente es el resultado de salud que sea de interés.

3. Falso. Por definición, la variable dependiente es la variable Y.
4. Verdadero
5. Verdadero
6. Logístico
7. Coeficiente de regresión
8. 0,25
9. 0,45
10. diferentes
11. promedio
12. cero
13. resultado de salud, de exposición
14. logit

LECCIÓN 14

ANÁLISIS ESTRATIFICADO

Esta opción de análisis para el control de variables extrañas implica los siguientes pasos:

1. *categorización de todas las variables;*
2. *formación de combinaciones de categorías (es decir, estratos);*
3. *realización de análisis específicos para cada estrato;*
4. *elaboración de una evaluación general de la relación entre la exposición y la enfermedad, si correspondiera.*

*Tanto los análisis específicos para cada estrato como la evaluación general requieren una **estimación puntual**, una **estimación de intervalo** y una **prueba de hipótesis**. En esta lección nos concentraremos en la **evaluación general**, que, desde el punto de vista conceptual y matemático, es el más complicado de los cuatro pasos. Para una evaluación general, la estimación puntual es un **cálculo ajustado** que normalmente consiste en un **promedio ponderado** de los cálculos específicos para cada estrato. El **intervalo de confianza** comúnmente es una estimación de intervalo de muestra grande en torno al cálculo ajustado (ponderado). La prueba de hipótesis es una generalización de la **prueba de chi cuadrado de Mantel-Haenszel**.*

14-1 Análisis estratificado

Ejemplo: una variable de control

Ilustramos los cuatro pasos del **análisis estratificado** mediante un ejemplo:

Análisis estratificado

1. Categorización de todas las variables
2. Formación de combinaciones de categorías (estratos)
3. Realización de análisis específicos para cada estrato
4. Una evaluación general de E → D

Los siguientes cuadros contienen datos de un estudio de cohorte retrospectivo hipotético cuyo objetivo es determinar el efecto de la exposición a un supuesto químico tóxico, denominado TCX, en el desarrollo de cáncer de pulmón. Suponga que la única variable de control de interés es el tabaquismo. A continuación delineamos una estrategia analítica en cuatro pasos. Primero, categorizamos la variable de tabaquismo en dos grupos: fumadores y no fumadores. Segundo, formamos cuadros 2 x 2 para cada estrato. Tercero, realizamos **análisis específicos para cada estrato** como se muestra a continuación. Estos datos ejemplifican la confusión. Los datos crudos que ignoran el control del tabaquismo dan una razón de riesgos moderadamente fuerte de 2,1. Este cálculo de la razón de riesgos es distinto de los dos cálculos que se obtienen cuando se controla el tabaquismo, los cuales indican que no hay asociación.

Estudio de cohorte retrospectivo **TCX** → Cáncer de pulmón

Variable de control: Tabaquismo

Datos crudos	TCX	no TCX	Total
CP	27	14	41
No CP	48	67	115
Total	75	81	156

$\hat{RR} = 2.1$

1. Categorización de todas las variables
2. Formación de combinaciones de categorías (estratos)
3. Realización de análisis específicos para cada estrato

No fumadores

	TCX	no TCX	Total
CP	1	2	3
No CP	24	48	72
Total	25	50	75

$\hat{RR} = 1.0$

Fumadores

	TCX	no TCX	Total
CP	26	12	38
No CP	24	19	43
Total	50	31	81

$\hat{RR} = 1.3$

El tercer paso también implica calcular **estimaciones de intervalos** y un **valor p** para cada estrato, y luego interpretar los resultados en forma separada para cada uno de ellos, así como para los datos crudos. Cada análisis **específico de acuerdo a cada estrato** es, básicamente, un análisis simple para un cuadro 2 x 2. Estos son los resultados calculados:

TCX → **Cáncer de pulmón**

Datos crudos $\hat{RR} = 2.1$

95% CI: (1.2,3.7)

P = 0.02

No fumadores $\hat{RR} = 1.0$

95% IC (0.1, 10.6)

P = 1.00

Fumadores $\hat{RR} = 1.3$

95% IC (0.8, 2.3)

P = 0.25

Preguntas de estudio (P14.1)

1. ¿Cuál es su interpretación de los resultados específicos para cada estrato?
2. ¿Parece haber interacción debido al tabaquismo?
3. ¿Parece haber un efecto general de la exposición al TCX después de controlar el tabaquismo?

En el cuarto paso, la evaluación general de la relación **E→D** se debe realizar solo cuando sea adecuado. Esta evaluación debe llevarse a cabo toda vez que existan pruebas de **confusión**. Sin embargo, cuando hay evidencia suficiente de **interacción** o **modificación de efecto**, este paso se considera inadecuado. En nuestro ejemplo, los valores de la razón de riesgos para ambos grupos de la condición de tabaquismo son básicamente iguales, lo que indica que es razonable continuar con una evaluación de resumen o evaluación general.

Para llevar a cabo el cuarto paso, se debe: (a) computar un **cálculo ajustado general** de efecto de la exposición sobre la enfermedad en todos los estratos; (b) realizar una **prueba de hipótesis** para ver si existe un efecto general controlando la estratificación; y (c) calcular e interpretar una **estimación de intervalo** alrededor del valor puntual ajustado.

4. Evaluación general de la relación E ⇒ D
a. Cálculo ajustado: promedio ponderado
b. Prueba de hipótesis: prueba de Mantel-Haenszel para el análisis estratificado
c. Estimación del intervalo: intervalo de confianza de muestra grande

El cálculo ajustado normalmente consiste en algún tipo de **promedio ponderado** de los cálculos específicos para cada estrato. El procedimiento de prueba es la prueba de **Mantel-Haenszel** para el análisis estratificado. La estimación del intervalo se calcula, en general, como un intervalo de confianza de muestra grande basado en los puntos porcentuales de la distribución normal. Estos tres componentes de la evaluación general se describirán en más detalle en las actividades siguientes.

Preguntas de estudio (P14.1) (continuación)

Un valor de la razón de riesgos ajustada basada en la precisión con respecto a la relación entre el TCX y el cáncer de pulmón es igual a 1,25. Un intervalo de confianza de 95% alrededor de este valor resulta ser (0,78; 2,00). El estadístico de prueba de Mantel-Haenszel tiene un valor p de 0,28.

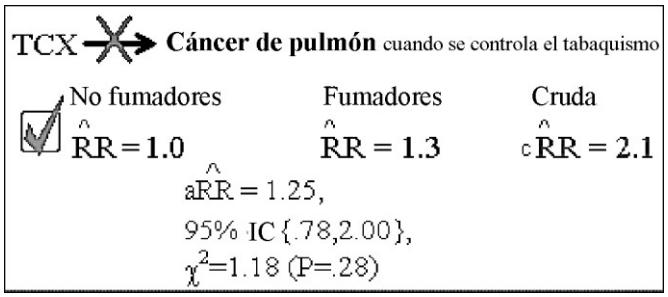
4. A partir de estos resultados, ¿a qué conclusión llega con respecto a la evaluación general de la relación entre la exposición y la enfermedad en este estudio?

Resumen

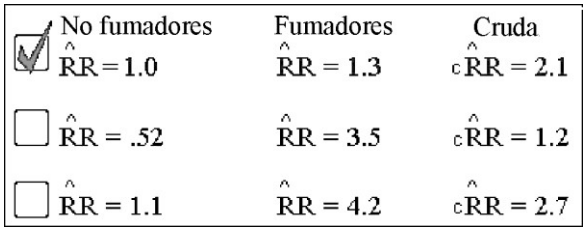
- La forma más simple de estratificación ocurre cuando hay una única variable dicotómica por controlar.
- En este ejemplo, solo se categoriza una variable (paso 1) y se obtienen dos estratos (paso 2).
- Normalmente, el paso 3 consiste en calcular una estimación puntual, una estimación de intervalo y un valor p para cada estrato.
- Es posible que sea inadecuado realizar una evaluación general (paso 4) si existe interacción/modificación de efecto.
- Cuando adecuado, el paso 4 involucra calcular un valor ajustado general de efecto, un intervalo de confianza de muestra grande para el efecto ajustado y una prueba de significación (Mantel-Haenszel).

¿Evaluación general?

Como los valores de la razón de riesgos para ambos grupos relacionados con la condición de tabaquismo son básicamente iguales, hemos llegado a la conclusión de que es razonable continuar con la evaluación general mediante un **cálculo ajustado**, un **intervalo de confianza** alrededor de dicho cálculo y una prueba de **Mantel-Haenszel** para los datos estratificados. A continuación se presentan los resultados. Claramente, indican que no hay un efecto importante o significativo del TCX en el desarrollo de cáncer de pulmón cuando se controla el tabaquismo.



Pero, ¿qué sucede si obtenemos un grupo distinto de cálculos específicos para cada estrato, por ejemplo, los resultados que se muestran a continuación (ejemplos 2 y 3)? ¿Todavía quisiéramos calcular un valor ajustado, obtener un intervalo de confianza alrededor del mismo y calcular una prueba de Mantel-Haenszel?



Nota: las filas de los resultados de las razones de riesgos son, desde arriba hacia abajo, los ejemplos 1, 2 y 3, respectivamente.

Estos dos supuestos muestran una interacción muy fuerte debido al tabaquismo. Además, el tipo de interacción en el ejemplo 2 es bastante diferente de la interacción en el ejemplo 3. Los valores de las razones de riesgos específicas para cada estrato de 0,52 y 3,5 en el ejemplo 2 están del lado opuesto al valor nulo de 1. Por el contrario, los valores de las razones de riesgos específicas para cada estrato de 1,1 y 4,2 del ejemplo 3 están del mismo lado que el valor nulo, aunque también son bastante distintos.

Cuando los efectos específicos para cada estrato están en lados opuestos de 1, como en el ejemplo 2, es posible que se puedan cancelar unos con otros en el cálculo de un efecto ajustado. En consecuencia, en esta situación **no** se recomienda el uso de dicho cálculo ajustado, ni el correspondiente intervalo de confianza ni la prueba de Mantel-Haenszel. Los resultados importantes en este caso están dados por los efectos contrastantes específicos para cada estrato y es probable que queden ocultos al realizar una evaluación general.

Cuando los efectos específicos para cada estrato están todos en la misma dirección, como en el ejemplo 3, no puede surgir una apariencia falsa de falta de asociación a partir de la cancelación de efectos opuestos. Por lo tanto, a pesar de la interacción, puede valer la pena realizar una evaluación general, según la opinión del investigador sobre qué tan grande es la diferencia entre los efectos específicos para cada estrato y qué tan estables son estos cálculos.

Resumen

- Es posible que sea inadecuado realizar una evaluación general (paso 4) si existe interacción/modificación de efecto.
- El caso más convincente para no realizar una prueba general es aquel en el que en lados opuestos del valor nulo hay efectos significativos específicos para cada estrato.
- Cuando todos los efectos específicos para cada estrato están del mismo lado que el valor nulo, la evaluación general puede ser adecuada, incluso si existe interacción.
- La situación más apropiada para realizar una evaluación general es aquella en la que los efectos específicos para cada estrato son todos aproximadamente iguales, indicando ausencia de interacción entre los estratos.

La prueba de Breslow-Day (BD) para la interacción

Esta prueba de interacción puede utilizarse para datos estratificados obtenidos de un estudio de cohorte con incidencia acumulada, un estudio de casos y controles o un estudio transversal. La prueba se basa en la siguiente disposición de datos para el estrato *i*:

Disposición de los datos para el estrato *i*	**E**	**no E**	
D	a_i	b_i	m_{1i}
no D	c_i	d_i	m_{0i}
	n_{1i}	n_{0i}	n_i

La **hipótesis nula** que se somete a prueba indica que **no hay interacción** entre los estratos o, en forma equivalente, que los cálculos de los efectos en los estratos son uniformes (es decir, iguales). El estadístico de prueba tiene la siguiente forma:

$$\chi^2_{BD} = \sum_{i=1}^{G} \frac{(a_i - a'_i)^2}{\hat{Var}_0(a_i)}$$

En esta fórmula, la cantidad a'_i denota el **valor esperado en la celda *a*** del cuadro para el estrato ***i*** suponiendo un efecto ajustado común $\hat{\theta}$ (por ejemplo, **aOR**, **aRR**, **mOR** o **mRR**). Este valor esperado se obtiene al resolver la siguiente ecuación para el término **a** en forma separada para cada estrato.

$$\frac{a'_i(m_{0i} - n_{1i} - a'_i)}{(m_{1i} - a'_i)(n_{1i} - a'_i)} = \hat{\theta}$$

El término de la varianza en la fórmula de **BD** está representado por la siguiente expresión:

$$\hat{Var}_0(a_i) = \frac{1}{\frac{1}{a'_i} + \frac{1}{m_{1i} - a'_i} + \frac{1}{n_{1i} - a'_i} + \frac{1}{m_{0i} - n_{1i} + a'_i}}$$

Ejemplificamos el cálculo de la prueba de BD mediante los datos hipotéticos de los dos cuadros siguientes:

Estrato 1

	E	**no E**	
D	15	5	20
no D	85	95	180
	100	100	200

Estrato 2

	E	**no E**	
D	5	15	20
no D	95	85	180
	100	100	200

$\hat{OR} = 3{,}35$	$\hat{OR} = 0{,}30$

Sobre la base de estos datos y dejando que θ = **aOR**, se calculan los términos en la fórmula de **BD**, que dan los siguientes resultados:

$$\hat{aOR} = 1,\ a_1' = a_2' = 10,\ \hat{Var}_0(a_1) = \hat{Var}_0(a_2) = 4{,}5$$

Al sustituir estos términos en la fórmula de BD se obtiene lo siguiente:

$$\chi^2_{BD} = \frac{(a_1 - a_1')^2}{\hat{Var}_0(a_1)} + \frac{(a_2 - a_2')^2}{\hat{Var}_0(a_2)} = \frac{(15-10)^2}{4{,}5} + \frac{(5-10)^2}{4{,}5} = 11{,}11$$

Puesto que hay **G = 2** estratos, los grados de libertad para el chi cuadrado se calculan como **G** - 1 = 1. El **valor p** para esta prueba es de 0,0008, que es muy pequeño. En consecuencia, llegaríamos a la conclusión de que existe una interacción significativa entre los dos estratos, lo que respalda la interacción observada a partir de la comparación de las estimaciones puntuales en los dos cuadros.

Ejemplo: diversas variables explicativas

En años recientes, la resistencia a los antibióticos se ha convertido en un problema importante en el tratamiento de infecciones bacterianas. Muchos antibióticos que solían resultar efectivos contra ciertas bacterias, especialmente el estafilococo aureus, ya no funcionan porque las nuevas cepas de esta bacteria son resistentes a las drogas antimicrobianas. Cuando se diagnostica a alguien una infección por estafilococo aureus, la primera línea de tratamiento comúnmente involucra drogas antimicrobianas basadas en la meticilina. Sin embargo, las cepas de esta bacteria que son resistentes a esas drogas ahora se consideran un problema importante para los pacientes que se atienden en las guardias. Las bacterias resistentes de este tipo se denominan infecciones por estafilococo aureus resistente a la meticilina o SARM.

Podemos preguntarnos cuáles son las características o factores de riesgo que se asocian a padecer la infección por SARM. Para estudiar esta pregunta, se llevó a cabo un estudio transversal en el Grady Hospital de Atlanta, Georgia, que incluyó a 297 pacientes adultos atendidos en una guardia y cuyos cultivos de sangre -tomados dentro de las 24 horas posteriores al ingreso- indicaron que poseían una infección por estafilococo aureus. Se obtuvo información sobre diversas variables, algunas de las cuales fueron factores de riesgo descritos anteriormente para la resistencia a la meticilina:

Análisis de diversas variables

- Condición de SARM: (1 = Sí, 0=No)
- Hospitalización previa : (1=Sí, 0=No)
- Edad (continua): ⟶ 1= edad>55, 0=edad ≤55
- Sexo : (1=hombre, 0=mujer)
- Drogas antimicrobianas: (1=Sí, 0=No)

Utilizamos esta información como ejemplo de un análisis estratificado para evaluar si la hospitalización previa está asociada a la resistencia a la meticilina, controlando la edad, el sexo y el uso previo de drogas

antimicrobianas. La edad es continua, así que la categorizaremos en dos grupos (1 = edad mayor a 55 años; 0 = edad menor o igual a 55 años).

Primero consideramos los datos crudos que relacionan la hospitalización previa con la condición de SARM:

Datos Crudos	Hosp Prev Sí	No	Total
SARM Sí	103	12	115
SARM No	75	102	177
Total	178	114	292

DATOS CRUDOS

Hosp prev → Condición SARM

$\widehat{cOR} = 11.67\ (5.99, 22.77)$

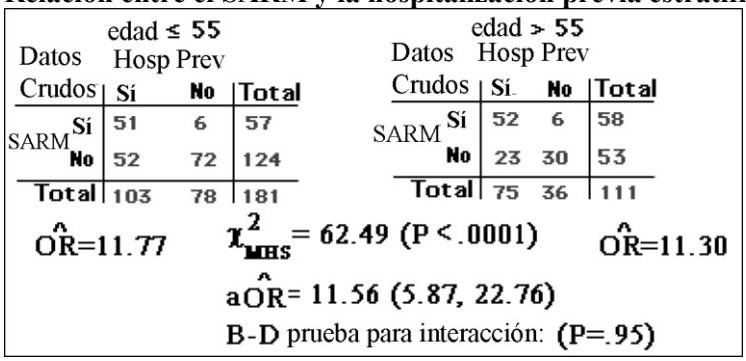
$\chi^2_{MH} = 65.01\ (p < 0.001)$

Preguntas de estudio (P14.2)

1. Observando sólo los datos crudos, ¿la hospitalización previa está asociada a la resistencia a la meticilina?
2. ¿Qué reservas debería tener sobre su respuesta a la pregunta 1?
3. ¿Debería estratificar automáticamente con respecto a la edad, el sexo y el uso anterior de droga antimicrobiana, ya que fueron medidas en el estudio?
4. Considerando que había 297 personas en el estudio, ¿por qué el total general es igual a 292?

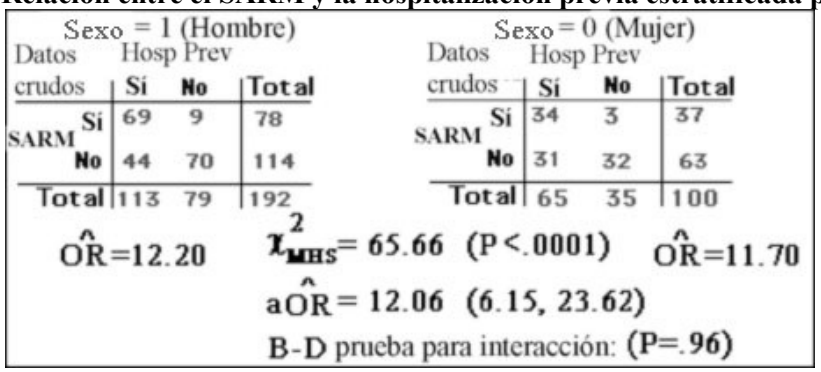
Ahora veamos qué sucede cuando estratificamos en forma separada por la edad, el sexo y el uso previo de droga antimicrobiana. A continuación se muestra cada cuadro estratificado en forma separada.

Relación entre el SARM y la hospitalización previa estratificada por la edad



5. Si sólo nos concentramos en la estratificación por la edad, ¿parece haber interacción/modificación de efecto debido a la edad en función de los resultados específicos para cada estrato?
6. La prueba de Breslow Day (BD) para la interacción brinda un valor p para probar la hipótesis nula de que no hay interacción entre los estratos. Sobre la base de esta prueba con estratificación con respecto a la edad, ¿hay evidencia de interacción?
7. En función de sus respuestas a las preguntas anteriores, ¿es adecuada una evaluación general de la relación entre la exposición y la enfermedad cuando se estratifica con respecto a la edad?
8. ¿Hay confusión debido a la edad? (Ayuda: $c\hat{O}R = 11{,}67$).
9. ¿Parece haber un efecto significativo de la hospitalización previa en el SARM cuando se controla la edad?
10. ¿Qué indica el intervalo de confianza para el valor ajustado sobre este cálculo?

Relación entre el SARM y la hospitalización previa estratificada por el sexo



11. ¿Es adecuado realizar una evaluación general de la relación entre la exposición y la enfermedad cuando solo se estratifica por el sexo?
12. ¿Hay confusión debido al sexo? (Ayuda: $c\hat{O}R = 11{,}67$).
13. ¿Parece haber un efecto significativo de la hospitalización previa en la condición de SARM cuando se controla el sexo?

Relación entre el SARM y la hospitalización anterior estratificada por el uso previo de droga antimicrobiana ("UPDAM")

UPDAM = 1

Datos crudos	Hosp Prev Sí	No	Total
SARM Sí	92	3	95
SARM No	47	13	60
Total	139	16	155

UPDAM = 0

Datos crudos	Hosp Prev Sí	No	Total
SARM Sí	10	9	19
SARM No	27	89	116
Total	37	98	135

$\hat{OR} = 8.48$ $\chi^2_{MHS} = 20.08$ $(P < .0001)$ $\hat{OR} = 3.66$

$\hat{aOR} = 5.00$ (2.26,11.04)

B-D prueba para interacción: (P=.31)

14. ¿Es adecuado realizar una evaluación general de la relación entre la exposición y la enfermedad cuando solo se estratifica por el UPDAM?
15. ¿Hay confusión debido al UPDAM? (Ayuda: $c\hat{OR} = 11,67$).
16. ¿Parece haber un efecto significativo de la hospitalización previa en la condición de SARM cuando se controla el UPDAM?

Resumen

- Cuando se controlan diversas variables mediante el análisis estratificado, el primer paso típico es analizar e interpretar los datos crudos.
- El próximo paso normalmente es estratificar, en forma separada, con respecto a cada variable de control, lo que incluye realizar una evaluación general de la relación entre la exposición y la enfermedad, si corresponde.
- El enfoque para determinar si la evaluación general es adecuada consiste en determinar si los efectos específicos para cada estrato son más o menos iguales.
- Otro enfoque es realizar una prueba de Breslow Day de la hipótesis nula que indica que no hay interacción/modificación de efecto debido a la o las variables que se estratifican.

Ejemplo: diversas variables explicativas (continuación)

A continuación se incluye un cuadro de resumen que surge de estratificar con respecto a cada variable de control en forma separada.

Análisis estratificado para evaluar si la hospitalización previa está asociada a la resistencia a la meticilina

Variable de Control	aOR (95% IC)	MH Valor-P	B-D Valor P
Ninguna	11.67 (5.99, 22.77)	P<.0001	-
Edad	11.56 (5.87, 22.76)	P<.0001	.95
Sexo	12.06 (6.15, 23.62)	P<.0001	.96
UPDAM	5.00 (2.26, 11.04)	P<.0001	.31

Odds ratio crudo = 11.67 (5.99, 22.77)

Preguntas de estudio (P14.3)

1. Sobre la base de la información en el cuadro anterior, ¿cuál de las variables edad, sexo y uso previo de droga antimicrobiana se debe controlar, si es que corresponde alguna?
2. ¿Se gana precisión al controlar alguna de las variables edad, sexo y uso previo de droga antimicrobiana?

Ahora agregamos al cuadro de resumen los resultados de controlar dos o tres variables al mismo tiempo.

Análisis estratificado para evaluar si la hospitalización previa está asociada a la resistencia a la meticilina

Variable de Control	aOR (95% IC)	MH Valor-P	B-D Valor P
Ninguna	11.67 (5.99, 22.77)	P<.0001	-
Edad	11.56 (5.87, 22.76)	P<.0001	.95
Sexo	12.06 (6.15, 23.62)	P<.0001	.96
UPDAM	5.00 (2.26, 11.04)	P<.0001	.31
Edad, Sexo	11.59 (5.91, 22.76)★	P<.0001	.90
Edad, UPDAM	4.63 (2.08, 10.29)	P<.0001	.59
Sexo , UPDAM	5.04 (2.31, 11.03)★	P<.0001	.60
Edad, Sexo , UPDAM	4.66 (2.14, 10.14)★	P<.0001	.74

Odds ratio crudo = 11.67 (5.99, 22.77)

★ **Es necesario utilizar un factor de corrección 0,5 en estratos que contienen frecuencia cero**

3. ¿El control de la edad, el sexo o ambas variables tienen un efecto en los resultados después de que ya se haya controlado el uso previo de droga antimicrobiana (UPDAM)?
4. Mediante el valor p de la prueba de BD, ¿hay evidencia de que existe interacción cuando se estratifica con respecto a cualquier o todas las variables (tres) que se controlan?
5. ¿Por qué cree que es necesario utilizar un factor de corrección como 0,5 en estratos que contienen frecuencia cero?
6. Sobre la base de toda la información del cuadro, ¿cuál es el cálculo más adecuado del *odds ratio* de interés? (Aquí puede elegir dos opciones).
7. ¿Hay evidencia de que la hospitalización previa tenga algún efecto si una persona tiene el estafilococo aureus resistente a la meticilina?

Los resultados específicos para cada estrato cuando se controlan simultáneamente la edad, el sexo y el uso previo de droga antimicrobiana se presentan en el recuadro que se encuentra al final de esta actividad. Existen ocho estratos porque se controlan tres variables y cada una de ellas tiene dos categorías.

Preguntas de estudio (P14.3) (continuación)

(Nota: no hay pregunta 8).

9. ¿Cuál es la característica más obvia que describe los resultados estratificados que se acaban de mostrar?
10. ¿Qué indica su respuesta a la pregunta anterior sobre los análisis específicos para cada estrato con estos estratos?
11. Sobre la base de la comparación de los cálculos del *odds ratio* específicos para cada estrato, ¿parece haber interacción dentro de los datos estratificados?
12. Dé tres razones que justifiquen realizar una prueba general de Mantel-Haenszel con estos datos.

Resumen

- Cuando se controlan simultáneamente diversas variables mediante el análisis estratificado, es posible que no sea necesario controlarlas a todas; dicho control dependerá de si una variable contribuye a la confusión o a la precisión.
- Normalmente, el control simultáneo de diversas variables genera estratos con cantidades pequeñas y, a menudo, frecuencias de celda cero.
- Cuando hay cantidades pequeñas en algunos estratos, las conclusiones específicas para cada estrato pueden no ser confiables.
- Hay tres aspectos por considerar cuando se evalúa la interacción en los datos estratificados:
    - ¿Los cálculos específicos para cada estrato son básicamente iguales?
    - ¿La prueba de Breslow-Day para la interacción es significativa?
    - ¿Los cálculos específicos para cada estrato son poco confiables debido a que hay cantidades pequeñas?

Resultados específicos para cada estrato

A continuación se presentan los resultados específicos para cada estrato cuando se controlan simultáneamente edad, sexo y uso previo de droga antimicrobiana.

1. Edad ≤ 55, masculino, UPDAM = sí

		Hosp. previa sí	Hosp. previa no	
SARM	sí	37	2	39
	no	22	7	29
		59	9	68

$\hat{OR}$ = 5,89

2. Edad ≤ 55, masculino, UPDAM = no

		Hosp. previa sí	Hosp. previa no	
SARM	sí	5	4	9
	no	13	49	62
		18	53	71

$\hat{OR}$ = 4,71

3. Edad ≤ 55, femenino, UPDAM = sí

		Hosp. previa sí	Hosp. previa no	
SARM	sí	9	0	9
	no	14	3	17
		23	3	26

$\hat{OR}$ = 4,59 con ajuste de 0,5

4. Edad ≤ 55, femenino, UPDAM = no

		Hosp. previa sí	Hosp. previa no	
SARM	sí	0	0	0
	no	2	13	15
		2	13	15

$\hat{OR}$ = 5,4 con ajuste de 0,5

5. Edad > 55, masculino, UPDAM = sí

		Hosp. previa sí	Hosp. previa no	
SARM	sí	24	1	25
	no	2	2	4
		26	3	29

$\hat{OR}$ = 24,00

6. Edad > 55, masculino, UPDAM = no

		Hosp. previa sí	Hosp. previa no	
SARM	sí	2	2	4
	no	7	12	19
		9	14	23

$\hat{OR}$ = 1,71

7. Edad > 55, femenino, UPDAM = sí

		Hosp. previa sí	Hosp. previa no	
SARM	sí	22	0	22
	no	9	1	10
		31	1	68

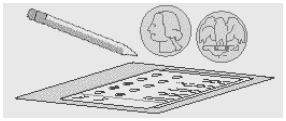
$\hat{OR}$ = 7,11 con ajuste de 0,5

8. Edad > 55, femenino, UPDAM = no

		Hosp. previa sí	Hosp. previa no	
SARM	sí	3	3	6
	no	5	15	20
		8	18	26

$\hat{OR}$ = 3,00

Auto-evaluación (P14.4)



Coloque **verdadero** o **falso** en cada una de las siguientes afirmaciones:

1. La estratificación solo implica categorizar variables en dos grupos y realizar análisis separados para cada uno de ellos. . . . . ???
2. Uno de los cuatro pasos del análisis estratificado es calcular la evaluación general de resumen correspondiente a la relación entre la exposición y la enfermedad cuando sea apropiado. ???
3. El cálculo del valor de resumen general puede considerarse inadecuado si hay muchas pruebas de interacción estadística. . . . . ???
4. El cálculo del valor de resumen general puede considerarse inadecuado si hay muchas pruebas de confusión. . . . . ???
5. Cuando se tiene en cuenta la conveniencia de calcular resultados de resumen generales, el investigador debe emitir su juicio sobre la importancia clínica de las diferencias observadas entre los cálculos específicos para cada estrato, así como sobre la estabilidad de dichos cálculos. . ???
6. Compare los cálculos de RR específicos para cada estrato en cada una de las tres situaciones siguientes. Complete los espacios vacíos con **sí**, **no** o **quizás** sobre el uso adecuado de un cálculo de resumen.

Situación:	RR: estrato 1	RR: estrato 2	Cálculo general
Dirección opuesta	0,7	3,5	???
Misma dirección	1,5	4,8	???
Efecto uniforme	2,3	2,9	???

Opciones
Quizás No Sí

14-2 Análisis estratificado (continuación)

Pruebas de asociación general – La prueba de Mantel-Haenszel

Propósito y características generales

La **prueba de Mantel-Haenszel** es el procedimiento que se utiliza con mayor frecuencia y se recomienda comúnmente para evaluar la asociación general en un análisis estratificado. La **hipótesis nula** que se somete a prueba indica que no hay asociación entre todos los estratos o, en forma equivalente, que la medida de efecto ajustada para todos los estratos es igual al valor nulo. Generalmente, la típica hipótesis alternativa es de una cola ya que, a menudo, el investigador desea determinar si el efecto ajustado es mayor que el valor nulo; sin embargo, también puede considerarse una hipótesis alternativa de dos colas o de una cola inferior. Conforme a la hipótesis nula, el estadístico de prueba tiene aproximadamente la distribución chi cuadrado con un grado de libertad.

Prueba de Mantel-Haenszel

H_0: No hay asociación entre todos los estratos , **aRR** $=1$

H_A: típica es de una cola

e.g., **aRR** > 1



dos colas

e.g., **aRR** $\neq 1$



una cola inferior

e.g., **aRR** < 1



Para variables dicotómicas de enfermedad y exposición, la disposición de los datos en un estrato dado puede describirse mediante uno de los dos formatos que se muestran a continuación. El formato depende del diseño del estudio:

Estudios de Cohorte (RR)
Casos y controles o
Transversales

	E	No E	Total
D	a_i	b_i	m_{1i}
no D	c_i	d_i	m_{0i}
Total	n_{1i}	n_{0i}	n_i

Estudio de cohorte con información de persona-tiempo (RDI)

	E	No E	Total
D	I_{1i}	I_{0i}	I_i
Persona-Tiempo	PT_{1i}	PT_{0i}	PT_i

Pero, independientemente del diseño del estudio, el estadístico de prueba tiene la estructura de cálculo básica que se indica aquí:

$$\chi^2_{MHS} = \frac{[\,a - E_0(a)\,]^2}{Var_0(a)}$$

a = cantidad total de casos expuestos

$E_0(a)$= cantidad total esperada de casos expuestos según la hipótesis nula

$Var_0(a)$= varianza de la cantidad total de casos expuestos según la hipótesis nula

En esta fórmula, **a** denota la cantidad total de casos expuestos; es decir, **a** es la suma de las frecuencias de la celda **a** de todos los estratos. El término $\mathbf{E_0(a)}$ en el numerador da la cantidad total **esperada** de casos expuestos según la hipótesis nula de ausencia de asociación entre la exposición y la enfermedad en cualquiera de los estratos. El término $\mathbf{Var_0(a)}$ describe la varianza de la cantidad total de casos expuestos según la misma hipótesis nula. La forma de cálculo específica para el valor esperado y para el término de la varianza depende del diseño del estudio. Para la primera de las dos disposiciones de datos, estos términos se derivan suponiendo la distribución hipergeométrica que presentamos en la lección 12. Para un estudio de cohorte con información de persona-tiempo, el valor esperado y los términos de la varianza se derivan mediante la distribución binomial. (Nota: véanse los dos recuadros al final de esta actividad para una descripción de cómo se derivan los valores esperados y los términos de la varianza si existe distribución hipergeométrica o distribución binomial).

Para los estudios de cohorte con incidencia acumulada, los de casos y controles, al igual que los transversales, la fórmula de cálculo se simplifica a la siguiente expresión:

Estudios de Cohorte (RR)
Casos y controles o
Transversales

	E	No E	Total
D	a_i	b_i	m_{1i}
no D	c_i	d_i	m_{0i}
Total	n_{1i}	n_{0i}	n_i

$$\chi^2_{MHS} = \frac{[a - E_0(a)]^2}{Var_0(a)}$$

$$\chi^2_{MHS} = \frac{\left(\sum_{i=1}^{G} \frac{a_i d_i - b_i c_i}{n_i}\right)^2}{\sum_{i=1}^{G} \frac{n_{1i} n_{0i} m_{1i} m_{0i}}{(n_i - 1) n_i^2}}$$

Esta fórmula se escribe en términos de las frecuencias de celda y las frecuencias totales dadas en la disposición por estrato. Cuando hay sólo dos estratos, de forma que **G** es igual a dos, la fórmula se simplifica aún más:

$$\chi^2_{MHS} = \frac{\left(\frac{a_1 d_1 - b_1 c_1}{n_1} + \frac{a_2 d_2 - b_2 c_2}{n_2}\right)^2}{\frac{n_{11} n_{01} m_{11} m_{01}}{(n_1 - 1) n_1^2} + \frac{n_{12} n_{02} m_{12} m_{02}}{(n_2 - 1) n_2^2}}$$

Ahora aplicamos esta fórmula a los datos del estudio de cohorte retrospectivo que relaciona el TCX con el cáncer de pulmón. Al sustituir los datos de los dos estratos en la fórmula de Mantel-Haenszel, el valor calculado resulta ser 1,18. Para una hipótesis alternativa de dos colas, el valor p es el área bajo la distribución chi cuadrado con 1 grado de libertad por arriba del valor 1,18. Esto es igual a 0,28.

Estudio de cohorte retrospectivo TCX → Cáncer de pulmón

No Fumadores

	TCX	no TCX	Total
CP	1	2	3
No CP	24	48	72
Total	25	50	75

$\hat{RR} = 1.0$

Fumadores

	TCX	no TCX	Total
CP	26	12	38
No CP	24	19	43
Total	50	31	81

$\hat{RR} = 1.3$

$$\chi^2_{MHS} = \frac{\left(\frac{1*48 - 2*24}{75} + \frac{26*19 - 12*24}{81}\right)^2}{\frac{25*50*3*72}{(75-1)75^2} + \frac{50*31*38*43}{(81-1)81^2}} = 1.18$$

Preguntas de estudio (P14.5)

1. Ya que el valor p para una prueba de Mantel-Haenszel de dos colas es 0,28, ¿cuál es el valor p para una hipótesis alternativa de una cola?
2. Sobre la base de su respuesta a las preguntas anteriores, ¿a qué conclusión llega sobre el efecto general del TCX en el cáncer de pulmón cuando se controla el tabaquismo?

(Esta cuestión se analizará en la próxima actividad). En la fórmula de MH para dos estratos, observe que el numerador suma dos términos, uno de cada estrato, antes de que esta suma se eleve al cuadrado.

3. ¿Qué podría suceder si, en lugar de los resultados descritos anteriormente, la razón de riesgos para los no fumadores estuviese del lado contrario del valor nulo, en vez de la razón de riesgos para los fumadores?

Resumen

- La prueba de Mantel-Haenszel (MH) es el procedimiento que se utiliza con mayor frecuencia y se recomienda comúnmente para evaluar la asociación general en un análisis estratificado.
- La hipótesis nula indica que no hay asociación general entre todos los estratos o, en forma equivalente, que la medida de efecto ajustada es igual al valor nulo.
- La típica hipótesis alternativa es de una cola, aunque puede ser de dos colas.
- Conforme a la hipótesis nula, el estadístico de prueba de MH tiene aproximadamente la distribución chi cuadrado con 1 grado de libertad.
- Para los estudios de cohorte con incidencia acumulada, los estudios de casos y controles, así como los estudios transversales, el estadístico de MH se deriva suponiendo la distribución hipergeométrica.
- Para los estudios de cohorte con información de persona-tiempo, el estadístico de MH se deriva mediante la distribución binomial.

Derivación de la prueba de Mantel-Haenszel para estudios de cohorte con incidencia acumulada, estudios de casos y controles, como también estudios transversales

Aquí describimos cómo se deriva la **prueba de Mantel-Haenszel (MH)** para **estudios de cohorte con incidencia acumulada, estudios de casos y controles**, así como **estudios transversales** donde la disposición de los datos para el estrato i-avo es la siguiente:

Disposición de los datos para el estrato *i*

	E	**no E**	
D	a_i	b_i	m_{1i}
no D	c_i	d_i	m_{0i}
	n_{1i}	n_{0i}	n_i

Para cada uno de los diseños de estudio mencionados anteriormente, la prueba de Mantel-Haenszel supone que las frecuencias marginales $\mathbf{n_{1i}}$, $\mathbf{n_{0i}}$, $\mathbf{m_{1i}}$, $\mathbf{m_{0i}}$ están **fijas** dentro de cada estrato. Surge que, al igual que en un análisis simple con márgenes fijos, la distribución de la cantidad de casos expuestos (en la celda **a**) es **hipergeométrica** según la **hipótesis nula que sostiene que no existe asociación general controlando las variables estratificadas**. De la distribución hipergeométrica, el valor esperado y la varianza nula de la cantidad de casos expuestos dentro de cada estrato están dados entonces por las siguientes fórmulas:

$$E_0(a_i) = \frac{n_{1i}m_{1i}}{n_i}$$

y

$$Var_0(a_i) = \frac{n_{1i}n_{0i}m_{1i}m_{0i}}{(n_i - 1)n_i^2}$$

de forma que, al sumar todos los estratos, obtenemos que la media esperada total y la media de la varianza total de los casos expuestos son:

$$E_0(a_i) = \sum_{i=1}^{G} E_0(a_i) = \sum_{i=1}^{G} \frac{n_{1i} m_{1i}}{n_i}$$

y

$$Var_0(a_i) = \sum_{i=1}^{G} Var_0(a_i) = \sum_{i=1}^{G} \frac{n_{1i} n_{0i} m_{1i} m_{0i}}{(n_i - 1) n_i^2}$$

Al sustituir las expresiones para la media esperada y la varianza del total de casos expuestos en la fórmula general del estadístico de prueba de Mantel Haenszel (MH):

$$\chi^2_{MH} = \frac{[a - E_0(a)]^2}{Var_0(a)}$$

el estadístico de prueba se simplifica (mediante la aplicación de principios algebraicos) en la siguiente fórmula de cálculo:

$$\chi^2_{MH} = \frac{\left[\sum_{i=1}^{G} \frac{a_i d_i - b_i c_i}{n_i}\right]^2}{\sum_{i=1}^{G} \frac{n_{1i} n_{0i} m_{1i} m_{0i}}{(n_i - 1) n_i^2}}$$

Según la hipótesis nula que sostiene que no existe asociación general, este estadístico de MH tiene aproximadamente una distribución chi cuadrado con 1 grado de libertad.

Derivación de la prueba de Mantel-Haenszel para los estudios con datos de persona-tiempo

Aquí describimos cómo se deriva la **prueba de Mantel-Haenszel (MH)** para los **estudios de cohorte con información de persona-tiempo**, donde la disposición de los datos para el estrato i-avo es la siguiente:

Estudio de cohorte con persona-tiempo
Disposición de los datos para el estrato i

	E	**no E**	
D	I_{1i}	I_{0i}	I_i
PT	PT_{1i}	PT_{0i}	PT_i

Según la **hipótesis nula que indica que no existe asociación general controlando las variables estratificadas**, esta versión de la prueba de MH supone que la distribución de la cantidad de casos expuestos (en la celda "a") dentro de cada estrato es **binomial**. En especial, se supone que la cantidad de casos expuestos en el estrato i es una variable aleatoria binomial con probabilidad de éxito representada por:

$$p_{0i} = \frac{PT_{1i}}{PT_i}$$

y con una cantidad de ensayos igual a la cantidad total de casos en el estrato, es decir:

$$N_i = I_{1i} + I_{0i} = I_i$$

de forma que la varianza de la probabilidad binomial es:

$$Var_0(p_{0i}) = N_i p_{0i}(1 - p_{0i}) = \frac{I_i PT_{1i} PT_{0i}}{PT_i^2}$$

Entonces, surge que la media y la varianza nulas para la celda "a" en el estrato i-avo están dadas por la fórmula:

$$E_0(a_i) = \frac{I_i PT_{1i}}{PT_i}$$

y

$$Var_0(a_i) = \frac{I_i PT_{1i} PT_{0i}}{PT_i^2}$$

de forma que, al sumar todos los estratos, obtenemos que la media esperada total y la media de la varianza total de los casos expuestos son:

$$E_0(a) = \sum_{i=1}^{G} E_0(a_i) = \sum_{i=1}^{G} \frac{I_i PT_{1i}}{PT_i}$$

y

$$Var_0(a) = \sum_{i=1}^{G} Var_0(a_i) = \sum_{i=1}^{G} \frac{I_i PT_{1i} PT_{0i}}{PT_i^2}$$

La forma específica de la prueba de MH para los datos de persona-tiempo se obtiene, entonces, mediante la sustitución de las expresiones para la media esperada y la varianza del total de casos expuestos en la fórmula general del estadístico de prueba de Mantel Haenszel (MH):

$$\chi^2_{MH} = \frac{[a - E_0(a)]^2}{Var_0(a)}$$

donde *a* denota la cantidad total de casos expuestos observados en todos los estratos, es decir:

$$a = \sum_{i=1}^{G} I_{1i}$$

Según la hipótesis nula que sostiene que no existe asociación general, este estadístico de MH tiene aproximadamente una distribución chi cuadrado con 1 grado de libertad.

Cuándo no se debe utilizar la prueba de MH

Consideremos un ejemplo en el que **no** sería adecuada la **prueba de Mantel-Haenszel**. Suponga que los resultados del estudio de cohorte retrospectivo que relaciona el TCX con el cáncer de pulmón fuesen los siguientes:

Resultados del estudio de cohorte retrospectivo TCX ⟶ LC

Datos crudos	TCX	no TCX	Total
CP	20	20	40
No CP	180	180	360
Total	200	200	400

$\hat{RR} = 1.0$ (P=1.00)

No fumadores

	TCX	no TCX	Total
CP	15	5	20
No CP	85	95	180
Total	100	100	200

$\hat{RR} = 3.0$ (P=.01)

Fumadores

	TCX	no TCX	Total
CP	5	15	20
No CP	95	85	180
Total	100	100	200

$\hat{RR} = .33$ (P=.01)

Estos datos indican, para los no fumadores, una razón de riesgos positiva moderadamente grande que es muy significativa y, para los fumadores, una razón de riesgos que describe una asociación negativa moderadamente grande, pero que también es muy significativa. Por el contrario, los datos crudos muestran una razón de riesgos no significativa, exactamente igual a 1.

Ambos efectos específicos de cada estrato son fuertes y significativos, pero indican **interacción en direcciones opuestas**.

Veamos lo que sucede si, de todas formas, realizamos, inadecuadamente, una prueba de Mantel-Haenszel. Aquí presentamos otra vez la fórmula de Mantel-Haenszel para el caso en que hay dos estratos:

$$\chi^2_{MHS} = \frac{\left(\frac{a_1 d_1 - b_1 c_1}{n_1} + \frac{a_2 d_2 - b_2 c_2}{n_2}\right)^2}{\frac{n_{11} n_{01} m_{11} m_{01}}{(n_1 - 1) n_1^2} + \frac{n_{12} n_{02} m_{12} m_{02}}{(n_2 - 1) n_2^2}}$$

Al sustituir los datos de los dos estratos en la fórmula de Mantel-Haenszel, el valor calculado resulta ser exactamente cero. En especial, observe que los dos términos en el numerador de la fórmula son el negativo del otro, de forma tal que, cuando se suman, el resultado es cero.

$$\chi^2_{MHS} = \frac{\left(\frac{15*95 - 5*85}{200} + \frac{5*85 - 15*95}{200}\right)^2}{\left(\frac{100*100*20*180}{(200 - 1)200^2} + \frac{100*100*20*180}{(200 - 1)200^2}\right)} = 0.0$$

Cuando hay interacción en direcciones opuestas, el empleo de la prueba de Mantel-Haenszel es, a menudo, inadecuado porque puede ocultar un efecto de interacción fuerte que refleja la verdadera relación entre la exposición y la enfermedad.

Preguntas de estudio (P14.6)

1. Si hay cuatro estratos o más involucrados en el cálculo del estadístico de MH, ¿es posible que aun la prueba arroje un resultado erróneo?
2. ¿Qué puede decir sobre el resultado de la prueba de MH si hay interacción fuerte, pero del mismo lado que el valor nulo?

Resumen

- La prueba de MH puede dar resultados erróneos y, generalmente, no es adecuada cuando los valores específicos para cada estrato están en lados opuestos del valor nulo.
- Si hay efectos fuertes y significativos en lados opuestos del valor nulo, la prueba de MH podría no ser significativa, dando, de esta manera, una impresión errónea de ausencia de efecto de la exposición en la enfermedad.
- El numerador en la fórmula de MH es el cuadrado de una suma, en lugar de una suma de términos al cuadrado.
- En consecuencia, los términos positivos y negativos en la suma pueden cancelarse unos con otros, dando, de este modo, una prueba de MH no significativa.

La prueba de MH para datos de cohorte con persona-tiempo

Consideremos los datos de un **estudio de cohorte con información de persona-tiempo** con respecto a la posible asociación entre la obesidad y la mortalidad por cualquier causa entre mujeres de 60 a 75 años de una población urbana del noreste de Estados Unidos. Consideramos 3 estratos que se definen mediante la categorización de la edad en 3 grupos.

Estudio de cohorte con información de persona-tiempo — Obesidad → Mortalidad

EDAD 60-64	OB	no OB	Total
DTH	7	9	16
P-años	234.5	544.5	779

$\hat{RDI} = 1.81$ (N.S.)

EDAD 65-69	OB	no OB	Total
DTH	11	11	22
P-años	264.5	444.5	709

$\hat{RDI} = 1.68$ (N.S.)

EDAD 70-74	OB	no OB	Total
DTH	12	16	28
P-años	200	410	610

$\hat{RDI} = 1.54$ (N.S.)

Cruda	no OB	no OB	Total
DTH	30	36	66
P-años	699	1399	2098

$\hat{RDI} = 1.67$ (P = .02)

Al observar estos datos, las razones de densidades de incidencia para los 3 estratos etarios tienen, aproximadamente, el mismo tamaño y son relativamente cercanas a la razón de densidades de incidencia cruda de 1,67. Por lo tanto, hay pocas pruebas de interacción, si es que hay alguna. Parece justificarse un cálculo ajustado general, al igual que un intervalo de confianza correspondiente alrededor de este valor y una prueba de Mantel-Haenszel para la asociación general. Aquí, solo nos concentramos en la prueba de Mantel Haenszel.

A continuación se presenta la disposición general de los datos para el estrato i-avo en los estudios de cohorte con información de persona-tiempo.

Estudio de cohorte con información de persona-tiempo (RDI)

	E	No E	Total
D	I_{1i}	I_{0i}	I_i
Persona tiempo	PT_{1i}	PT_{0i}	PT_i

Observe que la celda **a** en el cuadro anterior se denota con $\mathbf{I_{1i}}$ y la celda **b** con $\mathbf{I_{0i}}$, en lugar de $\mathbf{a_i}$ y $\mathbf{b_i}$, respectivamente. A continuación presentamos nuevamente la forma general del estadístico de Mantel-Haenszel:

$$\chi^2_{MHS} = \frac{[\,a - E_0(a)\,]^2}{Var_0(a)}$$

$\mathbf{a}$ = cantidad total de casos expuestos $= \sum_{i=1}^{G} I_{1i}$

$\mathbf{E_0(a)}$ = cantidad esperada de casos expuestos según la hipótesis nula de que no hay asociación general

$\mathbf{Var_0(a)}$ = varianza de la cantidad total de casos expuestos según la misma hipótesis nula

En este supuesto, **a** denota la cantidad total de casos expuestos. $\mathbf{E_0(a)}$ en el numerador denota la cantidad esperada de casos expuestos según la hipótesis nula de que no hay asociación general. $\mathbf{Var_0(a)}$ denota la varianza de la cantidad total de casos expuestos según la misma hipótesis nula.

El valor esperado y los términos de la varianza pueden derivarse mediante una distribución binomial, para la cual la probabilidad de éxito es la proporción de **PT** total en el estrato *i*-avo que proviene de personas expuestas, y la cantidad de ensayos es la cantidad total de casos en el estrato *i*-avo.

Distribución binomial (n,p)

$$p = \frac{PT_{1i}}{PT_i}$$

$$n = I_{1i} + I_{0i} = I_i$$

Ahora aplicamos esta fórmula a los datos de cohorte de persona-tiempo que relacionan la obesidad con la mortalidad. Al sustituir los datos de los tres estratos en la fórmula de Mantel-Haenszel, el valor chi cuadrado calculado resulta ser igual a 4,14. El área bajo la distribución chi cuadrado con 1 grado de libertad por arriba del valor 4,14 es 0,04 (valor p de dos colas).

$$\chi^2_{MHS} = \frac{[\,30 - E_0(a)\,]^2}{Var_0(a)} = 4.14 \quad a = 30$$

Preguntas de estudio (P14.7)

1. ¿Cuál es el valor p para una hipótesis alternativa de una cola?
2. Sobre la base de su respuesta a la pregunta anterior, ¿a qué conclusión llega sobre el efecto general de la obesidad en la mortalidad cuando se controla la edad?

Resumen

- Para los estudios de cohorte con datos de persona-tiempo, al igual que para los estudios de cohorte con incidencia acumulada, los de casos y controles, como también los transversales, el estadístico de MH tiene aproximadamente una distribución chi cuadrado con 1 grado de libertad según la hipótesis nula de que no hay asociación general.
- La forma particular que adopta la fórmula de MH para datos de persona-tiempo se deriva suponiendo que la cantidad de casos expuestos en un estrato dado tiene una distribución binomial.
- Para la distribución binomial anterior, la probabilidad de éxito es la proporción de PT total en el estrato g-avo que proviene de personas expuestas, y la cantidad de ensayos es la cantidad total de casos en el estrato g-avo.

Auto-evaluación (P14.8)

Complete los espacios vacíos

Se llevó a cabo un estudio de cohorte sobre la actividad física (AF) y la incidencia de la diabetes durante un período de seis años entre hombres japoneses-estadounidenses en Honolulu. A continuación se resumen los datos de ese estudio, estratificados según el índice de masa corporal (IMC):

	IMC Alto			IMC Bajo	
	AF alta	AF baja		AF alta	AF baja
Diabetes	48	62	Diabetes	54	71
Persona-años	1050	1067	Persona-años	1132	1134

1. ¿Cuál es la razón de tasas (RDI) para el grupo con IMC alto? . . . . **???**
2. ¿Cuál es la razón de tasas (RDI) para el grupo con IMC bajo? . . . . **???**
3. ¿Cuál es la RDI cruda para estos datos? (Sugerencia: combine los datos de ambos estratos). **???**
4. ¿Hay evidencia de confusión en estos datos? . . . . . **???**
5. ¿Hay evidencia de interacción en estos datos? . . . . . **???**

Opciones
0,76 0,77 0,79 1,27 1,30 1,32 no sí

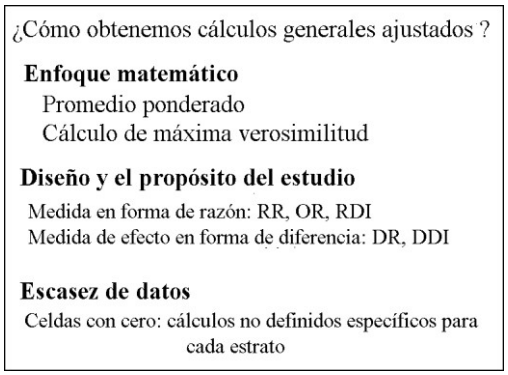
6. ¿Cuál sería la ventaja de realizar una evaluación general para estos datos? . . **???**
7. El estadístico chi cuadrado de Mantel-Haenszel es 3,83, que corresponde a un valor p de 2 colas de 0,051. Suponiendo un nivel de significación de 5%, ¿rechaza o no rechaza la hipótesis nula? **???**
8. Suponga que este estadístico hubiese sido una prueba de una cola, ¿cuál sería su conclusión? **???**

Opciones
menor precisión no se rechaza la hipótesis nula mayor precisión se rechaza la hipótesis nula

14-3 Análisis estratificado (continuación)

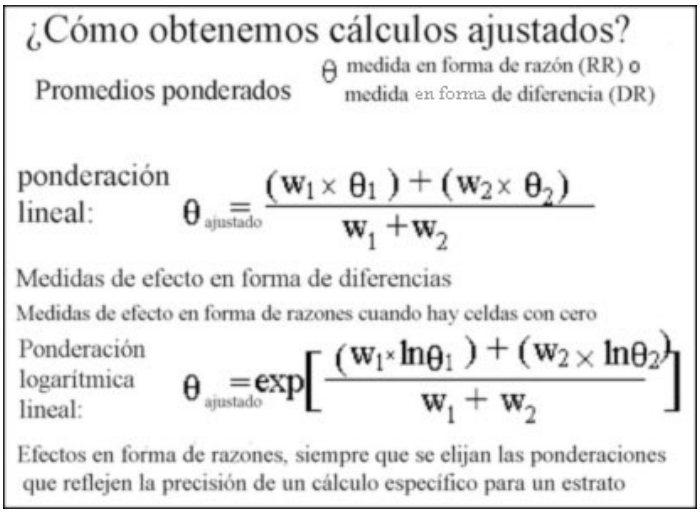
Evaluación general mediante cálculos ajustados

¿Cómo obtenemos cálculos generales ajustados de efecto en un análisis estratificado? La respuesta depende de: 1) la clase de **enfoque matemático** considerado; 2) el tipo de **medida de efecto** adecuada para el diseño y el propósito del estudio; y 3) la **escasez** de datos dentro de los estratos. La forma matemática puede ser un **promedio ponderado** o un cálculo de **máxima verosimilitud**. El diseño y el propósito del estudio pueden requerir una **medida de efecto de razones**, como, por ejemplo, una razón de riesgos, un *odds ratio* o una razón de tasas, o una **medida de efecto de diferencias**, como, por ejemplo, una diferencia de riesgos o una diferencia de tasas. Además, si algunos de los estratos contienen celdas con cero, el cálculo ajustado debe adaptarse a cálculos no definidos, específicos para cada estrato.



Los **promedios ponderados**, que son fáciles de comprender, se utilizan con mayor frecuencia que los cálculos de **máxima verosimilitud** para el resumen de análisis estratificados, de forma que aquí solo nos centraremos en los cálculos de promedios ponderados. Sin embargo, los cálculos de máxima verosimilitud se obtienen en forma práctica mediante la adaptación de modelos matemáticos, como, por ejemplo, el modelo logístico, y dichos cálculos se utilizan normalmente cuando la modelación se considera adecuada. Afortunadamente, en muchos casos, los resultados obtenidos de estos dos enfoques matemáticos son muy cercanos o idénticos, especialmente para datos equiparados.

Los cálculos ajustados que son promedios ponderados tienen dos formas, que se muestran después de este párrafo en el caso simple donde hay solo dos estratos. Denotamos la medida de efecto como **theta** (**θ**). Entonces, según el diseño y el propósito del estudio, **θ** puede ser una medida en forma de razón, como, por ejemplo, una razón de riesgos, o una medida en forma de diferencia, como, por ejemplo, una diferencia de riesgos.



Para las medidas de efecto en forma de diferencias, como la diferencia de riesgos y la diferencia de tasas, siempre se utiliza la ponderación lineal, la cual también puede emplearse en ciertos tipos de medidas de efecto en forma de razones, especialmente cuando hay celdas con cero en alguno de los estratos. Por el contrario, la ponderación logarítmica lineal se usa para obtener cálculos ajustados para efectos en forma de razones, siempre que se elijan las ponderaciones que reflejen la precisión de un cálculo específico para un estrato. Se dice que dichos cálculos ajustados están **basados en la precisión**.

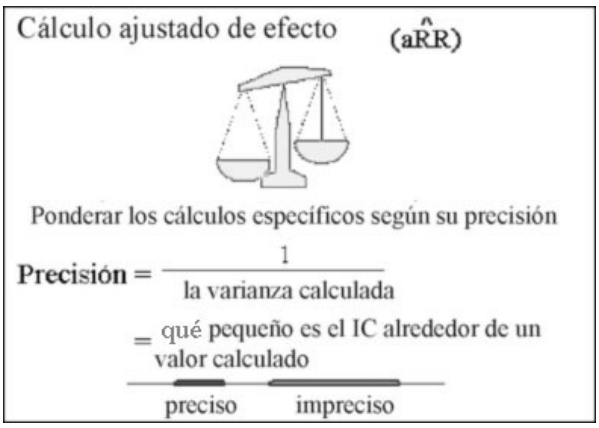
Cuando hay celdas con cero o incluso datos escasos sin celdas con cero, se utiliza normalmente una forma especial de promedio ponderado conocida como **cálculo ajustado de Mantel-Haenszel**. A menudo se emplean cálculos de Mantel-Haenszel en lugar de cálculos basados en la precisión, incluso cuando los datos específicos para cada estrato no son escasos. Esto se debe a que se demostró que los cálculos de Mantel Haenszel tienen propiedades estadísticas deseables, sin olvidar que son idénticos a los cálculos de máxima verosimilitud para datos de casos y controles. La forma específica que adoptan los cálculos basados en la precisión y los cálculos de Mantel-Haenszel para las medidas de efecto en forma de razones se describirá y ejemplificará en las actividades siguientes.

Resumen

- La forma matemática del cálculo ajustado puede ser un promedio ponderado o un cálculo de máxima verosimilitud.
- El diseño y el propósito del estudio pueden requerir una medida en forma de razón, como, por ejemplo, una razón de riesgos, un *odds ratio* o una razón de tasas, o una medida de efecto en forma de diferencia, como, por ejemplo, una diferencia de riesgos o una diferencia de tasas.
- Si algunos de los estratos contienen celdas con cero, el cálculo ajustado debe adaptarse a cálculos no definidos, específicos para cada estrato.
- El promedio ponderado puede involucrar una ponderación lineal o una ponderación logarítmica lineal, donde la última se utiliza para medidas en forma de razones cuando las ponderaciones están basadas en la precisión.
- Los cálculos de Mantel-Haenszel se utilizan cuando hay datos escasos (es decir, celdas con cero), pero también tienen buenas propiedades estadísticas incluso cuando no hay celdas con cero.

Cálculos ajustados mediante el uso de la ponderación por la precisión

Una manera común de encontrar un **cálculo ajustado de efecto** (p. ej., una razón de riesgos ajustada), es **ponderar los cálculos específicos para cada estrato** según su **precisión**. Este abordaje está disponible en la mayoría de los programas informáticos que calculan valores ajustados. Matemáticamente, definimos la precisión de un valor calculado como 1 sobre la varianza de ese valor calculado. También podemos conceptualizar la precisión como una medida de qué considerablemente estrecho es el intervalo de confianza alrededor de un valor calculado. Cuanto más estrecho es el intervalo de confianza, más preciso es el cálculo.



Para dos estratos, la fórmula de la razón de riesgos ajustada basada en la precisión es un promedio ponderado de los logaritmos naturales de las razones de risgos de acuerdo a cada estrato, en lugar de un promedio ponderado de los cálculos de las razones de riesgos en sí. Dadas ponderaciones adecuadas $\mathbf{w_1}$ y $\mathbf{w_2}$, primero se obtiene un promedio ponderado del logaritmo natural de las razones de riesgos específicas para cada estrato, y luego se exponencia este promedio ponderado para obtener un valor que esté en una escala de la razón de riesgos.

$\widehat{(aRR)}$ basado en la precisión (2 estratos)

$$\widehat{aRR} = \exp\left[\frac{w_1 \times \ln(\widehat{RR}_1) + w_2 \times \ln(\widehat{RR}_2)}{w_1 + w_2}\right]$$

Tomamos logaritmos naturales para simplificar algunos de los principios matemáticos subyacentes. Estas ponderaciones están basadas en la precisión y se obtienen computando el inverso de la varianza calculada del logaritmo natural de las razones de riesgos específicas para cada estrato. Mediante una disposición de datos de 2 x 2 para el estrato **i**, la fórmula real para las ponderaciones se muestra a continuación en términos de las frecuencias de celda **a**, **b**, **c** y **d**, y los totales de exposición $\mathbf{n_1}$ y $\mathbf{n_0}$ para el estrato **i**.

$$w_i = \frac{1}{\widehat{Var}(\ln \widehat{RR}_i)} = \frac{1}{\left(\frac{c_i}{n_{1i} a_i}\right) + \left(\frac{d_i}{n_{0i} b_i}\right)}$$

estrato i	E	No E	Total
D	a_i	b_i	m_{1i}
no D	c_i	d_i	m_{0i}
Total	n_{1i}	n_{0i}	n_i

Dado que las ponderaciones, así como la fórmula, son difíciles de calcular, especialmente cuando hay más de 2 estratos, recomendamos utilizar una computadora para realizar los cálculos. Para el ejemplo del estudio de cohorte retrospectivo que involucraba dos estratos de los antecedentes de tabaquismo, la fórmula es la siguiente:

$$\hat{aRR} = \exp\left[\frac{0.69 \times \ln(1.0) + 14.38 \times \ln(1.3)}{0.69 + 14.38}\right]$$

Estudio de cohorte retrospectivo de trabajadores de la química

Estrato 1: no fumadores

	TCX	No TCX	Total
CP	1	2	3
No CP	24	48	72
Total	25	50	75

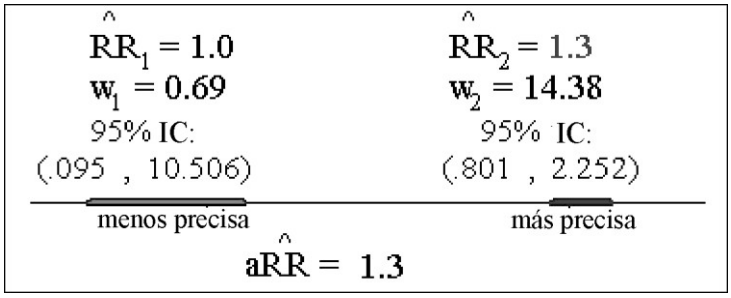
Estrato 2: fumdores

	TCX	No TCX	Total
CP	26	12	38
No CP	24	19	43
Total	50	31	81

$\hat{RR}_1 = 1.0$ $\quad w_1 = 0.69$

$\hat{RR}_2 = 1.3$ $\quad w_2 = 14.38$

Nuestro ejemplo del estudio de cohorte retrospectivo arroja amplias discrepancias entre las cuatro frecuencias de celda del cuadro. En especial, para los no fumadores, una de las cuatro celdas tiene sólo un individuo y otra tiene sólo dos, pero las otras dos celdas tienen frecuencias bastante grandes. Decimos que dicho cuadro está **desequilibrado**. Con un programa informático para realizar los cálculos, encontramos los siguientes intervalos de confianza de 95% para la razón de riesgos calculada en cada cuadro. El primer cuadro muestra un intervalo de confianza mucho más amplio y menos preciso que el segundo cuadro. Por lo tanto, cuando se calcula el valor ajustado basado en la precisión, se debería asignar más peso al segundo cuadro en lugar de hacerlo en el primero.

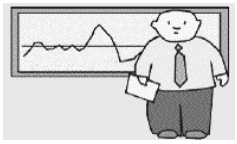


Con un programa informático para realizar los cálculos, encontramos que el valor ajustado de la razón de riesgos se redondea a 1,3. Este valor es el mismo que el que se obtuvo para la razón de riesgos de los fumadores. Por lo tanto, la razón de riesgos ajustada asigna casi todo el peso a los fumadores, cuya razón de riesgos es mucho más precisa que el valor obtenido para los no fumadores.

Cuando se pondera por la precisión, se les asigna más peso a las razones de riesgos de **cuadros equilibrados** que a las razones de riesgos de **cuadros desequilibrados** del mismo tamaño muestral total. Cuando calculamos valores ajustados basados en la precisión, la ponderación utilizada para cada estrato está influenciada no solo por el tamaño muestral total, sino también por el grado de equilibrio en las cuatro celdas del cuadro.

Resumen

- Cuando se realizan cálculos ajustados, las ponderaciones típicas están basadas en la precisión.
- Cuanto más pequeña es la varianza, mayor es la precisión. Asimismo, cuanto más estrecho es el intervalo de confianza, mayor es la precisión.
- La fórmula basada en la precisión para una aRR se obtiene exponenciando un promedio ponderado del logaritmo de los valores de la RR.
- Cada ponderación es el valor inverso de la varianza del logaritmo natural de los cálculos específicos para cada estrato.
- Cuando se pondera por la precisión, se les asigna más peso a las razones de riesgos de cuadros equilibrados que a las razones de riesgos de cuadros desequilibrados del mismo tamaño muestral total.



Auto-evaluación (P14.9)

Complete los espacios vacíos

1. Cuando se pondera por la precisión, ¿cuál de los siguientes ítems NO es un factor en la determinación de las ponderaciones específicas para cada estrato basadas en la precisión? . . **???**

    A. El tamaño de la muestra
    B. La magnitud de la razón de riesgos
    C. El equilibrio del cuadro
    D. La varianza del valor calculado

Verdadero o falso

2. Cuando se pondera por la precisión, cuanto más grande sea el intervalo de confianza alrededor del valor calculado, mayor será el peso. . . . . **???**
3. Cuando se pondera por la precisión, si las razones de riesgos específicas para cada estrato son iguales, los pesos de los estratos también serán iguales. . . . . **???**
4. Normalmente, los grupos de datos desequilibrados poseen pesos más pequeños que los grupos de datos equilibrados. . . . . **???**
5. El valor de la razón de riesgos ajustada estará entre los dos valores de las razones de riesgos específicas para cada estrato. . . . . **???**

Fórmula general para una razón de riesgos ajustada basada en la precisión

Podemos expandir la **fórmula basada en la precisión** para la razón de riesgos ajustada para dos estratos a una fórmula general basada en la precisión para **G** estratos. La fórmula general contiene **G** ponderaciones. Cada ponderación ofrece la precisión estimada del logaritmo de la razón de riesgos calculada para su estrato. Podemos replantear esta fórmula mediante la notación estándar de sumatoria. Calculamos la ponderación multiplicada por el logaritmo natural de la razón de riesgos para cada estrato y luego los sumamos para todos los estratos. Después, dividimos entre la suma de las ponderaciones de todos los estratos y exponenciamos el resultado:

Fórmula general basada en la precisión

$$\hat{aRR} = \exp\left[\frac{\sum w_i \ln(RR_i)}{\sum w_i}\right]$$

Considere datos con cuatro estratos; a continuación se brindan los cálculos de las cuatro razones de riesgos, sus correspondientes ponderaciones basadas en la precisión, los tamaños muestrales del cuadro, el valor de la razón de riesgos cruda y el tamaño muestral correspondiente. No se preocupe por cómo encontrar estos números, especialmente las ponderaciones, ya que una computadora lo puede hacer fácilmente. Observe que el orden de los pesos no corresponde directamente al orden de los tamaños muestrales específicos para cada estrato:

Ejemplo con 4 estratos

Estrato	1	2	3	4	cruda
$\hat{RR}$	2.01	1.49	1.88	1.54	2.45
ponderaciones	1.075	2.4974	6.9473	4.4856	----
n	282	76	161	90	609

Preguntas de estudio (P14.10)

1. En el cuadro de resumen, el mayor peso se asigna al estrato 3, que posee el segundo tamaño muestral más grande, mientras que el menor peso se asigna al estrato 1, que posee el tamaño muestral más grande. ¿Cómo puede suceder esto?

Utilizamos estas fórmulas del promedio ponderado para calcular la razón de riesgos ajustada mediante la sustitución en la fórmula de los valores por las **w** y los valores de la razón de riesgos específicos para cada estrato del cuadro. La razón de riesgos ajustada resultante es igual a 1,71. Este valor es un poco menor que la razón de riesgos cruda de 2,45. Por lo tanto, hay algo de confusión debido a la edad. Para resumir el efecto de la exposición en la enfermedad, debería utilizarse la razón de riesgos ajustada, en lugar de la razón de riesgos cruda.

$$\hat{aRR} = \exp\left[\frac{w_1 x \ln(\hat{RR}_1) + w_2 x \ln(\hat{RR}_2) + w_3 x \ln(\hat{RR}_3) + w_4 x \ln(\hat{RR}_4)}{w_1 + w_2 + w_3 + w_4}\right]$$

$$= \exp\left[\frac{1.075 x \ln(2.01) + 2.4974 x \ln(1.49) + 6.9473 x \ln(1.88) + 4.4856 x \ln(1.54)}{1.075 + 2.4974 + 6.9473 + 4.4856}\right]$$

$$= 1.71$$

Resumen

- Cuando hay **G** estratos, la fórmula basada en la precisión para una razón de riesgos ajustada es un promedio ponderado de los logaritmos naturales de los valores de la razón de riesgos para los **G** estratos.
- Las ponderaciones son estimaciones de la precisión de los logaritmos de los valores de la razón de riesgos específicos para cada estrato.
- La fórmula debería calcularse mediante un programa informático.

La razón de tasas ajustada basada en la precisión

Para los datos estratificados de persona-tiempo, la medida de efecto de interés normal es la razón de tasas o la RDI (es decir, la razón de densidades de incidencia). La siguiente fórmula se utiliza para obtener una razón de tasas ajustada basada en la precisión (es decir, **aRDI**).

$$a\hat{RDI} = \frac{\sum_{i=1}^{k} w_i \ln(\hat{RDI}_i)}{\sum_{i=1}^{k} w_i}$$

Las ponderaciones en esta fórmula son cálculos aproximados de la varianza inversa de la razón de tasas en cada estrato y están dadas por la expresión:

$$w_i = \frac{I_{1i} I_{0i}}{I_i}$$

donde $\mathbf{I_{1i}}$, $\mathbf{I_{0i}}$ e $\mathbf{I_i}$ para $\mathbf{i} = 1, 2, \ldots, \mathbf{k}$ denotan las frecuencias de enfermedad entre los expuestos, los no expuestos y la cantidad total de individuos, respectivamente, en el estudio.

14-4 Análisis estratificado (continuación)

***Odds ratio* ajustado basado en la precisión**

Hasta este momento solo hemos considerado datos estratificados de estudios de cohorte, donde la medida de efecto ajustada es una **razón de riesgos basada en la precisión** (**aRR**). Ahora describiremos cómo calcular un ***odds ratio* ajustado basado en la precisión** (**aOR**) para datos estratificados de estudios de casos y controles, transversales o de cohorte en donde el *odds ratio* es la medida de efecto de interés. Considere una disposición general de doble entrada de los datos para uno de varios estratos. El *odds ratio* ajustado basado en la precisión es el exponencial de un promedio ponderado del logaritmo natural de los *odds ratios* específicos para cada estrato. Esta fórmula se aplica ya sea que el diseño del estudio utilizado requiera un ***odds ratio* de riesgo**, un ***odds ratio* de exposición** o un ***odds ratio* de prevalencia**:

estrato i	Exposición: Sí	No	Total
D	a_i	b_i	m_{1i}
no D	c_i	d_i	m_{0i}
Total	n_{1i}	n_{0i}	n_i

aOR basada en la precisión

$$\hat{aOR} = \exp\left[\frac{\sum w_i \times \ln(\hat{OR}_i)}{\sum w_i}\right]$$

A continuación encontrará la fórmula para las ponderaciones, que no es la misma para el *odds ratio* ajustado y para la razón de riesgos ajustada, porque la precisión se calcula de manera distinta para medidas de efecto diferentes.

$$w_i = \frac{1}{\frac{1}{a_i} + \frac{1}{b_i} + \frac{1}{c_i} + \frac{1}{d_i}}$$

Concentrémonos en las ponderaciones para el *odds ratio* ajustado. Si todas las frecuencias de celda para un estrato dado son razonablemente grandes, entonces el denominador será pequeño; por lo tanto, su recíproca, que da la precisión, será relativamente grande.

	Exposición		
estrato i	Sí	No	Total
D	100	100	m_{1i}
no D	100	100	m_{0i}
Total	n_{1i}	n_{0i}	400

aOR basada en la precisión

$$\hat{aOR} = \exp\left[\frac{\sum w_i \times \ln(\hat{OR}_i)}{\sum w_i}\right]$$

$$w_i = \frac{1}{\frac{1}{100} + \frac{1}{100} + \frac{1}{100} + \frac{1}{100}} = 25$$

Por otro lado, si una de las frecuencias de celda es muy pequeña, la precisión tenderá a ser relativamente pequeña. Esto explica, a lo sumo para el *odds ratio*, el motivo por el cual un estrato desequilibrado con, al menos, una frecuencia de celda pequeña probablemente dé un peso relativamente pequeño, incluso si el tamaño total del estrato es grande.

	Exposición		
estrato i	Sí	No	Total
D	1	199	m_{1i}
no D	100	100	m_{0i}
Total	n_{1i}	n_{0i}	400

aOR basada en la precisión

$$\hat{aOR} = \exp\left[\frac{\sum w_i \times \ln(\hat{OR}_i)}{\sum w_i}\right]$$

$$w_i = \frac{1}{\frac{1}{1} + \frac{1}{199} + \frac{1}{100} + \frac{1}{100}} = 0.98$$

Resumen

- La medida de asociación utilizada para evaluar la confusión dependerá del diseño del estudio.
- La fórmula para la razón de riesgos ajustada se aplica cuando el diseño del estudio utilizado requiere una razón de prevalencias.
- La fórmula para el *odds ratio* ajustado se aplica ya sea que el diseño del estudio utilizado requiera un *odds ratio* de riesgo, un *odds ratio* de exposición o un *odds ratio* de prevalencia.
- Las ponderaciones se calculan de manera distinta para el *odds ratio* ajustado con respecto a la razón de riesgos ajustada.
- Un estrato desequilibrado con, al menos, una frecuencia de celda pequeña probablemente dé un peso relativamente pequeño, incluso si el tamaño total del estrato es grande.

Cálculo del OR ajustado: ejemplo

Estos cuadros muestran los resultados de un estudio de casos y controles que evalúa la relación entre el consumo de alcohol y el cáncer bucal. Los cuadros brindan los datos crudos cuando se ignora la edad y los datos estratificados cuando se ha categorizado la edad en tres grupos.

Un ejemplo: estudio de casos y controles

Consumo de alcohol → cáncer bucal

Crudo:

	Alc	no Alc	
CB	27	47	n = 607
no CB	90	443	

Estratificado:

	Edades 40-49		Edades 50-59		Edades 60+	
	Alc	no Alc	Alc	no Alc	Alc	no Alc
CB	4	25	12	10	11	12
no CB	22	309	37	67	31	67
	n = 360		n = 126		n = 121	

A continuación se presenta la fórmula para el *odds ratio* ajustado correspondiente a estos datos. Ahora, dado que las ponderaciones para cada estrato son las siguientes:

$w_1 = 2.95$
$w_2 = 4.44$
$w_3 = 4.52$

$\hat{OR} =$ **2.8 (crudo)**

$\hat{OR} =$ **2.2, 2.2, 2.0 (específico para un estrato)**

$$a\hat{OR} = \exp\left[\frac{w_1 \ln(\hat{OR}_1) + w_2 \ln(\hat{OR}_2) + w_3 \ln(\hat{OR}_3)}{w_1 + w_2 + w_3}\right]$$

1. ¿Cuál es el valor del *odds ratio* ajustado para estos datos? (Sugerencia: trate de responder sin cálculos).

2,1	2,3	1,9

El *odds ratio* calculado es 2,1.

2. ¿Los valores crudo y ajustado son relevantemente distintos?

sí	no

Sí, la razón cruda de 2,8 indica que hay casi el triple de exceso de riesgo, pero el valor ajustado de 2,1 indica aproximadamente el doble de exceso de riesgo.

3. ¿Qué valor es más adecuado, el crudo o el ajustado?

crudo	ajustado

El valor ajustado es más adecuado, porque es relevantemente distinto del valor crudo y controla la confusión causada por la edad.

Auto-evaluación (P14.11)

Los siguientes datos provienen de una encuesta transversal de seroprevalencia del VIH entre las prostitutas en relación con el uso de drogas intravenosas. El *odds ratio* crudo de prevalencia es 3,59. (Quizás desee utilizar una calculadora para responder a las siguientes preguntas).

	Negros o Hispanos		Blancos	
	Drogas intravenosas	No drogas intravenosas	Drogas intravenosas	No drogas intravenosas
VIH +	31	12	16	3
VIH -	93	144	141	124

1. ¿Cuál es el valor del POR para el grupo de negros o hispanos? . . . . **???**
2. ¿Cuál es el valor del POR para los blancos? . . . . . **???**
3. ¿Qué cuadro cree que es más equilibrado y, por lo tanto, brindará el mayor peso basado en la precisión? . . . . . **???**

Opciones
3,25 3,59 4,00 4,31 4,69 negros o hispanos blancos

En el estudio descrito en la pregunta anterior, el POR calculado para el grupo de negros o hispanos fue de 4,00 y para los blancos fue de 4,69. La ponderación basada en la precisión para el grupo de blancos o hispanos es de 7,503 y para los blancos es de 2,433.

4. Mediante la fórmula que se brinda a continuación, calcule el POR ajustado para este estudio. **???**
5. Recuerde que el POR crudo fue 3,59. ¿Hay evidencia de confusión? . . . **???**

$$\hat{aPOR} = \exp\left[\frac{\sum w_i \times \ln(\hat{OR}_i)}{\sum w_i}\right]$$

Opciones
1,00 4,16 4,35 4,97 discutible no sí

14-5 Análisis estratificado (continuación)

Cálculos ajustados de Mantel-Haenszel

El problema de las celdas con cero

Consideremos el siguiente grupo de resultados que se podrían haber presentado en un estudio de casos y controles que relaciona una variable de exposición **E** con una enfermedad **D**:

Resultados hipotéticos: estudio de casos y controles

E → D

Datos crudos	E	no E	Total
D	10	11	21
No D	1	14	15
Total	11	25	36

$\hat{OR} = 12.7$

Estrato 1	E	no E	Total
D	5	6	11
No D	0	5	5
Total	5	11	16

$\hat{OR}$ = indefinido

Estrato 2	E	no E	Total
D	5	5	10
No D	1	9	10
Total	6	14	20

$\hat{OR} = 9.0$

El *odds ratio* para el estrato 1 es indefinido debido a la frecuencia de celda cero en este estrato. Por lo tanto, no podemos determinar si hay confusión o interacción dentro de estos datos. Un enfoque que a veces se adopta para resolver dicho problema es sumar un número pequeño, generalmente 0,5, a cada celda de cualquier cuadro que posea una celda con cero. Si agregamos 0,5 aquí, el *odds ratio* para este cuadro modificado es 9,3.

Estrato 1	E	no E	Total
D	5.5	6.5	12.0
No D	.5	5.5	6.0
Total	6.0	12.0	18.0

$\hat{OR}$ = no definido

Estrato 2	E	no E	Total
D	5	5	10
No D	1	9	10
Total	6	14	20

$\hat{OR} = 9.3$ ← ¿Interacción? **NO** → $\hat{OR} = 9.0$

Ahora parece que hay pocas pruebas de interacción, ya que los *odds ratios* específicos para los dos estratos son muy cercanos. Pero, hay evidencia de cierto grado de confusión, ya que el *odds ratio* crudo es un poco más alto que cualquiera de los dos *odds ratios* específicos para cada estrato. Aunque este enfoque del problema de las celdas con cero es razonable, y generalmente se utiliza, podría preocuparnos que la elección de 0,5 sea arbitraria.

Preguntas de estudio (P14.12)

Estrato 1	**E**	**no E**	**Total**
D	**5**	**6**	**11**
No D	**0**	**5**	**5**
Total	**5**	**11**	**16**

1. Dados los valores a = 5, b = 6, c = 0 y d = 5 para el estrato 1, ¿cuál es el *odds ratio* calculado si se suma 0,1 a cada frecuencia de celda?
2. Dados los valores a = 5, b = 6, c = 0 y d = 5 para el estrato 1, ¿cuál es el *odds ratio* calculado si se suma 1,0 a cada frecuencia de celda?
3. ¿Qué indican sus respuestas a las preguntas anteriores sobre la práctica de sumar un número pequeño a todas las celdas en un estrato que contenga una frecuencia de celda cero?

También podríamos tener interés en calcular un *odds ratio* ajustado basado en la precisión que involucre un *odds ratio* específico para un estrato que se haya modificado en forma cuestionable.

Preguntas de estudio (P14.12) (continuación)

Los *odds ratios* específicos para cada estrato obtenidos cuando se suma 0,5 al estrato 1 son iguales a 9,31 para el estrato 1 y 9,00 para el estrato 2.

4. Si se calcula un *odds ratio* ajustado basado en la precisión sumando 0,5 a cada frecuencia de celda en el estrato 1, las ponderaciones son 0,3972 para el estrato 1 y 0,6618 para el estrato 2. ¿Qué valor obtiene para el aOR?

Los *odds ratios* específicos para cada estrato obtenidos cuando se suma 0,1 al estrato 1 son iguales a 41,64 para el estrato 1 y 9,00 para el estrato 2.

5. Si se calcula un *odds ratio* ajustado basado en la precisión sumando 0,1 a cada frecuencia de celda en el estrato 1, las ponderaciones son 0,0947 para el estrato 1 y 0,6618 para el estrato 2. ¿Qué valor obtiene para el aOR?
6. ¿Cómo se comparan sus respuestas a las dos preguntas anteriores?
7. ¿Qué indican sus respuestas a las preguntas anteriores sobre el uso de un *odds ratio* ajustado basado en la precisión?

Afortunadamente, hay otra opción de cálculo ajustado para abordar el problema de las celdas con cero que se denomina ***odds ratio* de Mantel-Haenszel**; esta alternativa se describe en la próxima actividad.

Resumen

- Cuando hay datos escasos en algunos estratos, especialmente celdas con cero, los *odds ratios* específicos para cada estrato dejan de ser confiables y posiblemente se tornen indefinidos.
- Cuando hay frecuencias de celda cero en un estrato, un enfoque que se puede adoptar es sumar un número pequeño, generalmente 0,5, a cada frecuencia de celda en el estrato.
- Una de las desventajas de este último enfoque es que el cálculo resultante de efecto modificado específico para cada estrato puede cambiar radicalmente según el número pequeño que se suma (por ejemplo, 0,5; 0,1).
- El uso de un cálculo ajustado basado en la precisión en dicha situación se vuelve entonces problemático.
- Afortunadamente, hay un enfoque alternativo para el problema de las celdas con cero, que implica utilizar el denominado cálculo ajustado de Mantel-Haenszel.

El *odds ratio* de Mantel-Haenszel

El **cálculo ajustado de Mantel-Haenszel** que se emplea en los estudios de casos y controles se denomina ***odds ratio* de Mantel-Haenszel**, o **mOR**. Su fórmula es la siguiente:

Odds ratio de Mantel-Haenszel **(mOR)**

$$mOR = \frac{\sum_{i=1}^{G} \frac{a_i d_i}{n_i}}{\sum_{i=1}^{G} \frac{b_i c_i}{n_i}}$$

G = # de estratos

Esta fórmula también se puede usar para calcular un *odds ratio* ajustado en estudios transversales y de cohorte. Una característica clave del **mOR** es que se puede recurrir a él sin tener que modificar ningún estrato que contenga una frecuencia de celda cero. Por ejemplo, si hay una frecuencia de celda cero, como se muestra a continuación para la celda **c**, entonces el cálculo simplemente incluye un cero en una suma en el denominador del **mOR**, pero la suma total no necesariamente será cero.

Odds ratio de Mantel-Haenszel **(mOR)**

$$mOR = \frac{\sum_{i=1}^{G} \frac{a_i d_i}{n_i}}{\sum_{i=1}^{G-1} \frac{b_i c_i}{n_i} + 0}$$

G = # de estratos

	E	$\overline{E}$	
D	a_G	b_G	m_{1G}
$\overline{D}$	$c_G = 0$	d_G	m_{0G}
	n_{1G}	n_{0G}	n_G

Preguntas de estudio (P14.13)

1. Suponga que hay G = 5 estratos y que la celda *b* o *c* tienen cero en cada una y en todos los estratos. ¿Qué sucederá si calcula el mOR para estos datos?
2. Suponga que hay G = 5 estratos y que la celda *a* o *d* tienen cero en cada una y en todos los estratos. ¿Qué sucederá si calcula el mOR para estos datos?
3. ¿Qué indican sus respuestas a las preguntas anteriores con respecto al uso del mOR cuando hay frecuencias de celda cero?

Otra buena característica del **mOR** es que, aunque no parezca, puede escribirse como un promedio ponderado de *odds ratios* específicos para cada estrato, siempre que no haya ninguna celda con cero en ningún estrato. Por lo tanto, el **mOR** brindará un valor que estará entre los *odds ratios* específicos mínimo y máximo de todos los estratos, al igual que cualquier promedio ponderado.

$$mOR = \frac{\sum_{i=1}^{G} \frac{a_i d_i}{n_i}}{\sum_{i=1}^{G} \frac{b_i c_i}{n_i}} = \frac{\sum_{i=1}^{G} \frac{b_i c_i}{n_i} \hat{OR}_i}{\sum_{i=1}^{G} \frac{b_i c_i}{n_i}}$$

* Puede escribirse como un promedio ponderado siempre que no haya ninguna celda con cero

Sin embargo, otra particularidad del **mOR** es que es igual a 1 solo cuando el estadístico chi cuadrado de Mantel-Haenszel es igual a cero. Es posible que el aOR basado en la precisión sea diferente de 1, aunque el estadístico chi cuadrado de Mantel-Haenszel sea exactamente igual a cero. En general, se ha demostrado que el **mOR** tiene buenas propiedades estadísticas, en particular, cuando se utiliza con datos de casos y controles equiparados. Por lo tanto, con frecuencia se emplea en lugar del **aOR** basado en la precisión, incluso cuando no hay celdas con cero o los datos no son escasos.

Ahora aplicamos la fórmula del mOR al ejemplo de datos estratificados que hemos considerado anteriormente. Al sustituir las frecuencias de celda de cada estrato en la fórmula del **mOR**, el **mOR** calculado resulta ser igual a 15,25:

Datos crudos	E	no E	Total
D	10	11	21
No D	1	14	15
Total	11	25	36

Datos de estudio de casos y controles

$\hat{OR} = 12.7$

Estrato 1	E	no E	Total
D	5	6	11
No D	0	5	5
Total	5	11	16

$\hat{OR}$ = no definido

Estrato 2	E	no E	Total
D	5	5	10
No D	1	9	10
Total	6	14	20

$\hat{OR} = 9.0$

$$\hat{mOR} = \frac{\frac{5\times5}{16} + \frac{5\times9}{20}}{\frac{6\times0}{16} + \frac{5\times1}{20}} = 15.25$$

Este valor ajustado es un poco mayor que el *odds ratio* crudo de 12,7 y mucho más alto que el *odds ratio* de 9,0 en el estrato 2. Como el estrato 1 tiene un *odds ratio* indefinido, no podemos determinar si hay evidencia de interacción. Sin embargo, debido a que los *odds ratios* crudo y ajustado difieren, hay evidencia de confusión.

Preguntas de estudio (P14.13) (continuación)

Si se calcula un *odds ratio* ajustado basado en la precisión sumando 0,5 a cada celda en el estrato 1, el aOR que se obtiene es 9,11. Si, en cambio, se suma 0,1 a cada frecuencia de celda en el estrato 1, el aOR es 10,65.

4. Compare los resultados del aOR anterior con el mOR de 15,25 que se obtuvo anteriormente. ¿Qué valor prefiere?

Resumen

- Para los estudios de casos y controles, así como para otros estudios que involucran un *odds ratio*, el *odds ratio* de Mantel-Haenszel (mOR) es una alternativa al *odds ratio* ajustado basado en la precisión.
- En concordancia con el mOR, la mRR o la mRDI pueden utilizarse en los estudios de cohorte.
- Una característica clave del mOR es que se puede emplear sin modificación cuando haya celdas con cero en alguno de los estratos.
- El mOR también se puede plantear como un promedio ponderado de los *odds ratios* específicos para cada estrato **siempre que** no haya celdas con cero en ningún estrato.
- El mOR es igual a 1 sólo cuando el estadístico chi cuadrado de Mantel-Haenszel es igual a cero.
- Se ha demostrado que el mOR tiene buenas propiedades estadísticas, en particular, cuando se utiliza con datos de casos y controles equiparados.

Fórmulas para la razón de riesgos de Mantel-Haenszel y la razón de tasas de Mantel-Haenszel

La razón de riesgos de Mantel-Haenszel (es decir, mRR) está representada por la siguiente fórmula:

$$\hat{mRR} = \frac{\sum_{i=1}^{k} \frac{a_i n_{0i}}{n_i}}{\sum_{i=1}^{k} \frac{b_i n_{1i}}{n_i}}$$

donde las cantidades en la fórmula provienen de las frecuencias de celda en el cuadro de 2 x 2 del estrato *i*-avo para un estudio de cohorte con incidencia acumulada:

Disposición de los datos para el estrato i

	E	**no E**	
D	a_i	b_i	m_{1i}
no D	c_i	d_i	m_{0i}
	n_{1i}	n_{0i}	n_i

La razón de tasas de Mantel-Haenszel (es decir, mRDI) es:

$$\hat{mIDR} = \frac{\sum_{i=1}^{k} \frac{I_{1i} PT_{0i}}{PT_i}}{\sum_{i=1}^{k} \frac{I_{0i} PT_{1i}}{PT_i}}$$

donde las cantidades en la fórmula provienen de las frecuencias de celda en el cuadro de 2 x 2 del estrato *i*-avo para un estudio con datos de persona-tiempo:

Estudio de cohorte con persona-tiempo
Disposición de los datos para el estrato i

	E	**no E**	
D	I_{1i}	I_{0i}	I_i
PT	PT_{1i}	PT_{0i}	PT_i



Auto-evaluación(P14.14)

Verdadero o **falso**

1. La práctica de sumar un pequeño número a cada una de las celdas de un cuadro de 2 x 2 para eliminar una frecuencia de celda cero es arbitraria y debería emplearse con precaución. . . **???**
2. Lo que más se afecta cuando se suma un número pequeño a cada celda es el cálculo ajustado, antes que los cálculos específicos para cada estrato. . . . . **???**
3. Cuando hay frecuencias de celda cero debería preferirse el uso de cálculos ajustados de Mantel-Haenszel, ya que generalmente pueden calcularse sin realizar ningún ajuste. . . **???**

Estimación de intervalos

Estimación de intervalos - Introducción

Ahora describimos cómo obtener un intervalo de confianza de muestra grande alrededor de un valor ajustado obtenido en un análisis estratificado. Esta estimación del intervalo puede adoptar una de las dos formas que se muestran a continuación:

Intervalo de confianza de muestra grande alrededor de un valor ajustado

Medidas de diferencia: $\hat{\theta} \pm Z_p \sqrt{\hat{\mathrm{Var}}(\hat{\theta})}$

Medidas de razones: $\hat{\theta} \exp\left[\pm Z_p \sqrt{\hat{\mathrm{Var}}(\ln \hat{\theta})}\right]$

La **Z** en cada expresión denota un punto porcentual de la distribución normal estándar. La **θ** (“theta”) en cada expresión denota la medida de efecto de interés. Puede ser una medida en forma de diferencia como, por ejemplo, una diferencia de riesgos, o una medida en forma de razón como, por ejemplo, una razón de riesgos. Normalmente, **θ** es un promedio ponderado de efectos específicos de cada estrato. En especial, para las **medidas en forma de diferencias de riesgos**, **θ** tendrá una ponderación lineal, como se muestra aquí:

Ponderación lineal:

$$\hat{\mathbf{a}}\hat{\mathrm{RD}} = \frac{\sum \mathbf{w}_i \hat{\mathrm{RD}}_i}{\sum \mathbf{w}_i}$$

donde $\hat{\mathrm{DR}}$ es la diferencia de riesgos calculada en estrato i

Los **cálculos ajustados de Mantel-Haenszel para las medidas de efecto en forma de razones** también tienen una ponderación lineal.

Ponderación lineal:

$$\hat{\mathbf{m}}\hat{\mathrm{OR}} = \frac{\sum \mathbf{w}_i \hat{\mathrm{OR}}_i}{\sum \mathbf{w}_i}$$

dado que $\hat{\mathrm{OR}}_i$ esté definido para cada estrato

Para los **cálculos en forma de razones basados en la precisión**, **θ** tendrá una ponderación logarítmica lineal.

Ponderación logarítmica lineal:

$$\theta = \exp\left[\frac{\sum w_i \ln(\theta_i)}{\sum w_i}\right] \quad \text{e.g.,} \quad \hat{aOR} = \exp\left[\frac{\sum w_i \ln(\hat{OR}_i)}{\sum w_i}\right]$$

El componente de la varianza dentro del intervalo de confianza adoptará una forma matemática específica según la medida de efecto. Para las **medidas basadas en la precisión**, el componente de la varianza se simplifica en forma práctica a expresiones que involucran la suma de ponderaciones. En especial, para las **medidas basadas en la precisión en forma de diferencias**, la fórmula del intervalo de confianza se reduce a la forma que se muestra a continuación:

$$\hat{\theta} \pm \frac{1.96}{\sqrt{\sum w_i}}$$

Para las **medidas basadas en la precisión en forma de razones**, la fórmula es la siguiente:

$$\hat{\theta}\, e^{\pm \frac{1.96}{\sqrt{\sum w_i}}}$$

Como ejemplo, nuevamente consideramos datos de cohorte que involucran cuatro estratos. Tenemos cuatro razones de riesgos, sus correspondientes ponderaciones basadas en la precisión y tamaños muestrales, el valor de la razón de riesgos y el respectivo tamaño muestral.

Ejemplo con 4 estratos

Estrato	1	2	3	4	crudo
$\hat{RR}$	2.01	1.49	1.88	1.54	2.45
w_i	1.0751	2.4974	6.9473	4.4856	----
n	282	76	161	90	609

$$\hat{aRR} = 1.71$$

La razón de riesgos ajustada resulta ser 1,71. Para obtener la estimación del intervalo de confianza de 95% basado en la precisión para esta razón de riesgos ajustada, comenzamos con la fórmula que se indica a continuación:

95% IC:

$$aRR\, e^{\pm \frac{1.96}{\sqrt{(w_1 + w_2 + w_3 + w_4)}}}$$

Luego sustituimos $\hat{aRR}$ en la fórmula por 1,71 y los valores que se muestran en el cuadro para las cuatro ponderaciones:

$$95\%\ \text{IC:}\quad 1.71\ e^{\pm \frac{1.96}{\sqrt{(1.0751+2.4974+6.9473+4.4856)}}}$$

Entonces, los límites inferior y superior del intervalo de confianza resultan ser 1,03 y 2,84, respectivamente.

$$95\%\ \text{IC:}\quad 1.71\ e^{\pm \frac{1.96}{\sqrt{15.0054}}} = (1.03, 2.84)$$

Preguntas de estudio (P14.15)

1. ¿Cómo interpretaría el intervalo de confianza anterior?

Ejemplo de diferencia de riesgos con 4 estratos

Estrato	1	2	3	4	crudo
$\hat{RD}$	0.0630	0.0578	0.1078	0.0851	0.1310
w_i	71.9424	96.8891	183.9742	137.4186	----
n	282	76	161	90	609

$$\hat{aRD} = 0.0850 \qquad \hat{aRD} \pm \frac{1.96}{\sqrt{\sum w_i}}$$

2. Sobre la base de la información anterior, calcule un intervalo de confianza de 95% para la diferencia de riesgos ajustada. (Deberá utilizar una calculadora para esta operación).
3. ¿Cómo interpreta este intervalo de confianza?

Resumen

- Una estimación de intervalo de muestra grande alrededor de un cálculo ajustado puede adoptar una de las dos formas siguientes:

  $\hat{\theta} \pm Z_p \sqrt{\hat{Var}(\hat{\theta})}$ para medidas de efecto en forma de diferencias, y

  $\hat{\theta} \exp\left[\pm Z_p \sqrt{\hat{Var}(\ln\hat{\theta})} \right]$ para medidas de efecto en forma de razones.
- Para las medidas en forma de diferencias de riesgos y para los cálculos de Mantel-Haenszel, **θ** tendrá una ponderación lineal.
- Para los cálculos en forma de razones basados en la precisión, **θ** tendrá una ponderación logarítmica lineal.
- Para las medidas basadas en la precisión, el componente de la varianza implica la suma de ponderaciones,

según se indica a continuación:

$$\hat{\theta} \pm \frac{Z_p}{\sqrt{\sum w_i}}$$ para medidas en forma de diferencias

$$\hat{\theta} \exp\left[\pm \frac{Z_p}{\sqrt{\sum w_i}} \right]$$ para medidas en forma de razones

Estimación de intervalos para el mOR

Considere nuevamente los datos estratificados de casos y controles con estratos escasos que hemos utilizado con anterioridad para calcular un ***odds ratio* de Mantel-Haenszel**. Para estos datos, el *odds ratio* de Mantel-Haenszel es igual a 15,25.

Datos estratificados de casos y controles con estratos escasos

Datos de casos y controles

Datos crudos

	E	no E	Total
D	10	11	21
no D	1	14	15
Total	11	25	36

$\hat{OR} = 12.7$

Estrato 1

	E	no E	Total
D	5	6	11
no D	0	5	5
Total	5	11	16

$\hat{OR} =$ no definido

Estrato 2

	E	no E	Total
D	5	5	10
no D	1	9	10
Total	6	14	20

$\hat{OR} = 9.0$

$\hat{mOR} = 15.25$

Encontramos un intervalo de confianza de 95% alrededor de este valor con la fórmula que se muestra a continuación:

95% intervalo de confianza

$$\hat{mOR} \exp\left[\pm 1.96 \sqrt{(\hat{Var}(\ln \hat{mOR}))} \right]$$

El término de la varianza en esta fórmula es una expresión compleja que involucra las frecuencias de cada estrato. Aquí presentamos la fórmula de la varianza principalmente para mostrarle qué tan compleja es. Utilice un programa informático para realizar los cálculos reales (aquí nosotros los realizaremos por usted).

$$\hat{\mathrm{Var}}(\ln \hat{\mathrm{mOR}}) = A + B + C$$

$$A = \frac{[\sum(P_i R_i)]}{2[\sum(R_i)]^2} \qquad P_i = \frac{(a_i + d_i)}{n_i}$$

$$B = \frac{[\sum(P_i S_i + Q_i R_i)]}{2[\sum(R_i)\sum(S_i)]} \qquad Q_i = \frac{(b_i + c_i)}{n_i}$$

$$C = \frac{[\sum(Q_i S_i)]}{2[\sum(S_i)]^2} \qquad R_i = \frac{a_i d_i}{n_i}$$

$$S_i = \frac{b_i c_i}{n_i}$$

Al sustituir $\mathbf{P_i}$, $\mathbf{Q_i}$, $\mathbf{R_i}$ y $\mathbf{S_i}$ en la fórmula por las frecuencias de cada estrato, obtenemos los valores que se incluyen a continuación. Aquí verá la varianza calculada:

$$\hat{\mathrm{Var}}(\ln \hat{\mathrm{mOR}}) = 0.0878 + 0.7533 + 0.6000 = 1.4411$$

$$A = \frac{[\sum(P_i R_i)]}{2[\sum(R_i)]^2} = \frac{2.5516}{2(3.8125)^2} = .0878$$

$$B = \frac{[\sum(P_i S_i + Q_i R_i)]}{2[\sum(R_i)\sum(S_i)]} = \frac{1.4359}{2(3.8125)(.2500)} = .7533$$

$$C = \frac{[\sum(Q_i S_i)]}{2[\sum(S_i)]^2} = \frac{.0750}{2(.2500)^2} = .6000$$

El intervalo de confianza de 95% se muestra a continuación. Aunque este intervalo no contiene el valor nulo de 1, es, sin embargo, extremadamente amplio, lo que no debería sorprender dados los estratos escasos.

95% intervalo de confianza $\hat{\mathrm{mOR}} = 15.25$

$$= \hat{\mathrm{mOR}} \exp\left[\pm 1.96 \sqrt{\hat{\mathrm{Var}}(\ln \hat{\mathrm{mOR}})}\right]$$

$$= 15.25 \exp\left[\pm 1.96 \sqrt{1.4411}\right]$$

$$= 15.25 \exp\left[\pm 2.3529\right]$$

$$= (1.4502, 160.3693)$$

Resumen

- Un intervalo de confianza de 95% alrededor de un *odds ratio* de Mantel-Haenszel es la fórmula:

$$\hat{\text{mOR}} \exp\left[\pm 1{,}96\sqrt{\text{Vâr}(\ln \hat{\text{mOR}})}\right]$$

- El término de la varianza en esta fórmula es una expresión compleja que involucra las frecuencias de cada estrato.
- Debería utilizar una computadora para calcular el término de la varianza en la fórmula.
- No debería sorprenderle encontrar un intervalo de confianza muy amplio si todos los estratos son escasos.

Estimación de intervalos de muestra grande para la razón de riesgos de Mantel-Haenszel y la razón de tasas de Mantel-Haenszel

Un intervalo de confianza de 95% de muestra grande para la razón de riesgos de Mantel-Haenszel (es decir, mRR) es la siguiente fórmula:

$$\hat{\text{mRR}} \exp\left[\pm 1{,}96\sqrt{\hat{\text{Var}}(\ln \hat{\text{mRR}})}\right]$$

donde

$$\hat{\text{mRR}} = \frac{\sum_{i=1}^{k} \frac{a_i n_{0i}}{n_i}}{\sum_{i=1}^{k} \frac{b_i n_{1i}}{n_i}}$$

y la expresión de la varianza en la fórmula es:

$$\hat{\text{Var}}(\ln \hat{\text{mRR}}) = \frac{\sum_{i=1}^{k} \left(\frac{m_{1i} n_{1i} n_{0i}}{n_i^2} - \frac{a_i b_i}{n_i} \right)}{\left(\sum_{i=1}^{k} \frac{a_i n_{0i}}{n_i} \right) \left(\sum_{i=1}^{k} \frac{b_i n_{1i}}{n_i} \right)}$$

Las cantidades en la fórmula de la varianza provienen de las frecuencias de celda en el cuadro de 2 x 2 del estrato *i*-avo para un estudio de cohorte con incidencia acumulada:

Disposición de los datos para el estrato *i*

	E	**no E**	
D	a_i	b_i	m_{1i}
no D	c_i	d_i	m_{0i}
	n_{1i}	n_{0i}	n_i

Un intervalo de confianza de 95% de muestra grande para la razón de tasas de Mantel-Haenszel (es decir, mRDI) es la siguiente fórmula:

$$\hat{\text{mRDI}} \exp\left[\pm 1{,}96\sqrt{\hat{\text{Var}}(\ln \hat{\text{mRDI}})}\right]$$

Estimación de intervalos de muestra grande para la razón de riesgos de Mantel-Haenszel y la razón de tasas de Mantel-Haenszel (continuación)

donde

$$\widehat{\text{mRDI}} = \frac{\sum_{i=1}^{k} \frac{I_{1i} PT_{0i}}{PT_i}}{\sum_{i=1}^{k} \frac{I_{0i} PT_{1i}}{PT_i}}$$

y la expresión de la varianza en la fórmula es:

$$\widehat{\text{Vâr}}(\ln \text{mIDR}) = \frac{\sum_{i=1}^{k} \frac{I_i PT_{1i} PT_{0i}}{PT_i^2}}{\left(\sum_{i=1}^{k} \frac{I_{1i} PT_{0i}}{PT_i} \right) \left(\sum_{i=1}^{k} \frac{I_{0i} PT_{1i}}{PT_i} \right)}$$

Las cantidades en la fórmula de la varianza provienen de las frecuencias de celda en el cuadro de 2 x 2 del estrato *i*-avo para un estudio de cohorte con datos de persona-tiempo:

Estudio de cohorte con persona-tiempo
Disposición de los datos para el estrato i

	E	**no E**	
D	I_{1i}	I_{0i}	I_i
PT	PT_{1i}	PT_{0i}	PT_i

Auto-evaluación (P14.16)

Se realizó un estudio de casos y controles para determinar la relación entre el tabaquismo y el cáncer de pulmón. Los datos se estratificaron en 3 categorías etarias. Los resultados fueron: aOR = 4,51 y la suma de las ponderaciones = 3,44.

1. Calcule un intervalo de confianza de 95% basado en la precisión para estos datos. . **???**
2. ¿Estos resultados brindan evidencia significativa de que el tabaquismo está relacionado con el cáncer de pulmón cuando se controla la edad? . . . . . **???**

Opciones
1,25; 10,78 **1,57; 12,98** **no** **sí**

$$\widehat{\theta}\exp\left[\frac{\pm 1.96}{\sqrt{\sum_{i=1}^{c} W_i}}\right]$$

14-6 Análisis estratificado (continuación)

Extensiones a más de 2 categorías de exposición

Cuadros de 2 x C

Tiene lugar una extensión natural del análisis estratificado para los cuadros de 2 x 2 cuando hay más de dos categorías de exposición. En dicho caso, la disposición básica de los datos tiene la forma de un cuadro de 2 x C, donde C denota la cantidad de categorías de exposición. Ahora brindamos una descripción general de cómo analizar cuadros de 2 x C estratificados.

Los cuadros que se muestran a continuación contienen el nivel de colesterol y la condición de enfermedad coronaria para 609 hombres blancos dentro de dos categorías etarias de un estudio de cohorte con una duración de 10 años realizado en el condado Evans, Georgia, desde 1960 a 1969.

Estudio de cohorte con una duración de 10 años realizado en el condado Evans, Georgia 1960-1969

Crudo

	COL-AL	COL-MED	COL-BA	
EC	22	40	9	
NO EC	132	267	139	
	154	307	148	609

EDAD ≥ 55

	COL-AL	COL-MED	COL-BA	
EC	13	25	5	
NO EC	48	106	54	
	61	131	59	251

EDAD < 55

	COL-AL	COL-MED	COL-BA	
EC	9	15	4	
NO EC	84	161	85	
	93	176	89	358

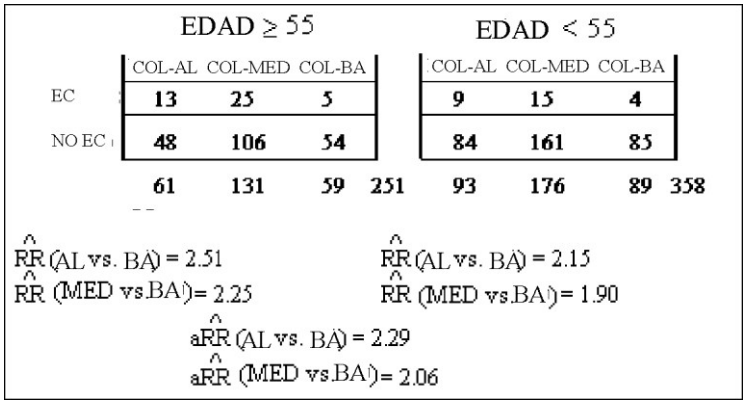
Estos datos muestran tres categorías de exposición, en lugar de dos, para cada uno de los dos estratos. ¿Cómo realizamos el análisis estratificado de dichos datos? Normalmente llevamos a cabo análisis específicos de cada estrato y una evaluación general, si corresponde, de la relación entre la exposición y la enfermedad en todos los estratos. Pero, como hay tres categorías de exposición, podemos preguntarnos cómo calcular la multiplicidad de posibles medidas de efecto para cada cuadro, cómo resumir dicha información para todos los estratos y cómo modificar el procedimiento de prueba de hipótesis para más de 2 categorías.

Normalmente, calculamos varias medidas de efecto, cada una de las cuales compara una de las categorías de exposición con otra empleada como referencia. En general, si hay **C** categorías de exposición, la disposición básica de los datos es un cuadro de **2 x C**. Entonces, el análisis típico genera C-1 valores ajustados, comparando **C-1** categorías de exposición con la categoría de referencia.

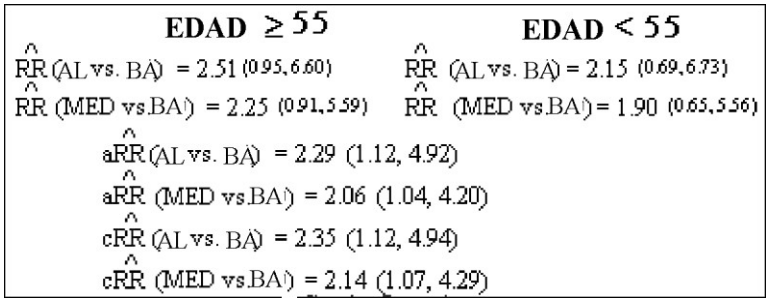
	E_1	E_2	E_3	…	E_c
D					
no D					

Cálculos ajustados:
$a\hat{R}R_1\ a\hat{R}R_2\ a\hat{R}R_3 \ldots a\hat{R}R_{c-1}$

En nuestro ejemplo estableceremos que el **colesterol bajo** será la categoría de referencia. Como tenemos datos de cohorte con incidencia acumulada, calculamos dos razones de riesgos por estrato, una comparando la categoría **colesterol alto** con la categoría **colesterol bajo** y la otra cotejando la categoría **colesterol moderado** con la categoría **colesterol bajo**. A continuación se brindan estos valores para cada grupo etario y los cálculos ajustados basados en la precisión para ambos grupos etarios. Estos resultados indican un leve efecto de dosis-respuesta del colesterol en el riesgo de desarrollar EC. Es decir, el efecto, medido por la razón de riesgos, disminuye a medida que el grupo índice cambia de colesterol alto a moderado, cuando se compara cada grupo con el grupo de colesterol bajo.



A continuación se muestran los intervalos de confianza de 95%, tanto para los *odds ratios* específicos de cada estrato como para los *odds ratios* ajustados. Los intervalos para las razones de riesgos ajustadas se obtienen mediante la fórmula descrita anteriormente, que involucra la suma de las ponderaciones.



Preguntas de estudio (P14.17)

1. Sobre la base de la información brindada con anterioridad, ¿hay evidencia de que la edad sea un modificador de efecto de la relación entre el nivel de colesterol y el riesgo de desarrollar EC?
2. En función de su respuesta a la pregunta anterior, ¿es adecuado realizar una evaluación general en este análisis estratificado?
3. ¿Hay evidencia de confusión debido a la edad?

Resumen

- Cuando hay más de dos categorías de exposición, la disposición básica de los datos tiene la forma de un cuadro de **2 x C**, donde **C** denota la cantidad de categorías de exposición.
- Al igual que con los cuadros estratificados de 2 x 2, el objetivo de la evaluación general es un cálculo ajustado general, una prueba de hipótesis general y una estimación de intervalo alrededor del valor ajustado.
- Cuando hay **C** categorías de exposición, el análisis típico produce **C-1** valores ajustados, que comparan **C-1** categorías de exposición con una de referencia.

Prueba de tendencia

Ahora describimos cómo probar hipótesis para datos estratificados con varias categorías de exposición. Nuevamente consideramos los datos del condado Evans que se indican a continuación, los cuales relacionan el nivel de colesterol con el desarrollo de enfermedad coronaria, estratificados por edad.

Prueba de hipótesis - varias categorías de exposición

Crudo

	COL-AL	COL-MED	COL-BA	
EC	22	40	9	
NO EC	132	267	139	
	154	307	148	609

EDAD ≥ 55

	COL-AL	COL-MED	COL-BA	
EC	13	25	5	
NO EC	48	106	54	
	61	131	59	251

EDAD < 55

	COL-AL	COL-MED	COL-BA	
EC	9	15	4	
NO EC	84	161	85	
	93	176	89	358

Riesgo .21 > .19 > .08 .10 > .09 > .04

e.g., 13/61 = .21

La variable de exposición, el nivel de colesterol, se clasificó en tres categorías ordenadas. Podemos observar que, para cada estrato, el riesgo de EC se reduce a medida que el nivel de colesterol disminuye.

Para ver si estos resultados específicos para cada estrato son estadísticamente significativos, debemos realizar una **prueba de tendencia**. Dicha prueba nos permite evaluar si existe una **relación de dosis-respuesta** significativa entre la variable de exposición y el resultado de salud. La prueba de tendencia puede realizarse mediante un procedimiento de **extensión de la prueba de Mantel-Haenszel**. Esta prueba requiere que se asigne un valor numérico o **puntaje** a cada categoría de exposición. Por ejemplo, se podrían conceder puntajes de 2 a las tres categorías ordenadas de exposición cuando el colesterol sea mayor que 233, 1 cuando el colesterol esté entre 184 y 233, y 0 si el colesterol fuera menor que 184.

$S_1 = 2$ si COL > 233
$S_2 = 1$ si COL 184 - 233
$S_3 = 0$ si COL < 184

De manera alternativa, los puntajes podrían determinarse por el valor medio del colesterol en cada una de las categorías ordenadas. Para estos datos, entonces, los puntajes que resultan son 265,0; 207,8 y 164,4 para las categorías de colesterol alto, moderado y bajo, respectivamente.

$S_1 = 265.0$ si COL > 233
$S_2 = 207.8$ si COL 184 - 233
$S_3 = 164.4$ si COL < 184

A continuación se incluye la fórmula general para el estadístico de prueba:

$$\chi^2_{MHT} = \frac{\left[\sum_{i=1}^{G} T_i - E(T_i)\right]^2}{\sum_{i=1}^{G} V(T_i)}$$

G = cantidad de estratos
T_i = puntaje total para el estrato i
$E(T_i)$ = puntaje total esperado para el estrato i
$V(T_i)$ = varianza del puntaje total para el estrato i

La cantidad de estratos es **G**. Normalmente, esta fórmula debería calcularse con una computadora. En nuestro ejemplo, hay dos estratos relacionados con la edad, por lo tanto, G = 2. La fórmula de la prueba de tendencia, entonces, se simplifica como se muestra a continuación (el recuadro al final de esta actividad indica cómo realizar los cálculos):

$$\chi^2 = \frac{\left[T_1 - E(T_1) + T_2 - E(T_2)\right]^2}{V(T_1) + V(T_2)}$$

Utilizaremos esta fórmula simplificada para calcular la prueba de tendencia donde usaremos los valores medios redondeados del colesterol para cada categoría de colesterol como nuestros puntajes. Estos son los resultados:

$$\chi^2 = \frac{\left[9465 - 9094.9283 + 6161 - 5932.3240\right]^2}{44088.6783 + 33914.7168} = 4.56 \quad (P = 0.0320)$$

	EDAD ≥ 55				EDAD < 55			
	COL-AL	COL-MOD	COL-BA		COL-AL	COL-MOD	COL-BA	
EC-Sí	13	25	5		9	15	4	
EC-No	48	106	54		84	161	85	
	61	131	59	251	93	176	89	358
Risk	.21	.19	.08		.10	.09	.04	

Valores medios
$S_1 = 265$
$S_2 = 208$
$S_3 = 164$

$T_1 = 9465$ $V(T_1) = 44088.6783$
$T_2 = 6161$ $V(T_2) = 33914.7168$
$E(T_1) = 9094.9283$
$E(T_2) = 5932.3240$

Preguntas de estudio (P14.18)

1. Brinde dos formas equivalentes de formular la hipótesis nula para la prueba de tendencia en este ejemplo.
2. En función de los resultados de la prueba de tendencia, ¿a qué conclusión llega?
3. Si se utiliza un método distinto de asignación de puntaje, ¿cuáles cree que serían los resultados de la prueba de tendencia?

Veamos cuáles serían los resultados chi cuadrado cuando empleamos un sistema de asignación de puntaje distinto. Aquí están los resultados cuando usamos 2, 1 y 0.

$S_1 = 265, S_2 = 208, S_3 = 164$:

$\chi^2 = 4.56$
(P = 0.0320)

$S_1 = 2, S_2 = 1, S_3 = 0$:

$\chi^2 = 5.04$
(P = .0248)

Preguntas de estudio (P14.18) (continuación)

4. ¿Cuál es la diferencia entre los resultados basados en los puntajes 2, 1 y 0 y los basados en 265, 208, 164?
5. ¿Cuáles serían los resultados de la prueba de tendencia que podría esperar si los puntajes fuesen 3, 2, 1, en lugar de 2, 1, 0?

Resumen

- Si la variable de exposición es ordinal, puede realizarse una prueba de tendencia lineal chi cuadrado con 1 grado de libertad mediante un procedimiento de extensión de la prueba de Mantel-Haenszel.
- Dicha prueba "de tendencia" puede utilizarse para evaluar si existe una relación de dosis-respuesta lineal entre la exposición y el riesgo de enfermedad.
- Para realizar esta prueba de tendencia, debe asignarse un valor numérico o puntaje a cada categoría de exposición.
- La hipótesis nula indica que el riesgo del resultado de salud es el mismo para cada categoría de exposición.
- La hipótesis alternativa es que el riesgo del resultado de salud aumenta o disminuye a medida que el nivel de exposición se incrementa.
- La mejor forma de calcular el estadístico de prueba es con una calculadora.

Cálculo de la prueba de tendencia de Mantel-Haenszel incluyendo varias categorías de exposición

A continuación describiremos cómo se extiende la **prueba de Mantel-Haenszel** (**MH**) para probar si existe una tendencia significativa en varias categorías de exposición (ordinales). Para esta prueba, la disposición de los datos para el estrato *i*-avo es la siguiente:

	E_1	E_2	E_3	…	E_C	
D	a_{i1}	a_{i2}	a_{i3}		a_{iC}	m_{1i}
no D						m_{0i}
	n_{i1}	n_{i2}	n_{i3}		n_{iC}	n_i

El estadístico de prueba está representada por la siguiente fórmula general:

$$\chi^2_{\text{MH=Tendencia}} = \frac{[T_1 - E_0(T_i)]^2}{\text{Var}_0(T_i)}, i = 1, 2, ..., G$$

Los términos en la fórmula anterior se definen de la siguiente manera:

$$T_i = \text{puntaje total para el estrato } i = \sum_{j=1}^{C} a_{ij} S_j$$

S_j = el puntaje asignado para la *j*-ava categoría de exposición

$$E(T_i) = \text{puntaje total esperado para el estrato } i = \frac{m_{1i}}{n_i}\left(\sum_{i=1}^{C} n_{ij} S_j\right)$$

$$\text{Var}_0(T_i) = \text{varianza del puntaje total para el estrato } i = \frac{m_{1i} m_{0i}}{n_i}\sigma_i^2$$

$$\sigma_i^2 = \frac{1}{n_i - 1}\left\{\sum_{j=1}^{C} n_{ij} S_j^2 - \left[\sum_{j=1}^{C} n_{ij} S_j\right]^2 \Big/ n_i\right\}$$

Según la **hipótesis nula que enuncia que no hay una tendencia general significativa** de la medida de efecto en todas las categorías de exposición, el estadístico de MH de la prueba de tendencia tiene aproximadamente una distribución **chi cuadrado** con **1 grado de libertad**. Para el caso especial en que hay **G = 2** estratos, el estadístico de MH de la prueba de tendencia se simplifica a:

$$\chi^2_{\text{MH-Tendencia}} = \frac{[\{T_1 - E(T_1)\} + \{T_2 - E(T_2)\}]^2}{\text{Var}_0(T_1) + \text{Var}_0(T_2)}$$

donde

$$T_1 = \text{puntaje total para el estrato } 1 = \sum_{j=1}^{C} a_{1j} S_j$$

$$T_2 = \text{puntaje total para el estrato } 2 = \sum_{j=1}^{C} a_{2j} S_j$$

$$E(T_1) = \frac{m_{11}}{n_1}\sum_{i=1}^{C} n_{1j} S_j$$

$$E(T_2) = \frac{m_{12}}{n_2}\sum_{i=1}^{C} n_{2j} S_j$$

$$\text{Var}_0(T_1) = \frac{m_{11} m_{01}}{n_1}\sigma_1^2$$

$$\mathrm{Var}_0(T_2) = \frac{m_{12}m_{02}}{n_2}\sigma_2^2$$

$$\sigma_1^2 = \frac{1}{n_1 - 1}\left\{\sum_{j=1}^{C} n_{1j}S_j^2 - \left[\sum_{j=1}^{C} n_{1j}S_j\right]^2 \Big/ n_1\right\}$$

$$\sigma_2^2 = \frac{1}{n_2 - 1}\left\{\sum_{j=1}^{C} n_{2j}S_j^2 - \left[\sum_{j=1}^{C} n_{2j}S_j\right]^2 \Big/ n_2\right\}$$

Prueba de tendencia mediante la regresión logística

La **prueba de tendencia** que describimos en la actividad anterior puede realizarse en forma alternativa mediante el empleo de un modelo de **regresión logística**. Aquí se presenta nuevamente el modelo logístico y su forma logit equivalente, como se definió en una lección anterior:

Prueba de tendencia - La regresión logística alternativa

Modelo logístico:

$$\Pr(Y=1|X_1,X_2,\ldots,X_p) = \frac{1}{1+e^{-(b_0 + b_1X_1 + b_2X_2 + \ldots + b_pX_p)}}$$

$$\text{logit } P = \ln\frac{P}{1-P} = b_0 + b_1X_1 + b_2X_2 + \ldots + b_pX_p$$

donde $P = \Pr(Y=1|X_1,X_2,\ldots,X_p)$

X = variables independientes o los predictores

Y = variable dependiente o de resultado dicotómica

Las **X** en el modelo son las **variables independientes** o los **predictores**; la **Y** es la variable **dependiente** o **de resultado dicotómica**, que indica si una persona desarrolla o no la enfermedad. Ahora describimos la forma específica de este modelo para los datos del condado Evans que se consideraron con anterioridad, los cuales relacionan las categorías de colesterol con el desarrollo de EC dentro de 2 categorías etarias. A continuación se vuelven a mostrar estos datos:

	EDAD > 55				EDAD ≥ 55			
	COL-AL	COL-MOD	COL-BA		COL-AL	COL-MOD	COL-BA	
EC-Sí	13	25	5		9	15	4	
EC-No	48	106	54		84	161	85	
	61	131	59	251	93	176	89	358

COL-AL: > 233 (Medio = 265)

COL-MED: 184-233 (Medio = 208)

COL-BA: <184 (Medio = 164)

El modelo logístico adecuado para la prueba de tendencia correspondiente a estos datos adopta la siguiente forma logit:

$$\text{logit } P = \ln \frac{P}{1-P} = b_0 + b_1 E + b_2 A$$

La variable **E** en este modelo representa el colesterol. La variable **A** representa la edad. Más específicamente, **E** es una variable ordinal que asigna puntajes a las tres categorías de exposición. Los puntajes pueden definirse como el valor medio del colesterol en cada una de sus categorías:

$$E = \begin{cases} 265 \textbf{ si } \textbf{COL-AL} \\ 203 \textbf{ si } \textbf{COL-MED} \\ 164 \textbf{ si } \textbf{COL-BA} \end{cases}$$

Por otra parte, si los puntajes son 2, 1 y 0, entonces definimos la variable **E** como se muestra a continuación:

$$E = \begin{cases} 2 \textbf{ si } \text{COL} > 233 \\ 1 \textbf{ si } \text{COL} = 184\text{-}233 \\ 0 \textbf{ si } \text{COL} < 184 \end{cases}$$

La variable **A** se denomina **variable ficticia** o **indicadora**; y distingue los dos estratos etarios que se consideran:

$$A = \begin{cases} \textbf{1 si } \text{edad} \geq 55 \\ \textbf{0 si } \text{edad} < 55 \end{cases}$$

Preguntas de estudio (P14.19)

En general, si hay **S** estratos, entonces se requieren **S** - 1 **variables ficticias**. Por ejemplo, si hubiesen tres estratos, entonces se necesitarían dos variables ficticias. Una manera de definir las dos variables sería:

A1 = 1 si estrato 1, sino 0, A2 = 1 si estrato 2, sino 0

Para dicha codificación:

A1 = 1, A2 = 0 para el estrato 1
A1 = 0, A2 = 1 para el estrato 2
A1 = 0, A2 = 0 para el estrato 3

Mediante dicha codificación, el estrato 3 se denomina **grupo de referencia**.

Suponga que queremos estratificar por dos categorías de edad (por ejemplo, 1 = edad $\geq$ 55 vs. 0 = edad < 55) y por sexo(1 = mujeres, 0 = hombres).

1. ¿Cuántas variables ficticias debería definir?
2. ¿Cómo definiría las variables ficticias (por ejemplo, D1, D2, etc.) si el grupo de referencia involucrara hombres menores de 55 años?

3. Defina la forma logit del modelo logístico que incorporaría esta situación y permitiría una prueba de tendencia que utilice los valores medios del colesterol como puntajes.

Para el modelo que incluye solo dos estratos etarios, la hipótesis nula para la prueba de tendencia es que el verdadero coeficiente de la variable de exposición, E, es cero. Esto equivale a decir que no hay tendencia lineal en el riesgo de desarrollar EC, después de controlar la edad. La hipótesis alternativa es que hay tendencia lineal en el riesgo de desarrollar EC, después de controlar la edad.

Prueba de tendencia - la regresión logística alternativa

$$\text{logit } P = \ln \frac{P}{1-P} = b_0 + b_1 E + b_2 A$$

$H_0: b_1 = 0 \iff$ no hay tendencia lineal en el riesgo de desarrollar EC (después de controlar la edad)

$H_A: b_1 \neq 0 \iff$ hay tendencia lineal en el riesgo de desarrollar EC (después de controlar la edad)

Utilizamos un programa informático para adaptar el modelo logístico en los datos del condado Evans. Cuando definimos la variable **E** a partir del colesterol medio en cada categoría de exposición, el estadístico chi cuadrado que prueba esta hipótesis nula tiene el valor de 4,56.

$$E = \begin{cases} 265 & \text{si COl-AL} \\ 203 & \text{si COL-MED} \\ 164 & \text{si COL-BA} \end{cases}$$

Razón de verosimilitud $\chi^2_{1\,d.f.} = 4.56$

Éste se denomina estadístico de la **razón de verosimilitud** y tiene aproximadamente una distribución chi cuadrado con 1 grado de libertad. El valor p para esta prueba resulta ser 0,0327. Debido a que el valor p es menor que 0,05, rechazamos la hipótesis nula al nivel de significación de 5% y llegamos a la conclusión de que hay una tendencia lineal significativa en estos datos.

Preguntas de estudio (P14.19) (continuación)

Cuando se asignan los puntajes de exposición de 2, 1 y 0 para COL-AL, COL-MOD y COL-BA, respectivamente, la prueba de la razón de verosimilitud correspondiente para la prueba de tendencia mediante modelación logística da un valor chi cuadrado de 5,09 (p = 0,0240).

4. ¿Cuál es la diferencia entre estos resultados y los obtenidos mediante el uso de los puntajes de colesterol medio? (chi cuadrado = 4,56; p = 0,0327)

Cuando se calculó la prueba de tendencia en la actividad anterior (mediante una fórmula de sumatoria), los resultados fueron los siguientes:

puntajes de colesterol medio: chi cuadrado = 4,60 (p = 0,0320)
puntajes 2, 1, 0: chi cuadrado = 5,04 (p = 0,0248)

5. ¿Por qué cree que estos últimos resultados chi cuadrado son diferentes de los resultados obtenidos mediante la regresión logística? ¿Esto le debería preocupar?

Resumen

- La prueba de hipótesis que involucra varias categorías de exposición puede realizarse mediante regresión logística.
- Si la exposición es nominal, el modelo logístico requiere variables ficticias para distinguir las categorías de exposición.
- Si la exposición es ordinal, el modelo logístico incluye un término lineal que asigna puntajes a categorías de exposición.
- Para variables de exposición nominal u ordinal, la prueba involucra el estadístico chi cuadrado de 1 grado de libertad.
- La hipótesis nula es que *no hay asociación general entre la exposición y la enfermedad controlando covariables estratificadas*.
- De manera equivalente, la hipótesis nula es que *el coeficiente de la variable de exposición en el modelo es cero*.
- La prueba puede realizarse utilizando una prueba de la razón de verosimilitud o una prueba de Wald, las que, normalmente, brindan respuestas similares, aunque no siempre.

Auto-evaluación (P14.20)

		FA	FB	NF
	D	38	42	20
no	D	25	50	35

Los datos a la derecha corresponden a un estudio de casos y controles que se llevó a cabo para investigar la posible asociación entre fumar cigarrillos y el infarto de miocardio (IM). Todos los individuos eran hombres blancos de entre 50 y 54 años. El consumo actual de cigarrillos se dividió en tres categorías: no fumadores (NF), fumadores con bajo consumo de cigarrillos (FB) que fuman un paquete o menos por día, y fumadores con alto consumo de cigarrillos (FA) que fuman más de un paquete por día.

1. ¿Cuál es el *odds ratio* para FA vs. NF? **???**
2. ¿Cuál es el *odds ratio* para FB vs. NF? **???**

Opciones
1,47 1,78 2,66 2,70 3,06

Se categorizó a todos los individuos en función del estatus social (ES) "alto" o "bajo" según su ocupación, educación e ingreso. A continuación se muestran los datos específicos para cada estrato.

ES 'Alto'

		FA	FB	NF
	D	9	6	5
no	D	10	20	20

ES 'Bajo'

		FA	FB	NF
	D	29	36	15
no	D	15	30	15

3. Calcule los *odds ratios* específicos para cada estrato:
   a. ES alto: FA vs. NF **???**
   b. ES alto: FB vs. NF **???**
   c. ES bajo: FA vs. NF **???**

d. ES bajo: FB vs. NF **???**

4. ¿Es adecuado realizar una evaluación general para estos datos? **???**

Opciones
0,01 1,20 1,93 3,50 3,60 5,40 quizás no sí

Considere los siguientes resultados:
OR crudo, FA vs. NF = 2,66
OR crudo, FB vs. NF = 1,47
OR ajustado, FA vs. NF = 2,38
OR ajustado, FB vs. NF = 1,20

5. ¿Estos resultados ofrecen pruebas de tendencia? **???**

Opciones
no sí

La prueba de tendencia de Mantel-Haenszel se realizó con puntajes de 0, 1, 2 para los no fumadores, los fumadores con bajo consumo de cigarrillos y los fumadores con alto consumo de cigarrillos, respectivamente. El estadístico chi cuadrado de Mantel-Haenszel = 5,1. Esto corresponde a un valor p de una cola de 0,012.

6. ¿A qué conclusión llega al nivel de significación de 0,05? **???**
7. ¿A qué conclusión llega al nivel de significación de 0,01? **???**

Opciones
no se rechaza la H_0 **se rechaza la H_0**

8. Una prueba de tendencia alternativa para estos datos se puede llevar a cabo con un modelo logístico. Defina la forma logit del modelo logístico que incorporaría esta situación y permitiría una prueba de tendencia. **???**

Opciones
Logit P = $b_0 + b_1$TAB + b_2SES
Logit P = $b_0 + b_1$TAB + b_2SES1 + B_3SES2
Logit P = $b_0 + b_1$TAB + b_2TAB2 + B_3TAB3 + b_4SES

Referencias

Clayton DG. *Some odds ratio statistics for the analysis of ordered categorical data*. Biometrika 1974; 61:525-31.

Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic Research: Principles and Quantitative Methods*. John Wiley and Sons Publishers, Nueva York, 1982.

Mantel N. *Chi-square tests with one degree of freedom: Extensions of the Mantel-Haenszel procedure*. J Am Stat Assoc 1963; 58:690-700.

Mantel N, Haenszel W. *Statistical aspects of the analysis of data from retrospective studies of disease*. J Natl Cancer Inst 1959; 22(4):719-48.

Rezende NA, Blumberg HM, Metzger BS, Larsen NM, Ray SM y McGowan JE, Jr. *Risk factors for methicillin-resistance among patients with Staphylococcus aureus bacteremia at the time of hospital admission*. Am J Med Sci 2002; 323(3):117-23.

Ejercicios

EPI-1. Observación de datos estratificados

Se realizó un estudio de casos y controles para evaluar la relación entre el consumo de alcohol (CAL) y el cáncer bucal (CB). A continuación se presentan los resultados, estratificados por la condición de tabaquismo:

Fumador actual

	CAL	no CAL
CB	42	7
sin CB	3	4

Ex fumador

	CAL	no CAL
CB	100	2
sin CB	48	5

Nunca fumó

	CAL	no CAL
CB	158	4
sin CB	125	8

a. ¿Cuál es el *odds ratio* crudo para estos datos?
b. Calcule los OR específicos para cada estrato. ¿Hay evidencia de modificación de efecto por la condición de tabaquismo? Justifique su respuesta.
c. ¿Cuál de las siguientes opciones serían adecuadas para el análisis de estos datos? [Puede elegir MÁS de una opción].
   i. Calcular un *odds ratio* general de resumen para la relación entre el alcohol y el cáncer bucal, ajustado por la condición de tabaquismo.
   ii. Informar los efectos específicos para cada estrato.
   iii. Calcular pruebas $\chi 2$ de asociación para la relación entre el alcohol y el cáncer bucal DE FORMA SEPARADA para los tres estratos.
   iv. Comparar los OR crudo y ajustado para determinar si hay confusión por la condición de tabaquismo.
   v. Calcular una prueba de heterogeneidad para poder establecer si hay modificación de efecto por la condición de tabaquismo.

EPI-2. Evaluación general

Suponga que se realizó otro estudio de casos y controles de la misma relación entre la exposición y la enfermedad (es decir, alcohol y cáncer bucal), pero este estudio quedó restringido a personas que nunca fumaron. A continuación se presentan los respectivos resultados, estratificados por edad:

	40-49 años	
	CAL	**no CAL**
CB	4	25
sin CB	22	309

	50-59 años	
	CAL	**no CAL**
CB	12	10
sin CB	37	67

	60 años o más	
	CAL	**no CAL**
CB	11	12
sin CB	31	67

a. ¿Cuál es el *odds ratio* crudo para estos datos?
b. Calcule los OR específicos para cada estrato. ¿Hay evidencia de modificación de efecto por la edad? Justifique su respuesta.
c. ¿La asociación general observada entre el CAL y el CB, estratificada por la edad, es estadísticamente significativa? (Responda a esta pregunta realizando la prueba χ^2 adecuada. Asegúrese de formular la hipótesis nula y el valor p de la prueba).
d. Calcule un *odds ratio* de resumen basado en la precisión. Incluya el desarrollo de sus cálculos. ¿Qué estrato tiene la menor varianza? Justifique su respuesta.
e. Calcule un *odds ratio* de Mantel-Haenszel de resumen. Incluya el desarrollo de sus cálculos. Sobre la base de este valor de resumen, ¿hay evidencia de confusión por la edad en estos datos? Justifique su respuesta.

EPI-3. Exposición paterna a la radiación y defectos de nacimiento

Se realizó un estudio de casos y controles para evaluar si la exposición paterna a la radiación en el lugar de trabajo estaba asociada a defectos de nacimiento. A los investigadores les interesaba la edad materna como posible factor de confusión, de forma que estratificaron los datos de la siguiente manera:

	Edad materna > 35	
	Radiación	**Sin radiación**
Defectos de nacimiento	21	26

	Edad materna $\leq$ 35	
	Radiación	**Sin radiación**
Defectos de nacimiento	18	88

Controles	17	59	**Controles**	7	95

a. Calcule un *odds ratio* de Mantel-Haenszel de resumen. Incluya el desarrollo de sus cálculos.
b. ¿Hay evidencia de confusión por la edad materna en estos datos? Justifique su respuesta.
c. ¿La asociación observada entre la exposición paterna a la radiación y los defectos de nacimiento, controlando la edad materna, es estadísticamente significativa? (Responda esta pregunta realizando una prueba estadística adecuada. Asegúrese de formular las hipótesis nula y alternativa y brinde un valor p).
d. Calcule un intervalo de confianza de 95% basado en la prueba para el valor de resumen del ítem a. (Véase la "ayuda" al final de toda esta pregunta).
e. En los casos posibles, la información sobre la exposición paterna a la radiación (para el estudio descrito anteriormente) se tomó de los registros de empleo, en lugar de entrevistas a los individuos. Esto se hizo en un esfuerzo por MINIMIZAR una de las siguientes opciones: ¿cuál? [Seleccione SÓLO la mejor respuesta]:

    i. sesgo de detección
    ii. clasificación errónea diferencial del resultado
    iii. sesgo de recuerdo
    iv. sesgo de selección

Ayuda: un intervalo de confianza basado en la prueba para una medida de efecto □□(por ejemplo, □□es un OR o una RR) brinda un método alternativo para calcular un intervalo de confianza que evita tener que calcular una expresión matemática compleja para la varianza calculada de $\hat{\theta}$. Para una medida de efecto en forma de razón (por ejemplo, OR, RR, RDI), la fórmula del intervalo de confianza de 95% basado en la prueba es la siguiente:

$$\hat{\theta}^{(1 \pm 1,96/\chi_{MHS})}$$

donde χ_{MHS} es la raíz cuadrada del estadístico chi cuadrado de Mantel-Haenszel para un análisis estratificado).

EPI-4. Exposición al plomo y bajo peso al nacer

La relación entre la exposición paterna al plomo en el lugar de trabajo y el bajo peso al nacer se examinó en un estudio de cohorte retrospectivo. Los hombres con un nivel de plomo en sangre (NPS) > 50 microgramos por decilitro se consideraron expuestos. El bajo peso al nacer (BPN) se definió como el peso al nacer de menos de 2500 gramos. A continuación se brindan los datos del estudio, estratificados con respecto a la edad materna al momento del nacimiento del bebé:

	Edad materna ≤ 20			**Edad materna 21+**	
	NPS alto	**NPS bajo**		**NPS alto**	**NPS bajo**
BPN	45	30	**BPN**	68	48
Sin BPN	169	214	**Sin BPN**	257	354

a. ¿La edad materna es un factor de riesgo independiente del BPN en estos datos? Incluya el desarrollo de los cálculos y justifique su respuesta.
b. Utilice el método basado en los datos para determinar si hay confusión por la edad materna en este estudio. Incluya el desarrollo de los cálculos y justifique su respuesta.

c. ¿La asociación observada entre el nivel de plomo en sangre paterna y el bajo peso al nacer, controlando la edad materna, es estadísticamente significativa? Responda esta pregunta realizando una prueba estadística adecuada. Asegúrese de formular las hipótesis nula y alternativa y brinde un valor p o un rango del valor p.
d. Calcule la razón de riesgos general de resumen que asigne igual peso a los dos estratos definidos anteriormente (es decir, un promedio simple).
e. Calcule un intervalo de confianza de 95% basado en la prueba para el valor de resumen que se determinó en el ítem d anterior. (Véase la "ayuda" al final de la pregunta 3 para obtener la fórmula del intervalo de confianza basado en la prueba).
f. Suponga que, antes de analizar los datos, a los investigadores les preocupaba la posibilidad de confusión residual. ¿Cuál de las siguientes opciones hubiese constituido un método útil para abordar este tema? [Seleccione la mejor respuesta]:

__ Dividir el nivel de plomo en sangre (NPS) en categorías adicionales, más acotadas.
__ Dividir el bajo peso al nacer (BPN) en categorías adicionales, más acotadas.
__ Dividir la edad materna en categorías adicionales, más acotadas.
__ Todas las anteriores.

g. ¿Hay evidencia de modificación de efecto (por la edad materna) a escala aditiva en estos datos? Incluya el desarrollo de los cálculos y justifique su respuesta.
h. Sobre la base de su respuesta al ítem g, ¿cuál de las siguientes alternativas sería adecuada para el próximo paso en el análisis de estos datos? [Seleccione la mejor respuesta]:
__ Informar diferencias de riesgos específicas para cada estrato.
__ Calcular una diferencia de riesgos ajustada general y el intervalo de confianza asociado.
__ Informar la diferencia de riesgos cruda.
__ Realizar una prueba estadística para evaluar si hay confusión por la edad materna.
__ Ninguna de las opciones anteriores es adecuada.

EPI-5. Análisis estratificado: modificación de efecto

Cuando hay modificación de efecto en un análisis estratificado, ¿cómo deberían presentarse los datos? (Seleccione la mejor respuesta).

_____ Mediante la medida de asociación cruda.

_____ Mediante la medida de asociación ajustada.

_____ Mediante las medidas de asociación específicas para cada estrato.

_____ Con ninguna de las anteriores.

EPI-6. Análisis estratificado: confusión

Cuando hay confusión en un análisis estratificado, ¿cómo deberían presentarse los datos? **[Seleccione la mejor respuesta]**.

_____ Mediante la medida de asociación cruda.

_____ Mediante la medida de asociación ajustada.

_____ Mediante las medidas de asociación específicas para cada estrato.

_____ Con ninguna de las anteriores.

EPI-7. Consumo de alcohol y cáncer de vejiga: raza

Se realizó un estudio de casos y controles para evaluar la relación potencial entre el consumo de alcohol y el cáncer de vejiga. A continuación se resumen los datos de ese estudio, estratificados en tres categorías de raza.

Blancos

	CAL	no CAL
Casos	72	41
Controles	106	105

Negros

	CAL	no CAL
Casos	93	54
Controles	113	113

Asiáticos

	CAL	no CAL
Casos	68	33
Controles	78	142

a. Calcule los OR específicos para cada estrato. ¿Hay evidencia de modificación de efecto por la raza? Justifique su respuesta. (Aquí podría emplear una prueba estadística, es decir, la prueba de Breslow-Day, además de comparar las estimaciones puntuales para los tres estratos).
b. ¿Debería realizar una prueba de asociación de Mantel-Haenszel general que controle la raza mediante los tres estratos? Justifique su respuesta.
c. Considerando sólo la información sobre los individuos negros y blancos, ¿la asociación general observada entre CAL y el cáncer de vejiga, estratificada por la raza, es estadísticamente significativa? (Asegúrese de formular la hipótesis nula y el valor p de la prueba).
d. ¿Debería calcular un *odds ratio* ajustado general que controle la raza mediante los tres estratos? Justifique su respuesta.
e. Considerando solo la información sobre individuos negros y blancos, calcule un aOR basado en la precisión y un mOR. Para el aOR, ¿qué grupo recibe más peso? ¿los negros o los blancos?
f. Calcule y compare intervalos de confianza de 95% para el aOR y el mOR.
g. ¿A qué conclusión llega sobre la relación entre el CAL y el cáncer de vejiga?

EPI-8. Actividad física e incidencia de la diabetes

Se llevó a cabo un estudio de cohorte sobre la actividad física (AF) y la incidencia de la diabetes durante un período de seis años entre hombres japoneses-estadounidenses en Honolulu. A continuación se resumen los datos de ese estudio, estratificados según el índice de masa corporal.

	IMC alto	
	Mucha AF	**Poca AF**
Diabetes	48	62
Personas-años	1050	1067

	IMC bajo	
	Mucha AF	**Poca AF**
Diabetes	54	71
Personas-años	1132	1134

a. ¿Hay evidencia de modificación de efecto por el IMC? Justifique su respuesta.
b. Calcule y compare una aRDI basada en la precisión y una mRDI.
c. ¿Hay evidencia de confusión por el IMC en estos datos?
d. ¿Debería calcular un *odds ratio* ajustado general que controle el IMC utilizando los tres estratos? Justifique su respuesta.
e. ¿La asociación observada entre la actividad física y la diabetes, controlando el IMC, es estadísticamente significativa? Asegúrese de formular las hipótesis nula y alternativa que se someten a prueba, el estadístico de prueba y el valor p.
f. Calcule y compare intervalos de confianza de 95% para la aRDI y la mRDI.
g. ¿A qué conclusión llega sobre la relación entre la AF y la diabetes en función de estos datos?

Respuestas a las preguntas de estudio y auto-evaluaciones

P14.1

1. Para los fumadores y no fumadores en forma separada, no parece haber asociación entre la exposición al TCX y el desarrollo de cáncer de pulmón. Sin embargo, puede deducirse que la RR de 1,3 para los fumadores indica una asociación moderada; no obstante, este valor es altamente no significativo.
2. No. Las dos razones de riesgos específicas para cada estrato son básicamente iguales. Nuevamente, la RR de 1,3 para los fumadores indica un efecto pequeño, pero es altamente no significativa.
3. No. Aunque el valor de efecto crudo es 2,1, el análisis correcto requiere que se controle el tabaquismo, para el cual los datos muestran una ausencia de efecto de la exposición al TCX. Un cálculo ajustado para los dos estratos brindaría un estadístico de resumen adecuado que controla el tabaquismo.
4. Ya que el valor puntual ajustado es cercano al valor nulo de 1 y el estadístico de prueba de Mantel-Haenszel no es significativo, debería llegar a la conclusión de que no hay evidencia de una relación entre la exposición y la enfermedad a partir de estos datos.

P14.2

1. Sí. El *odds ratio* de 11,67 es muy alto y la prueba de MH es muy significativa; además, aunque el intervalo de confianza es amplio, no incluye el valor nulo.
2. La asociación puede cambiar cuando se controlan una o más variables. Si esto sucede y las variables de control son factores de riesgo; entonces lo más adecuado sería uno o más cálculos ajustados.
3. No necesariamente. Si una o más de estas variables no son factores de riesgo de la condición de SARM previamente conocidos, entonces es posible que dichas variables no se puedan controlar.
4. Debe de haberse perdido información sobre la condición de SARM o la hospitalización previa de algunos individuos del estudio (n = 5). De hecho, faltaban cinco observaciones en la última variable.
5. No. Los *odds ratios* específicos para cada estrato dentro de diferentes grupos etarios son muy cercanos (alrededor de 11).
6. No. El valor p de 0,95 es muy alto, lo que indica que no hay evidencia de interacción debido a la edad.
7. Sí. La evaluación general es adecuada, porque

no hay evidencia de interacción debido a la edad.
8. No. Los *odds ratios* crudo y ajustado son básicamente iguales.
9. Sí. La prueba de Mantel-Haenszel para los datos estratificados es muy significativa (P < 0,0001).
10. El intervalo de confianza es bastante amplio, lo que indica que aunque el cálculo ajustado es tanto estadísticamente significativo como importante, tiene poca precisión.
11. Sí. La evaluación general es adecuada, porque no hay evidencia de interacción debido al sexo.
12. No hay confusión ya que, cuando se controla el sexo, los *odds ratios* crudo y ajustado son básicamente iguales.
13. Sí. La prueba de Mantel-Haenszel para los datos estratificados es muy significativa (P < 0,0001).
14. La respuesta a esta pregunta es "quizás". Parece haber interacción porque el *odds ratio* es 8,48 con uso previo de drogas, pero sólo 3,66 sin uso previo de drogas. Sin embargo, los dos valores del *odds ratio* están del mismo lado que 1, de forma que un cálculo ajustado no sería el resultado de efectos opuestos que se cancelan uno con el otro. Asimismo, la prueba de BD para la interacción es no significativa, lo que respalda la idea de realizar una evaluación general.
15. Sí. Cuando controlamos el uso previo de drogas, el *odds ratio* crudo de 11,67 es bastante distinto del *odds ratio* mucho más pequeño de 5,00.
16. Sí. La prueba de Mantel-Haenszel para datos estratificados es muy significativa (P < 0,0001) y, aunque el intervalo de confianza es amplio, sigue sin incluir el valor nulo.

P14.3

1. Es necesario controlar el uso previo de drogas antimicrobianas porque es un factor de confusión.
2. Sí. Se gana precisión al controlar el uso previo de drogas antimicrobianas, ya que la amplitud del intervalo de confianza para el cálculo ajustado es mucho más acotada que la amplitud del intervalo de confianza que corresponde a los datos crudos.
3. No. Ni el *odds ratio* ajustado ni el intervalo de confianza, ni el valor p de MH cambian de forma considerable cuando se comparan los resultados que controlan solo el UPDAM con los resultados que controlan variables adicionales.
4. No. Todos los valores p son bastante grandes, lo que indica que la hipótesis nula de ausencia de interacción no debería rechazarse. Sin embargo, probablemente una comparación de cálculos específicos para cada estrato puede sugerir interacción cuando se controla más de una variable.
5. Porque el *odds ratio* es indefinido en un estrato con frecuencia de celda cero.
6. OR = 4,66 es una elección adecuada, ya que controla las tres variables que se consideran para el control. De manera alternativa, OR = 5,00 también es adecuada, puesto que surge de controlar solo el uso previo de drogas antimicrobianas, que es la única variable que afecta a la confusión y a la precisión.
7. Sí. El *odds ratio* ajustado (cercano a 5,00) indica un efecto fuerte que también es estadísticamente significativo. El intervalo de confianza de 95% indica una falta de precisión, pero los resultados revelan, en general, un efecto fuerte.
8. (Nota: no hay pregunta 8).
9. Hay cantidades pequeñas, incluso ceros en casi todos los cuadros.
10. Incluso cuando no hay celdas con cero, los análisis específicos para cada estrato son, en general, poco confiables debido a las cantidades pequeñas.
11. Sí. El valor del *odds ratio* en el cuadro 5 es 24,00, mientras que el *odds ratio* en el cuadro 6 es 1,71 y el *odds ratio* en el cuadro 1 es 5,89, todos valores bastante distintos.
12. La prueba BD no es significativa, aun cuando todos los *odds ratios* son distintos, están del mismo lado que el valor nulo y los estratos involucran cantidades muy pequeñas.

P14.4

1. Falso. La estratificación también implica realizar una evaluación general cuando sea adecuada.
2. Verdadero
3. Verdadero
4. Falso. Un valor de resumen general puede considerarse inadecuado si hay evidencia considerable de interacción.
5. Verdadero
6. No. Quizás. Sí.

P14.5

1. Para una hipótesis alternativa de una cola, el área en la cola derecha es el doble del valor p, por lo cual el valor p de una cola es igual a la mitad de 0,28, ó 0,14.
2. No hay que rechazar la hipótesis nula de ausencia de efecto general. No hay evidencia de que la exposición al TCX esté asociada al desarrollo de cáncer de pulmón cuando se controla el tabaquismo.
3. Si hubiese interacción en lados opuestos del valor nulo, entonces uno de los dos términos de la suma sería negativo y el otro sería positivo. En consecuencia, la suma de estos términos podría ser cercana a cero, dando una prueba chi cuadrado no significativa, incluso si las pruebas específicas para cada estrato fueran ambas significativas.

P14.6

1. Sí, definitivamente. El numerador siempre es una suma de cantidades que luego se elevan al cuadrado, en contraposición a una suma de cantidades que se elevan al cuadrado antes de la sumatoria. En consecuencia, valores positivos grandes de algunos estratos pueden cancelar valores negativos grandes de otros estratos, dando una prueba de MH no significativa.
2. Si los efectos específicos de cada estrato están todos del mismo lado que el valor nulo, entonces todas las cantidades en el numerador del estadístico de prueba tienen el mismo signo y, por lo tanto, no pueden cancelarse unas con otras.

P14.7

1. Para una hipótesis alternativa de una cola, el área en la cola derecha es el doble del valor p, por lo cual el valor p de una cola es la mitad de 0,04, ó 0,02.
2. La hipótesis nula de ausencia de efecto general se rechazaría al nivel de 5%, pero no al nivel de 1%.

P14.8

1. 0,79
2. 0,76
3. 0,77
4. No (para calcular la RDI cruda, hay que combinar los datos de ambos estratos).
5. No
6. Mayor precisión
7. No se rechaza la hipótesis nula
8. Se rechaza la hipótesis nula

P14.9

1. B
2. Falso. Un intervalo de confianza más grande implica menos precisión y, por lo tanto, un peso más pequeño.
3. Falso. La magnitud de la razón de riesgos no es un factor en la determinación de ponderaciones basadas en la precisión.
4. Verdadero
5. Verdadero

P14.10

1. Recuerde que, para determinar la precisión, el tamaño de la muestra no es tan importante como contar con un grupo de datos equilibrados. Los pesos corresponden a qué considerablemente equilibrados están los conjuntos de datos. Cuanto más equilibrio hay, mayor es la ponderación.

P14.11

1. 4,00
2. 4,69
3. Negros o hispanos
4. 4,16
5. Discutible: el POR crudo de 3,59 y el POR ajustado de 4,16 son diferentes, pero no tan alejados, por lo cual, decidir si hay una diferencia relevante es discutible. Observe, sin embargo, que si requerimos un margen de diferencia de 10% para evaluar la confusión, entonces 3,59 es menor que un cambio negativo de 10% (3,64) en el POR ajustado. Por lo tanto, con la regla de 10% de margen de cambio, llegaríamos a la conclusión de que hay confusión.

P14.12

1. OR(modificado por 0,1) = (5,1 x 5,1) / (0,1 x 6,1) = 42,6
2. OR(modificado por 1,0) = (6 x 6) / (1 x 7) = 5,1
3. El *odds ratio* modificado específico para cada estrato puede cambiar radicalmente en función de qué valor se utiliza para ajustar todas las frecuencias de celda en un estrato con una frecuencia de celda cero. En consecuencia, dicho enfoque es bastante problemático.
4. $\hat{aOR} = \exp\left[\frac{\{0,3972 \times \ln(9,31)\} + \{0,6618 \times \ln(9,00)\}}{0,3972 + 0,6618}\right] =$

   $\exp\left[\frac{0,8862 + 1,4541}{1,0590}\right] = 9,11$
5. $\hat{aOR} = \exp\left[\frac{\{0,0947 \times \ln(42,64)\} + \{0,6618 \times \ln(9,00)\}}{0,0947 + 0,6618}\right] =$

   $\exp\left[\frac{0,3354 + 1,4541}{0,7565}\right] = 10,65$
6. Son distintas (9,11 mediante un ajuste de 0,5 vs. 10,65 con un ajuste de 0,1), pero no muy diferentes.
7. No parece que la elección del factor de ajuste tenga un gran impacto en el aOR resultante, aunque puede tener un gran impacto en el valor del *odds ratio* en el estrato que se ajusta.

P14.13

1. El mOR será indefinido, porque cada término de la suma en el denominador será cero.
2. El mOR será indefinido, porque cada término de la suma en el numerador será cero.
3. Es posible, aunque poco probable, que tener frecuencias de celda cero en cada estrato pueda hacer que el mOR sea indefinido. Sin embargo, si dicha situación se presenta, el mOR no funcionará. En lugar de ello, probablemente tenga que usar, como alternativa, el enfoque que suma 0,5 a cada frecuencia de celda en algún estrato con una celda con cero.
4. Los tres cálculos ajustados son distintos y el mOR está bastante separado de los valores de aOR. Una razón clave para preferir el mOR es que evita el empleo de un valor modificador arbitrario como 0,5 ó 0,1. Además, el mOR tiene buenas propiedades estadísticas, como se describió anteriormente.

P14.14

1. Verdadero
2. Falso. Lo que más se afecta cuando se suma un número pequeño a cada celda son los cálculos específicos para cada estrato.
3. Verdadero

P14.15

1. El valor nulo de 1 está apenas contenido en el intervalo de confianza. Aunque el intervalo no es muy amplio, el valor puntual de 1,71 es, en cierta medida, poco confiable, ya que el intervalo ronda entre básicamente una ausencia de asociación y una asociación moderadamente fuerte.
2. IC del 95% para aDR:

   $0,0850 \pm \frac{1,96}{\sqrt{490,2243}} =$

   $0,0850 \pm 0,0885 =$

   $(-0,0035, 0,1735)$
3. El valor nulo de 0 está apenas contenido en el intervalo. Aunque el intervalo no es muy amplio, el valor puntual de 0,0850 es, en cierta medida, poco confiable, ya que el intervalo ronda entre básicamente una ausencia de asociación y una asociación moderadamente fuerte de 0,17 para la diferencia de riesgos.

P14.16

1. 1,57; 12,98
2. Sí

P14.17

1. Las razones de riesgos que comparan COL-AL con COL-BA no difieren mucho entre los dos

grupos etarios (2,51 vs. 2,15). De forma similar, cuando se compara COL-MOD con COL-BA, las razones de riesgos no difieren mucho entre los dos grupos etarios (2,25 vs. 1,90). En general, estos resultados indican que no hay interacción significativa de la edad con el nivel de colesterol.
2. La evaluación general es adecuada, porque no hay evidencia de interacción, especialmente a lados opuestos del valor nulo.
3. No hay evidencia de confusión, ya que el valor crudo y el valor ajustado son básicamente equivalentes cuando se comparan los grupos COL-AL vs. COL-BA y los grupos COL-MOD vs. COL-BA.

P14.18

1. (1) No hay tendencia significativa en el riesgo de EC en las tres categorías. (2) Los riesgos en cada estrato son iguales.
2. Ya que el valor p es 0,0320 (de dos colas), se llega a la conclusión de que hay una tendencia significativa en los riesgos de EC en las tres categorías mediante el método de asignación de puntaje que involucra los valores medios de COL dentro de cada categoría de COL.
3. Es probable que el estadístico chi cuadrado para la tendencia cambie un poco, pero probablemente se llegaría a la misma conclusión obtenida para otros métodos de asignación de puntaje, aunque no se garantiza que esto ocurra siempre.
4. El estadístico chi cuadrado y los valores p correspondientes son apenas diferentes, pero ambos estadísticos rechazarían la hipótesis nula de que no existe tendencia lineal al nivel de significación de 0,05.
5. Estadísticos chi cuadrado idénticos e valores p idénticos correspondientes. La razón: los puntajes son equidistantes. Éste también sería el caso si los puntajes fuesen 15, 10 y 5.

P14.19

1. Tres variables ficticias, porque habría 2 x 2 = 4 estratos de edad por sexo.
2. D1 = 1 si es mujer $\geq$ 55 años, sino 0; D2 = 1 si es mujer < 55 años, sino 0; y D3 = 1 si es hombre $\geq$ 55 años, sino 0.
3. logit $P = b_0 + b_1E + b_2D1 + b_3D2 + b_4D3$ donde E adopta los tres valores 265, 208 y 164 para los tres estratos del colesterol.
4. Son levemente distintos, como se podía esperar, pero llevan a la misma conclusión sobre la hipótesis nula. Sin embargo, es posible que diferentes sistemas de asignación de puntaje puedan generar otras conclusiones.
5. La perspectiva de la regresión logística utiliza un enfoque apenas distinto (llamado ***máxima verosimilitud***) que el punto de vista anterior (fórmula de sumatoria) para obtener los cálculos del modelo y las pruebas correspondientes. Aunque ambos enfoques pueden conducir a respuestas apenas diferentes, son equivalentes si los tamaños de las muestras son lo suficientemente grandes. Sin embargo, la mayoría de los estadísticos prefieren el enfoque de regresión logística debido a las propiedades de los cálculos de máxima verosimilitud.

P14.20

1. 2,66
2. 1,47
3. a. 3,60; b. 1,20; c. 1,93; d. 1,20
4. Quizás: aquí hay algunas pruebas de interacción, pero como hay interacción del mismo lado, el investigador debe decidir si la diferencia de 3,6 versus los otros valores es relevante.
5. Sí.
6. Se rechaza la H_0.
7. No se rechaza la H_0.
8. Logit $P = b_0 + b_1TAB + b_2SES$

LECCIÓN 15

EQUIPARAMIENTO

El equiparamiento es una opción de control que está disponible en la etapa de diseño del estudio. Anteriormente presentamos el equiparamiento en la sección 13-2 de la lección 13. Sugerimos que revise esa actividad antes de proceder con la presente. El objetivo principal del equiparamiento es ganar ***precisión*** *al calcular la medida de efecto de interés. El equiparamiento tiene otras ventajas, al igual que desventajas. En esta lección se brinda una definición general, se detallan los diferentes tipos de equiparamiento, se analiza si conviene equiparar o no y se describe cómo analizar los datos equiparados.*

15-1 Equiparamiento

Definición y ejemplo de equiparamiento

El síndrome de Reye es una enfermedad poco frecuente que afecta al cerebro y al hígado, y puede producir delirios, coma y la muerte. Suele afectar a niños y, por lo general, se manifiesta después de una enfermedad viral.

A fin de descubrir si la aspirina era un determinante del síndrome de Reye, en 1982, los investigadores llevaron a cabo un estudio de casos y controles equiparados que empleaba un sistema de vigilancia estatal para identificar todos los casos incidentes que padecían el mencionado síndrome en Ohio. Como grupo de comparación se seleccionaron controles equiparados basados en la población. Primero se identificaron los controles potenciales mediante un muestreo, realizado en todo el estado, de niños que habían sufrido enfermedades virales pero que no habían desarrollado el síndrome de Reye. Luego se eligieron los controles del estudio equiparando individualmente con cada caso a uno o más niños de la misma edad y con la misma enfermedad viral que éste. A los padres de los casos y los controles se les preguntó sobre los medicamentos que tomaron sus hijos durante la enfermedad, incluso el consumo de aspirinas.

Preguntas de estudio (P15.1)

1. ¿Por qué cree que el **tipo de enfermedad viral** se consideró una de las variables de equiparamiento en este estudio?
2. ¿Por qué cree que la **edad** se seleccionó como una variable de equiparamiento?

Este es un ejemplo clásico del uso de equiparamiento individual en un estudio de casos y controles. Aunque la forma más simple es el **equiparamiento uno a uno** o el **equiparamiento formando parejas** (también, emparejamiento o apareamiento), este estudio permite más de un control por caso.

Por lo general, el equiparamiento implica la comparación de dos grupos: el grupo índice y el grupo de comparación. En un estudio de casos y controles, el grupo índice es el conjunto de casos, por ejemplo, niños con el síndrome de Reye, y el grupo de comparación es el conjunto de controles.

Si el diseño del estudio fuese un estudio de cohorte o un ensayo clínico, el grupo índice sería, en tal caso, el conjunto de personas expuestas, y el grupo de comparación sería el conjunto de personas no expuestas. Como raramente se utiliza el equiparamiento en los estudios de cohorte o en los ensayos clínicos, nuestro enfoque aquí

estará puesto en los estudios de casos y controles.

Equiparamiento	Casos y Controles	Cohorte/ Ensayo clín.
Grupo índice	Casos	Expuestos
versus		
Grupo de comparación	Controles	No expuestos

Sin importar el tipo de diseño equiparado que se utilice, la **característica clave** es que el **grupo de comparación se restringe para que sea similar al grupo índice** con respecto a los factores que se equiparan. Por lo tanto, en el estudio del síndrome de Reye, se limitó a los controles para que tengan la misma distribución que los casos con respecto a las variables edad y tipo de enfermedad viral. Pero no se restringió la distribución de la edad o la enfermedad viral para los casos. Ésta es la razón por la cual decimos que el equiparamiento impone una restricción **parcial** sobre el grupo de control en un estudio de casos y controles.

Resumen

- El estudio de 1982 que analizó la relación entre la aspirina y el síndrome de Reye en niños es un ejemplo clásico de equiparamiento individual en un estudio de casos y controles.
- La forma más simple de equiparamiento individual es el equiparamiento uno a uno o el equiparamiento formando parejas, pero también puede implicar más de un control por caso.
- Normalmente, el equiparamiento compara un ***grupo índice*** con un ***grupo de comparación***.
- En un estudio de casos y controles, el grupo índice y el grupo de comparación son los casos y los controles, respectivamente.
- En un estudio de cohorte o en un ensayo clínico, el grupo índice y el grupo de comparación son los expuestos y los no expuestos, respectivamente.
- La característica clave del equiparamiento es que el grupo de comparación se restringe para que sea similar al grupo índice con respecto a la distribución de los factores de equiparamiento.

Tipos de equiparamiento

Existen dos tipos de equiparamiento: el **equiparamiento individual** y el **equiparamiento por frecuencias**. El **equiparamiento individual**, por ejemplo, en un estudio de casos y controles, se realiza seleccionando en forma secuencial, para un caso por vez, uno o más controles para cada caso, de forma que los controles tengan características iguales o similares al caso con respecto a cada una de las variables que se equiparan. Por ejemplo, si equiparamos por la edad, la raza y el sexo, entonces los controles para un caso dado se seleccionan de manera que tengan igual o similar edad, raza o sexo que el caso.

Cuando se equipara con respecto a variables continuas, como la edad, es necesario especificar una regla para decidir cuándo el valor de la variable que se equipara es lo "suficientemente cercano". El enfoque más generalizado para las variables continuas es el **equiparamiento por categorías**. (Nota: el equiparamiento por categorías es una de las diversas formas de realizar equiparamiento individual involucrando una variable continua. Véase el recuadro al final de esta actividad para obtener una descripción de otras maneras de equiparar con respecto a una variable continua). Las categorías elegidas para este tipo de equiparamiento deben especificarse antes del proceso de equiparamiento. Por ejemplo, si se establece que las categorías de equiparamiento para la edad son bandas de 10 años, entonces el equiparamiento de un control para un caso de 40 años debe provenir del rango de edad de 36-45 años.

El primer paso es categorizar cada una de las variables de equiparamiento, ya sea continua o discreta. Luego, para cada individuo índice, se equipara eligiendo a uno o más individuos de comparación que estén en la

misma categoría que el individuo índice, para cada una de las variables que se equiparan.

Preguntas de estudio (P15.2)

Considere un estudio de casos y controles que conlleva el equiparamiento individual por categorías con respecto a las variables edad, sexo, condición de tabaquismo, presión arterial y tamaño corporal.

1. ¿Qué debe hacer primero antes de realizar el equiparamiento?
2. ¿Cómo efectúa el equiparamiento para un caso dado?
3. Si el caso es un hombre fumador de 40 años que es obeso y tiene presión arterial alta, ¿su control equiparado puede ser un hombre fumador de 40 años de tamaño corporal normal y presión arterial baja? Desarrolle su respuesta.

En el equiparamiento por frecuencias, se equipara de forma grupal, en lugar de individual. Los controles se eligen como un grupo que posea la misma distribución que los casos con respecto a las variables que se equiparan. Por ejemplo, podríamos equiparar por frecuencias con respecto a la presión arterial y la edad en un estudio de casos y controles donde los casos poseen el desglose de la categoría presión arterial por edad que se muestra a continuación, asegurando que los controles tengan la misma disociación:

	PAlta, Edad>55	PAAlta, Edad<55	PANorm, Edad>55	PANorm, Edad<55	**Total**
Casos	**40**	**60**	**50**	**50**	**200**

Preguntas de estudio (P15.2) (continuación)

Suponga que desea tener el triple de controles totales en comparación con los casos. Responda a las siguientes preguntas.

4. ¿Cuál es el desglose del grupo de presión arterial según el grupo de edad para la cantidad de controles?
5. ¿Cuál es el desglose del porcentaje según el grupo de presión arterial y el grupo de edad combinados para los controles?

¿Cómo decide si debe emplear equiparamiento individual o por frecuencias? La elección depende principalmente de qué tipo de equiparamiento es más conveniente en términos de tiempo, costo y clase de información disponible con respecto a las variables que se equiparan. La selección también se funda en la cantidad de variables que están involucradas en el equiparamiento. Cuantas más variables hay, más difícil es crear grupos equiparados sin formar pares en forma individual.

Preguntas de estudio (P15.2) (continuación)

Suponga que los casos son mujeres de más de 55 años con cáncer de ovario de diferentes hospitales y que usted quiere que los controles sean mujeres hospitalizadas con fracturas óseas traumáticas equiparadas por edad y hospital.

6. ¿Qué sería más conveniente, el equiparamiento por frecuencias o el individual?

Supongamos que los casos son mujeres de un hospital de más de 55 años con cáncer de ovario y los controles son mujeres hospitalizadas con fracturas óseas traumáticas equiparadas por edad, raza, cantidad de hijos, edad de la primera relación sexual y de la primera menstruación.

7. ¿Qué sería más conveniente, el equiparamiento por frecuencias o el individual?

Resumen

- Existen dos tipos generales de equiparamiento: el **equiparamiento individual** y el **equiparamiento por frecuencias**.
- El **equiparamiento individual** en un estudio de casos y controles se efectúa de a un caso por vez.
- Con el equiparamiento individual se seleccionan en forma secuencial uno o más controles para cada caso, de forma que los controles tengan características iguales o similares al caso dado con respecto a cada una de las variables involucradas en el equiparamiento.
- Para las variables continuas, el equiparamiento puede llevarse a cabo mediante el equiparamiento calibrado, el equiparamiento con el vecino más cercano o el equiparamiento por categorías (el más común).
- El **equiparamiento por frecuencias** implica equiparar por categorías de manera grupal, en lugar de usar el equiparamiento individual.

Otros tipos de equiparamiento con respecto a variables continuas

Aunque el equiparamiento por categorías es el enfoque más difundido para equiparar variables continuas, como la edad, existen otros dos enfoques que a veces se adoptan en su lugar. Un procedimiento, que se denomina **equiparamiento calibrado**, conlleva la elección del control dentro de cierto intervalo definido donde se encuentra el caso con respecto al factor de equiparamiento. Por ejemplo, cuando se equipara con respecto a la edad, este intervalo, también llamado calibre, puede requerir que el control tenga hasta 5 años más o menos que la edad del caso.

En general, el equiparamiento calibrado se define como la especificación de un valor c y el requisito de que dos sujetos se equiparen si el individuo de comparación (por ejemplo, el control) se encuentra dentro de más o menos c unidades de la variable de equiparamiento que el individuo índice (por ejemplo, el caso).

Uno de los problemas del equiparamiento calibrado puede presentarse si se elige una c muy pequeña (es decir, reducida) ya que, entonces, podría ser difícil encontrar individuos equiparados. Otro inconveniente puede ocurrir si se selecciona una c muy grande; de ser así, el grupo de control no queda necesariamente restringido de manera que tenga la misma distribución que los casos con respecto al factor de equiparamiento. Si, por ejemplo, se equiparó con respecto a la edad con c = 30 años, entonces la distribución de la edad de los controles podría ser muy distinta de la distribución de la edad de los casos.

Una segunda alternativa para equiparar variables continuas se denomina equiparamiento **con el vecino más cercano**. En este procedimiento se procede al equiparamiento mediante la selección de uno o más individuos más cercanos al caso con respecto a la variable que se equipara. El equiparamiento con el vecino más cercano puede llevar mucho tiempo porque puede demandar la comparación de todos los controles potenciales para un caso dado (siempre que ningún control potencial tenga exactamente el mismo valor que el caso respecto a la variable que se equipara). Asimismo, cuando hay diversas variables de equiparamiento, es probable que el vecino más cercano para una variable no sea el vecino más cercano para otra variable correspondiente al mismo control potencial. Por lo tanto, puede ser difícil encontrar el vecino más cercano para varias variables en forma simultánea.

Razones de equiparamiento

Un tema importante de diseño para el estudio equiparado es la razón de la cantidad de individuos de comparación sobre la cantidad de individuos índice en cada estrato equiparado. Ésta se denomina **razón de equiparamiento** para el estudio equiparado. A continuación presentamos una lista de posibles razones de equiparamiento:

$$\text{Razón de equiparamiento} = \frac{\#\text{ individuos de comparación}}{\#\text{ individuos índice}}$$

Nombre	#Casos	#Controles	Tipo	¿Por qué?
1 a 1 (formación de pares)	1	1	Individual	Ganar precisión (vs no equiparar)
R a 1	1	R	Individual	Ganar precisión (aumenta n)
R_i a 1	1	R_i	Individual	Tratar para R, pero encontrar < R
R_i a S_i	S_i	R_i	Frecuencia o Individual	Conveniencia o 'agrupamiento'

La razón más pequeña y simple es **1 a 1**, también denominada **formación de pares**. En un estudio de casos y controles, la formación de pares equipara 1 control con cada caso y requiere equiparamiento individual. ¿Por qué recurrir a la formación de pares? La formación de pares puede llevarnos a ganar precisión en la medida de efecto calculada, en comparación con la ausencia de equiparamiento en un estudio del mismo tamaño total. Asimismo, es más fácil encontrar un par antes que hallar varias coincidencias por cada individuo índice.

El **equiparamiento R a 1** en un estudio de casos y controles implica elegir **R** controles para cada caso mediante el equiparamiento individual. Por ejemplo, el equiparamiento 3 a 1 en un estudio de casos y controles requeriría tres controles para cada caso. Se prefiere el equiparamiento R a 1 antes que la formación de pares porque se puede ganar aún más precisión del mayor tamaño muestral total. Sin embargo, desde un punto de vista práctico, puede ser difícil encontrar más de varios controles equiparados para cada caso.

El **equiparamiento R_i a 1** permite una cantidad variada de individuos equiparados para casos distintos mediante el equiparamiento individual. Por ejemplo, pueden encontrarse 3 controles para un caso, pero solo 2 controles para otro caso y quizás solo 1 control para otro caso. El equiparamiento R_i a 1 a menudo no se planea en forma inicial sino que, por el contrario, surge de intentar realizar un equiparamiento R a 1 y luego encontrar menos que **R** combinaciones para algunos casos.

El **equiparamiento R_i por S_i** permite que uno o más controles se equiparen grupalmente con varios casos, también considerados como un grupo. La letra **i** aquí denota el grupo o estrato equiparado **i-avo** que contiene R_i controles y S_i casos. Normalmente, esta razón de equiparamiento surge del equiparamiento por frecuencias, pero también puede presentarse a partir del equiparamiento individual cuando **se agrupan estratos equiparados intercambiables**.

Preguntas de estudio (P15.3)

En una actividad posterior se brinda un análisis detallado sobre el **agrupamiento**. Considere un estudio de casos y controles equiparados de manera individual que involucra el equiparamiento 2 a 1, donde la única variable equiparada es la condición de tabaquismo (es decir, TAB = 0 para los no fumadores y TAB = 1 para los fumadores). Suponga que hay 100 conjuntos equiparados en los que 30 conjuntos están compuestos por todos los fumadores y 70 conjuntos están compuestos por todos los no fumadores. Suponga, además, que ***agrupamos*** los 30

conjuntos ("intercambiables") correspondientes a los fumadores en un estrato combinado y a los 70 conjuntos ("intercambiables") que contienen a los no fumadores en otro estrato combinado.

1. ¿Cuántos casos y controles hay en el primer estrato equiparado (que combina 30 conjuntos equiparados)?
2. ¿Cuántos casos y controles hay en el segundo estrato equiparado (que combina 70 conjuntos equiparados)?
3. ¿Qué tipo de esquema de razón de equiparamiento se emplea cuando se usan datos combinados? ¿R a 1 o R_i a S_i?

Considere el siguiente cuadro determinado por el **equiparamiento por frecuencias** con respecto a raza y sexo en un estudio de casos y controles.

Equiparamiento por frecuencias con respecto a raza y sexo en un estudio de casos y controles

	Hombre blanco	Mujer blanca	Hombre negro	Mujer negra
Casos	100 (33%)	100 (33%)	40 (13%)	60 (20%)
Controles	200 (33%)	200 (33%)	80 (13%)	120 (20%)

4. ¿Cuántos estratos equiparados hay en este estudio equiparado por frecuencias?
5. ¿Qué tipo de razón de equiparamiento describe este diseño: R a 1 o R_i a S_i?
6. ¿Cuáles son las cantidades de controles y casos en cada estrato?

Resumen

- La **razón de equiparamiento** para un diseño equiparado es la razón de la cantidad de individuos de comparación sobre la cantidad de individuos índice en cada estrato equiparado.
- Las razones de equiparamiento pueden ser **1 a 1**, **R a 1**, **R_i a 1** o **R_i a S_i**.
- La razón de equiparamiento más simple es 1 a 1, que también se denomina formación de pares.
- El **equiparamiento R a 1** otorga más precisión que el equiparamiento 1 a 1 debido al tamaño muestral más grande, pero encontrar R combinaciones para cada individuo índice puede ser difícil.
- El **equiparamiento R_i a 1** normalmente ocurre cuando se trata de efectuar un equiparamiento R a 1, pero se encuentran menos de R individuos de comparación para algunos individuos índice.
- El **equiparamiento R_i a S_i** comúnmente surge del equiparamiento por frecuencias, pero también puede resultar de la agrupación de estratos equiparados artificialmente a partir del equiparamiento individual.

¿Cuántos equiparados se deben seleccionar?

Si se recurre al **equiparamiento R a 1**, ¿qué tan grande debería ser R? La respuesta ampliamente aceptada para esta pregunta es que no se gana mucho en términos de precisión si se emplea una $R > 4$. La justificación normal se basa en el **criterio de eficiencia de Pittman**, que es aproximadamente el cociente de dos varianzas: la varianza calculada por emparejamiento de un *odds ratio* ajustado *dividida entre* la varianza correspondiente calculada por equiparamiento R a 1. Aquí está la fórmula de la eficiencia de Pittman:

Criterio de la eficiencia de Pittman aproximadamente

$$\frac{\text{Var}(a\hat{O}R_{1\ to1})}{\text{Var}(a\hat{O}R_{R\ to1})} \approx \frac{2R}{(R+1)}$$

El cálculo de este criterio para varios valores de **R** brinda el siguiente cuadro:

R	1	2	3	4	5	6	...	∞
2R/(R+1)	1.000	1.333	1.500	1.600	1.667	1.714	...	2
aumento en el %	-	33.3	12.5	6.7	4.2	2.8		-

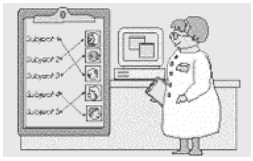
Este cuadro muestra retornos cada vez más pequeños cuando **R** es mayor que 4. La eficiencia de Pittman aumenta un 33,3% a medida que **R** pasa de 1 a 2, pero solo el 4,3% cuando **R** pasa de 4 a 5. El aumento en el porcentaje de la eficiencia es, claramente, bastante pequeño a medida que **R** supera 4. Asimismo, la máxima eficiencia posible es 2 y en R = 4 la eficiencia es 1,6.

Preguntas de estudio (15.4)

1. ¿Por qué la eficiencia de Pittman aumenta a medida que **R** crece?
2. ¿Qué indica el cuadro anterior sobre la eficiencia del equiparamiento 2 a 1 en comparación con la formación de pares?

Resumen

- Para el **equiparamiento R a 1**, hay poca ganancia en precisión si se elige una R > 4.
- Un enfoque que se usa para evaluar qué tan grande debe ser R es el **criterio de eficiencia de Pittman**, que compara la precisión del equiparamiento R a 1 con relación a la precisión del emparejamiento.
- Un cuadro de los valores de la **eficiencia de Pittman** para diferentes valores de **R** indica un menor retorno con respecto a la eficiencia una vez que R supera 4.



Auto-evaluación (15.5)

Verdadero o **falso**.

1. Si equiparamos en un estudio de casos y controles, el grupo índice está compuesto por individuos expuestos. . . . . **???**
2. Si equiparamos en un estudio de cohorte, el grupo de comparación está compuesto por no casos. . . . . **???**
3. Si equiparamos individualmente por categorías con respecto a la edad y el sexo en un estudio de casos y controles, entonces el control para un caso dado debe estar en la misma categoría de edad o debe tener el

mismo sexo que el caso. . . . . ???

4. Cuando se efectúa el equiparamiento por frecuencias con respecto a edad y raza en un estudio de casos y controles, la distribución de la edad de los controles se restringe de modo que sea igual que la distribución de la edad de los casos. . . . . ???
5. El equiparamiento 5 a 1 dará como resultado un cálculo más preciso de efecto que el que se obtenga con el equiparamiento 4 a 1 para la misma cantidad de casos. . . . . ???
6. El equiparamiento R_i a 1 puede surgir del intento de realizar un equiparamiento R a 1. . ???
7. La formación de pares es un caso especial de equiparamiento R_i a S_i. . . . ???
8. No se puede ganar mucha precisión si se elige más de un control por caso. . . ???

15-2 Equiparamiento (continuación)

Razones a favor del equiparamiento

¿Por qué deberíamos valernos del equiparamiento para controlar las variables externas cuando se diseña un estudio epidemiológico? La principal ventaja del equiparamiento es que se puede emplear para ganar **precisión** al calcular la medida de efecto de interés. El equiparamiento puede permitirle obtener un **intervalo de confianza más acotado** alrededor de la medida de efecto que el que podría lograr sin aplicarlo.

¿Por qué el equiparamiento ?

- Obtención de mayor precisión
- Control de variables difíciles de medir
- Aspectos prácticos: conveniencia, ahorro de tiempo, ahorro de dinero
- Posibilidad de controlar la confusión (?)

Otra razón para utilizar el equiparamiento es controlar variables que son difíciles de medir. Por ejemplo, el equiparamiento con respecto al barrio de residencia brindaría una forma de control de la clase social, que es difícil de medir como variable de control. Equiparar con respecto a las personas de la misma familia, por ejemplo, hermanos y hermanas, podría ser un modo práctico para controlar los factores genéticos, sociales y ambientales que, de otra forma, serían difíciles de medir.

Un tercer motivo para usar el equiparamiento es aprovechar las características prácticas de la recolección de datos, incluso los aspectos de conveniencia, así como el ahorro de tiempo y dinero. Por ejemplo, si los casos provienen de distintos hospitales, puede ser útil elegir a los controles equiparando con respecto al hospital del caso en el mismo momento en que se identifiquen a los casos en los registros del hospital. En un estudio ocupacional que incluya a distintas empresas de la misma industria, los controles pueden equipararse de manera conveniente con casos de la misma empresa. Dichos controles probablemente tendrán características sociales y ambientales similares a los casos.

Otra causa que a menudo justifica el equiparamiento es el control de **confusión**. En la imagen anterior hemos colocado un signo de interrogación después de este ítem en la lista de viñetas porque, aunque no se utilice el equiparamiento, la confusión puede controlarse mediante el análisis estratificado o la modelación matemática. Asimismo, si usa el equiparamiento en un estudio de casos y controles, debe asegurarse de realizar lo que luego describiremos como **análisis equiparado** para poder controlar adecuadamente la confusión.

Normalmente el equiparamiento se limita a un conjunto restringido de variables de control. En general, hay otras variables no involucradas en el equiparamiento que desearíamos controlar. El equiparamiento no

descarta el control de confusión de esos factores de riesgo que se miden y que no se equiparan.

Resumen

Las razones a favor del equiparamiento incluyen:

- obtención de mayor precisión
- control de variables difíciles de medir
- aspectos prácticos: conveniencia, ahorro de tiempo, ahorro de dinero
- posibilidad de controlar confusión, tanto para las variables equiparadas como las no equiparadas

Razones contra el equiparamiento

¿Por qué podríamos decidir no emplear el equiparamiento cuando diseñamos un estudio epidemiológico? Uno de los motivos es que equiparar con respecto a un factor de riesgo débil o un factor que no presenta riesgo brinda pocas probabilidades de ganar precisión e incluso ésta podría perderse, en comparación con lo que sucedería si no utilizáramos el equiparamiento. Si todas las variables potenciales de control son, en el mejor de los casos, débiles factores de riesgo, el uso del equiparamiento no va a brindar un aumento de la precisión.

Preguntas de estudio (P15.6)

Suponga que equipara con respecto al color de cabello en un estudio de casos y controles que analiza la exposición ocupacional al cáncer de vejiga.

1. ¿Cree que el color de cabello es un factor de riesgo de contraer cáncer de vejiga?
2. Sobre la base de su respuesta a la pregunta anterior, ¿esperaría ganar precisión en su cálculo equiparando con respecto al color de cabello? Desarrolle su respuesta.

Otra causa para no utilizar el equiparamiento es el costo de tiempo y trabajo que implica encontrar las combinaciones adecuadas, especialmente cuando se recurre al equiparamiento individual. En verdad, para efectuar el equiparamiento, se debe preparar un archivo que enumere los controles potenciales y sus valores con respecto a todas las variables que se equiparan; además, se debe especificar y aplicar un procedimiento de selección para equiparar los controles con los casos. Esto consume tiempo y dinero que no se requerirían si los controles se eligieran mediante un muestreo aleatorio de una población de origen.

Una tercera razón para no utilizar el equiparamiento es evitar la posibilidad de lo que se denomina **sobre-equiparamiento**. Esta situación se puede presentar si una o más variables de equiparamiento guardan una gran correlación con, o se aproximan a, la variable de exposición de interés. Por ejemplo, en un estudio ocupacional, si equiparamos con respecto al cargo laboral, y éste es una aproximación de la exposición bajo estudio, entonces sería como si equiparásemos con respecto a la variable de exposición. Es decir, cuando sobre-equiparamos, efectivamente equiparamos artificialmente con respecto a la exposición, lo que llevaría a no encontrar ningún efecto de la exposición sobre la enfermedad, incluso si dicho efecto estuviese realmente presente.

Preguntas de estudio (P15.6) (continuación)

3. Si la variable de exposición es el nivel de colesterol, ¿cómo podría presentarse el **sobre-equiparamiento** si se equipara con respecto a una amplia variedad de características dietarias, incluso la cantidad promedio de productos de comida rápida que registra la dieta de una persona?

Otra desventaja del equiparamiento es que el tamaño del estudio puede reducirse si **no** se pueden encontrar combinaciones para algunos individuos índice. La precisión que esperaba ganar con el equiparamiento podría verse comprometida por dicha reducción en el tamaño planeado del estudio.

Preguntas de estudio (P15.6) (continuación)

4. ¿Qué problema relacionado con el tamaño del estudio podría plantearse si realiza un equiparamiento por categorías con respecto a diversas variables valiéndose de rangos de categorías muy acotados?

Resumen

Razones **contra** el equiparamiento:

- Es poco probable que el equiparamiento con respecto a factores de riesgo débiles otorgue precisión, incluso puede provocar una pérdida de ésta.
- El equiparamiento puede ser costoso en términos del tiempo y dinero requeridos para realizar el proceso de equiparamiento.
- Es posible que sobre-equipare de forma inadecuada y, por lo tanto, equipare efectivamente con respecto a la exposición.
- Le puede resultar difícil encontrar combinaciones; en consecuencia, es posible que comprometa el tamaño muestral y, así, pierda la precisión que esperaba ganar con el equiparamiento.

¿Equiparar o no equiparar?

¿Cómo decidimos si debemos recurrir o no al equiparamiento cuando diseñamos un estudio epidemiológico? Y si resolvemos utilizarlo, ¿cómo establecemos con respecto a qué variables equiparar? La respuesta a ambas preguntas es "**depende**". Consideremos la lista de razones a favor y en contra del equiparamiento que describimos en las actividades anteriores:

Razones a favor del equiparamiento	Razones contra el equiparamiento
Obtención de mayor precisión	Equiparamiento con respecto a factores de riesgo débiles otorgue precisión
Control de variables difíciles de medir	Puede ser costoso en términos del tiempo y dinero
Aspectos prácticos	Es posible que sobre-equipare de forma inadecuada
Posibilidad de controlar la confusión, tanto para las variables equiparadas como no equiparadas	Puede resultar difícil encontrar combinaciones

Su determinación sobre emplear o no el equiparamiento debe basarse en una revisión cuidadosa de los puntos en ambas columnas de esta lista y en cómo sopesa estos diferentes motivos en el contexto del estudio que está planeando.

Preguntas de estudio (P15.7)

Suponga que el costo de tiempo y dinero que implica efectuar el equiparamiento supera los aspectos prácticos de llevarlo a cabo. Presuma además que se desconoce si los factores de riesgo identificados previamente para el resultado de salud son predictores muy fuertes de este resultado.

1. ¿Debería valerse del equiparamiento?

Suponga que edad y tabaquismo se consideran factores de riesgo muy fuertes para el resultado de salud.

2. ¿Debería equiparar o no con respecto a la edad y el tabaquismo?

Suponga que quiere controlar factores sociales y ambientales.

3. ¿Debería equiparar o no con respecto a dichos factores?

Si bien todos los puntos a favor y en contra del equiparamiento son importantes, la principal razón estadística para equiparar es ganar precisión. Los primeros puntos de ambas listas se relacionan con la precisión y sugieren que la posibilidad de ganar o no precisión con el equiparamiento depende del conocimiento previo que tenga el investigador sobre las relaciones importantes entre la enfermedad, la exposición y las posibles factores de confusión. Si dicho conocimiento previo está disponible, por ejemplo, a través de la literatura y se utiliza en forma adecuada, se puede tomar una decisión razonable sobre el equiparamiento.

En lo que respecta a la precisión, se recomienda ampliamente que la estrategia más segura sea equiparar sólo con respecto a fuertes factores de riesgo que, según se espera, surjan como factores de confusión en un estudio. Esta recomendación requiere, claramente, un juicio subjetivo relativo a lo que probablemente sucederá en el estudio con respecto a la distribución de factores potenciales de confusión. En la práctica, la decisión de utilizar el equiparamiento para ganar precisión se aplica sólo a aquellos factores identificados en la bibliografía como fuertes predictores del resultado de salud.

Resumen

- La respuesta a la pregunta "¿equiparar o no equiparar?" es "**depende**".
- Su decisión queda sujeta a una revisión cuidadosa de las razones a favor y en contra del equiparamiento y la manera en que sopesa esos distintos motivos.
- La posibilidad de ganar o no precisión depende del conocimiento previo que tengan los investigadores sobre las relaciones de las variables que se miden.
- Recomendación sobre la precisión: sólo equipare con respecto a fuertes factores de riesgo que se espera que surjan como factores de confusión en el estudio.



Auto-evaluación (15.8)

Verdadero o **falso**.

1. Una razón para optar por equiparar en un estudio de casos y controles es obtener un cálculo válido del *odds ratio* de interés. . . . . **???**

2. Una ventaja de equiparar en comparación con no hacerlo es que el tamaño muestral puede ser más pequeño si no usa el equiparamiento. . . . . **???**
3. El equiparamiento con respecto a débiles factores de riesgo puede provocar una pérdida de precisión cuando se compara con el no equiparamiento. . . . . **???**

Complete los espacios vacíos

4. ¿Cuáles de las siguientes alternativas son razones para no utilizar el equiparamiento? **???**
   a. Se equipara con respecto a un factor que no presenta riesgo.
   b. Se busca controlar una variable difícil de medir.
   c. Se desea controlar tanto las variables equiparadas como las no equiparadas.
   d. La variable de equiparamiento está muy correlacionada con la variable de exposición.
   e. Es costoso efectuar el equiparamiento.

Opciones
solo a **a, b y c** **a, b y d** **a, d y e** **solo b** **b, c y d** **solo c** **solo d** **solo e**

5. Si el equiparamiento es práctico y poco costoso, siempre se debería preferir a no equiparar. **???**
6. Si la razón principal para considerar el equiparamiento es ganar precisión en el cálculo del *odds ratio*, ¿debería equiparar o no en un estudio de casos y controles? . . . **???**

Opciones
Falso **Depende** **Verdadero** **No equiparar si es costoso** **Equiparar siempre** **Equiparar si es práctico** **Nunca equiparar**

15-3 Equiparamiento (continuación)

Análisis de datos equiparados – Opciones y principios generales

Existen dos alternativas para analizar datos equiparados con resultados dicotómicos: el **análisis estratificado** usando los métodos de **Mantel-Haenszel** y la **modelación matemática** empleando la **regresión logística**. Los métodos de Mantel-Haenszel son adecuados siempre que todas las variables que se controlen estén involucradas en el equiparamiento. Los métodos de regresión logística son apropiados si no se ha equiparado con respecto a algunas variables que se controlan, pero sí con respecto a otras.

Dos opciones para analizar datos equiparados:

Análisis estratificado
(usando los métodos de Mantel-Haenszel)

- Todas las variables que se controlen estén involucradas en el equiparamiento

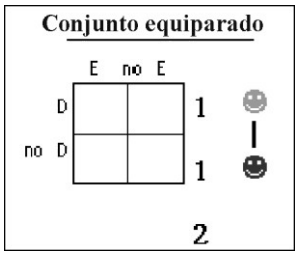
Modelación matemática
(empleando la regresión logística)

- No se ha equiparado con respecto a algunas variables que se controlan

Por ejemplo, si equiparamos con respecto a edad, raza y sexo en un estudio de casos y controles y estas tres variables son las únicas que se consideran para el control, entonces los métodos de Mantel-Haenszel para el análisis estratificado son convenientes. Por el contrario, si equiparamos con respecto a la edad, la raza y el sexo y también queremos controlar otras variables no equiparadas, como el nivel de actividad física, el tamaño corporal y la presión arterial, entonces es necesario emplear métodos de regresión logística.

Cuando se realiza un análisis equiparado, se deben considerar **cuatro** principios importantes. **En primer lugar**, el análisis equiparado requiere que, en verdad, "controle" las variables que equipara. En especial, si no controla dichas variables en un estudio de casos y controles, no habrá abordado la confusión que surge a raíz de estas variables y su *odds ratio* calculado estará sesgado hacia el valor nulo. Además, si no controla las variables que se equiparan en un estudio de seguimiento, probablemente no gane la precisión que pretendía lograr en su cálculo de la razón de riesgos mediante el equiparamiento.

En segundo lugar, el **análisis equiparado** es un **análisis estratificado**. Los estratos son los conjuntos equiparados o conjuntos equiparados combinados. Por ejemplo, si forma pares en un estudio de casos y controles y posee 100 casos, entonces hay 100 conjuntos o estratos equiparados para analizar. Cada conjunto equiparado contendría dos personas, el caso y el control.



En tercer lugar, cuando se recurre a la regresión logística para efectuar el análisis equiparado, los estratos se definen mediante variables ficticias o indicadoras. El número de variables ficticias será uno menos que la cantidad de estratos que se equiparan. Por ejemplo, si formamos pares en un estudio de casos y controles y tenemos 100 casos, un modelo logístico para dichos datos requerirá 99 variables ficticias para incorporar los 100 estratos que se equiparan.

Preguntas de estudio (P15.9)

1. Enuncie la forma logit de un modelo logístico que permita el análisis de 100 pares formados por un caso y un control para describir la relación entre una variable de exposición dicotómica, E, y una variable de resultado dicotómica, D.

En cuarto lugar, una ventaja clave de adoptar el modelo logístico con datos equiparados es que puede controlar las variables involucradas en el equiparamiento y las variables que no lo están. Si, por el contrario, aplica el análisis estratificado, normalmente tendrá que dejar de lado algunos estratos equiparados en el análisis y, en consecuencia, perderá precisión en su cálculo.

Preguntas de estudio (P15.9) (continuación)

Suponga que equipara con respecto a la edad, la raza y el sexo en un estudio de casos y controles que incluye 100 pares equiparados. Para su análisis, desea controlar la presión arterial sistólica (PAS) y el nivel de colesterol (COL), ninguna de las cuales se encuentra incluida en el equiparamiento; además, quiere controlar las variables con respecto a las cuales se equipara.

2. Indique la forma logit de un modelo logístico para realizar el análisis descrito anteriormente.
3. ¿Qué información se perderá si se efectúa un análisis estratificado para controlar las variables del equiparamiento, como también PAS y COL?

Resumen

- Dos opciones para analizar los datos equiparados son el análisis estratificado aplicando los métodos de Mantel-Haenszel y la modelación matemática mediante la regresión logística.
- El análisis equiparado requiere que controle las variables que se equiparan.
- El análisis equiparado es un análisis estratificado.
- Cuando se recurre a la regresión logística para efectuar el análisis equiparado, los estratos se definen mediante variables ficticias (es decir, indicadoras).
- Cuando se adopta el modelo logístico con datos equiparados, se pueden controlar las variables involucradas en el equiparamiento y las variables que no lo están.

¿El equiparamiento controla la confusión?

Claramente, si deseamos controlar las variables con respecto a las cuales no se equiparó, además de las variables involucradas en el equiparamiento, la respuesta a esta pregunta es **no**. Sin embargo, si **suponemos que las únicas variables que se controlan están comprendidas en el equiparamiento**, entonces la respuesta requiere que consideremos en forma separada los estudios de cohorte, al igual que los de casos y controles.

En un estudio de cohorte, el equiparamiento automáticamente controla la confusión sin necesidad de controlar las variables que se equiparan. Sin embargo, aún debe controlar el equiparamiento para ganar la precisión que esperaba obtener al equiparar (suponiendo que tomó una buena decisión sobre qué variables involucrar en el equiparamiento).

En un estudio de casos y controles, el equiparamiento no controla automáticamente la confusión, por lo cual es necesario controlar las variables que se equiparan para poder controlar la confusión.

Para explicar aún más las afirmaciones anteriores, debemos describir **las condiciones de confusión basada en los datos**. A fin de explicarlo en forma simple, supondremos que sólo se equipara con respecto a una variable, denotada aquí como **F**. Anteriormente, hemos establecido que hay confusión basada en los datos si:

$$c\hat{\theta} \neq a\hat{\theta}$$

donde $c\hat{\theta}$ y $a\hat{\theta}$ denotan el cálculo crudo y cálculo ajustado, respectivamente, de la medida de efecto θ. También pueden darse condiciones de confusión equivalentes a la descrita para una razón de riesgos (**RR**) y para un *odds ratio* (**OR**), respectivamente, en términos de la relación entre la variable de enfermedad **D** y la variable de equiparación **F**, y la relación entre la variable de exposición **E** y la variable de equiparación **F**.

Condiciones para la confusión de **RR**:

$$\hat{OR}_{DF|noexpuestos} \neq 1 \text{ y } \hat{OR}_{EF} \neq 1$$

Condiciones para la confusión de **OR**:

$$\hat{OR}_{DF|noexpuestos} \neq 1 \text{ y } \hat{OR}_{EF|sanos} \neq 1$$

Si aplicamos las condiciones para la confusión en la **RR** a un estudio de cohorte equiparado, podemos observar que estas condiciones **no se satisfacen**, ya que

¿El equiparamiento controla la confusión? (continuación)

$$\widehat{OR}_{EF} = 1$$

Es decir, un estudio de cohorte equiparado hace que los diferentes grupos de exposición tengan la misma distribución con respecto a la variable de equiparamiento **F**. Por lo tanto, el equiparamiento en un estudio de cohorte controla automáticamente la confusión en la razón de riesgos, es decir, la razón de riesgos cruda siempre será igual a la razón de riesgos ajustada; por ende, podemos ignorar el control del equiparamiento y aún obtener una razón de riesgos correcta.

Por otro lado, si aplicamos las condiciones para la confusión del OR a un estudio de casos y controles equiparados, podemos ver que:

$$\widehat{OR}_{DF} = 1$$

que no es equivalente al siguiente requisito:

$$\widehat{OR}_{DF|no\,expuestos} = 1$$

Es decir, aunque el estudio de casos y controles equiparados hace que los diferentes grupos de enfermedad tengan la misma distribución con respecto a la variable de equiparamiento **F**, no surge automáticamente que el *odds ratio* que relaciona **F** con la enfermedad **D**, dado el hecho de no estar expuesto, sea igual a 1. Tampoco implica que el *odds ratio* que relaciona **F** con la exposición **E**, dado el hecho de no estar enfermo, sea igual a 1. Por lo tanto, el equiparamiento en un estudio de casos y controles no controla automáticamente la confusión del *odds ratio*.

Asimismo, puede demostrarse (se omiten aquí los detalles) que el ***odds ratio* crudo para los datos de casos y controles equiparados siempre está sesgado hacia el valor nulo de 1**. Por ende, es necesario controlar la variable involucrada en el equiparamiento **F** (mediante el análisis estratificado o la regresión logística), ya que ignorar el control de **F** (mediante un *odds ratio* crudo) dará una respuesta sesgada que tenderá hacia la conclusión de que hay ausencia de efecto.

¿Cuáles son las consecuencias de no realizar un análisis equiparado (datos de casos y controles)?

Hay dos formas de no llevar a cabo el análisis equiparado:

1. **ignorar el equiparamiento**
2. **romper el equiparamiento**

1. Si equiparamos con respecto a una o más variables, pero analizamos los datos resultantes del estudio sin controlar ninguna de las variables equiparadas, entonces estamos **ignorando el equiparamiento**. Como ejemplo, suponga que tenemos variables **E** y **D** dicotómicas, 100 casos y formamos pares con respecto a la edad y el sexo. Si ignoramos el equiparamiento y el control de cualquier otro factor de riesgo con respecto al cual no se está equiparando, entonces, efectivamente, estamos efectuando un "análisis crudo" de los datos, es decir, nuestro cálculo es un *odds ratio* crudo, o **cOR** con acento circunflejo. Existen dos críticas fundamentales con respecto a ignorar el equiparamiento:
   a. El **cOR** calculado siempre está sesgado hacia el valor nulo de 1 en los estudios equiparados (de casos y controles y de cohorte). Por lo tanto, se espera que el **cOR** calculado brinde un *odds ratio* distinto

(sesgado) del *odds ratio* (es decir, el **mOR**) previsto en un análisis equiparado.

b. Si el equiparamiento cumple su función (es decir, mejora la precisión), se espera que el cálculo del **mOR** brinde mayor precisión que el cálculo del **cOR** correspondiente.

¿Cuáles son las consecuencias de no realizar un análisis equiparado (datos de casos y controles)? (Continuación)

2. Si equiparamos, pero controlamos las variables equiparadas sin realizar un análisis equiparado, entonces estamos **rompiendo el equiparamiento**. Como ejemplo, suponga que como anteriormente, tenemos: (1) variables **E** y **D** dicotómicas; (2) 100 casos; y (3) pares con respecto a edad y sexo. Si rompemos el equiparamiento, entonces controlamos edad y sexo formando estratos a partir de combinaciones de estas 2 variables, y luego realizamos un análisis estratificado. Es probable que la cantidad de estratos resultantes sea considerablemente menor a 100; por ejemplo, si edad posee 3 categorías y sexo 2, entonces la cantidad de estratos es 6. Si no rompemos el equiparamiento, entonces controlamos edad y sexo tratando a cada conjunto equiparado como un estrato único. Ya que hay 100 pares de casos y controles, el análisis equiparado, entonces, sería un análisis estratificado que incluiría 100 estratos con 2 personas en cada estrato.

   Argumentos a favor de romper el equiparamiento:
   En el ejemplo anterior, donde edad posee 3 categorías y sexo 2, romper el equiparamiento equivale a agrupar conjuntos equiparados intercambiables, lo que es más adecuado que suponer que hay 100 estratos distintos involucrados en el equiparamiento. [Véase la actividad sobre "agrupamiento" de la sección 15-4 de la lección].

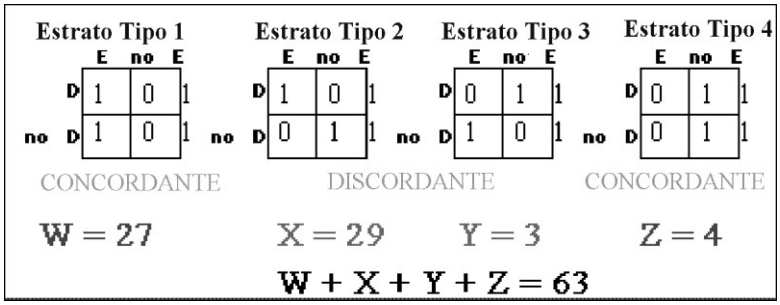
   Argumentos en contra de romper el equiparamiento:
   a. Si se rompe el equiparamiento, es posible que la precisión del *odds ratio* calculado pueda ser menor que la precisión obtenida al realizar un análisis equiparado (con o sin agrupamiento).
   b. Los estratos que surgen al romper el equiparamiento pueden no ser equivalentes a los estratos que se generarían al agrupar conjuntos equiparados intercambiables (el análisis correcto).
   c. Como consecuencia del ítem b, el *odds ratio* obtenido de un análisis estratificado que resulta al romper el equiparamiento puede ser distinto del *odds ratio* obtenido de un análisis agrupado.
   d. Si desea controlar variables que no se equipararon además de variables de equiparamiento, entonces romper el equiparamiento requerirá que desarme conjuntos equiparados que están en categorías distintas de la(s) variable(s) no equiparada(s).

Análisis de datos de casos y controles apareados

*Ahora ilustramos un análisis equiparado mediante datos **de casos y controles apareados (también, pareados o emparejados)**. Los datos se pueden examinar mediante un análisis estratificado para obtener un **mOR**, una **prueba de hipótesis de MH** y un **intervalo de confianza** alrededor del mOR.*

En la década de los setenta, se llevaron a cabo varios estudios para evaluar si el uso de estrógenos como terapia hormonal sustitutiva en mujeres menopáusicas estaba asociado con cáncer de endometrio. Uno de esos estudios empleó **equiparamiento individual** para realizar un **estudio de casos y controles apareados** que comprendía mujeres que vivían en una comunidad de jubilados de Los Ángeles entre 1971 y 1975. Había 63 casos. Los controles se emparejaron individualmente a los casos en función a edad, estado civil y fecha de ingreso a la comunidad para jubilados. Los 63 pares formados representan 63 estratos que contienen 2 personas por estrato. Para cada estrato, formamos el cuadro de 2 x 2 que relaciona la exposición (el uso de estrógenos) y el resultado de salud (cáncer de endometrio). Cada uno de estos estratos puede adoptar una de las cuatro formas que

se muestran a continuación, según el estatus de exposición determinado para el caso y el control en un estrato dado.



El estrato de tipo 1 posee cualquier par donde **tanto** el caso **como** el control están expuestos, es decir, ambos utilizaron estrógenos. Un par formado de esta manera se denomina par **concordante**. Denotamos la cantidad de pares concordantes de este tipo con **W**. En el estudio se encontraron 27 pares de este tipo.

El estrato de tipo 4 está compuesto por aquellos pares donde **ni** el caso **ni** el control usaron estrógenos. Este tipo de estrato también tiene pares concordantes, ya que tanto los casos como los controles tienen el mismo estatus de exposición, en este caso, no expuestos. Denotamos la cantidad de pares **concordantes** de este tipo con **Z**. En el estudio se hallaron 4 pares de este tipo.

Los otros dos tipos de estratos poseen lo que denominamos **pares discordantes**. En el estrato de tipo 2, el caso usó estrógenos, pero el control no. En el estrato de tipo 3, el caso no usó estrógenos, pero el control sí. En estos dos tipos de estratos, el caso tiene una exposición diferente de su control equiparado. Las cantidades de pares discordantes de cada tipo se denominan **X** y **Y**, respectivamente. En el estudio se determinó que **X** era igual a 29 pares y **Y** era igual a 3 pares. Observe que la suma de **W**, **X**, **Y** y **Z** equivale a 63, la cantidad de pares formados en el estudio.

¿Cómo analizamos estos datos? Una respuesta simple a esta pregunta es que utilizamos una computadora para realizar el análisis estratificado de estos 63 estratos para obtener un ***odds ratio* de Mantel-Haenszel**, **una prueba de hipótesis de Mantel-Haenszel** y un **intervalo de confianza de 95%** alrededor del *odds ratio* calculado. Efectuaremos este análisis en la siguiente actividad.

Preguntas de estudio (P15.10)

1. ¿Por qué es necesario calcular un mOR en lugar de un *odds ratio* ajustado basado en la precisión (es decir, aOR)?

Resumen

- Ilustramos un análisis equiparado mediante datos de casos y controles apareados.
- El estudio incluyó 63 pares o estratos con 2 personas por estrato.
- Había 4 tipos de estratos, 2 de los cuales poseían pares **concordantes** y 2 tenían pares **discordantes**.
- **W** = cantidad de pares concordantes donde tanto el caso como el control están expuestos.
- **X** = cantidad de pares discordantes donde el caso está expuesto y el control no.
- **Y** = cantidad de pares discordantes donde el caso no está expuesto y el control sí.
- **Z** = cantidad de pares concordantes donde ni el caso ni el control están expuestos.
- Los datos se pueden examinar mediante un análisis estratificado para obtener un mOR, una prueba de hipótesis de MH y un intervalo de confianza alrededor del mOR.

Análisis de datos de casos y controles apareados (continuación)

Se puede efectuar un análisis estratificado de los 63 estratos con una computadora para obtener un ***odds ratio* de Mantel-Haenszel**, una **prueba de hipótesis de Mantel-Haenszel** y un **intervalo de confianza de 95%** para el **mOR**. Cada uno de los 63 estratos corresponde a uno de los cuatro tipos que se muestran en el cuadro de la actividad anterior. Una forma práctica de realizar este análisis sin usar una computadora es formar el siguiente cuadro mediante las cantidades de pares **concordantes** y **discordantes W**, **X**, **Y** y **Z**.

Cuadro de McNemar (casos y controles apareados)

		Control: E	Control: no E
Caso	E	W	X
	no E	Y	Z

Este cuadro se denomina **cuadro de McNemar para datos de casos y controles apareados**. Los números en este cuadro representan pares de observaciones, en lugar de observaciones individuales. Con este cuadro pueden escribirse fórmulas simples para el mOR, el estadístico de prueba de Mantel-Haenszel y un intervalo de confianza de 95% para el mOR (véase el cuadro después de este párrafo). Observe que todas estas fórmulas solo tienen información sobre las cantidades de pares **discordantes**, **X** y **Y**; los datos sobre pares **concordantes** no se utilizan.

$$\hat{mOR} = \frac{X}{Y}$$

$$\chi^2_{MHS} = \frac{(X - Y)^2}{X + Y}$$

$$\hat{mOR} \exp\left[\pm 1.96\sqrt{\frac{1}{X} + \frac{1}{Y}}\right]$$

Al sustituir los valores de **X** y **Y** en cada una de estas fórmulas obtenemos los resultados que se muestran a continuación:

$$\hat{mOR} = \frac{29}{3} = 9.67$$

$$\chi^2_{MHS} = \frac{(29 - 3)^2}{29 + 3} = 21.13$$

95% IC: **(2.94, 31.73)**

Pregunta de estudio (P15.11)

1. ¿Cómo interpreta los resultados anteriores en términos de la relación entre el uso de estrógenos (E) y el cáncer de endometrio (D) que se abordó en el estudio de casos y controles apareados?

Resumen

- Una forma práctica de realizar un análisis de datos de casos y controles apareados es usar el cuadro de **McNemar** que contiene pares concordantes y discordantes.
- Solo mediante los pares discordantes **X** y**Y**, pueden emplearse fórmulas simples para calcular el mOR, la prueba de hipótesis de MH y el IC del 95% alrededor del mOR.

Diseño de casos alternantes o cruzados

El **estudio de casos alternantes o cruzados** es una variante del **estudio de casos y controles equiparados** diseñado para reducir las probabilidades de sesgo del diseño de casos y controles convencional mediante el método de selección de los controles. El diseño incorpora elementos tanto de un estudio de casos y controles equiparados como de un **estudio cruzado retrospectivo no experimental**. (Nota: en un **diseño cruzado**, cada individuo recibe, al menos, dos exposiciones/tratamientos diferentes en distintas ocasiones). El aspecto fundamental del diseño de casos cruzados es que cada caso sirve como su propio control. Las exposiciones que varían en términos de tiempo se comparan entre los intervalos en que se produjo el resultado (intervalos de casos) y los intervalos en que no se produjo el resultado en el mismo individuo.

El diseño de casos cruzados se creó para evaluar el efecto de exposiciones breves con efectos transitorios en resultados de salud graves cuando no se dispone de un grupo de control tradicional. La principal ventaja del diseño de casos cruzados radica en su capacidad para ayudar a controlar la confusión. Los individuos que se equiparan consigo mismos automáticamente eliminan la confusión entre individuos y la confusión de las covariables fijas medidas y no medidas.

Un ejemplo de diseño de casos cruzados es el estudio de Redelmeier y Tibshirani, quienes intentaron determinar si el uso del teléfono celular al conducir un vehículo motorizado incrementa el riesgo de sufrir un accidente de tránsito. Sus datos consideraron 699 conductores que tenían teléfonos celulares y que participaron en accidentes de tránsito que produjeron daños materiales importantes aunque no lesiones personales. Mediante el uso de registros de facturación detallados, se analizaron las llamadas por celular de cada persona correspondientes al día del accidente y a la semana anterior.

En general, 170 de las 699 personas habían utilizado un teléfono celular durante los 10 minutos anteriores al accidente; 37 personas lo habían usado durante el mismo período el día anterior al choque; y 13 personas lo habían hecho durante el mismo período ambos días. Esta información generó el siguiente cuadro de McNemar para el análisis:

		Día anterior	
Día del accidente	**Usó**	**No usó**	
Usó	13	157	170
No usó	24	505	529
	37	662	699

A partir de estos datos, se obtuvieron los siguientes resultados del análisis equiparado:

$$\hat{\text{mOR}} = \frac{X}{Y} = \frac{157}{24} = 6{,}5 \qquad \text{y } \chi^2 = \frac{(X-Y)^2}{X+Y} = \frac{(157-24)^2}{157+24} = 97{,}7 \ \ (P < 0{,}001)$$

Intervalo de confianza de 95% para el **mOR**:

$$\hat{mOR}\exp\left[\pm 1{,}96\sqrt{\frac{1}{X}+\frac{1}{Y}}\right] = 6{,}5\exp\left[\pm 1{,}96\sqrt{\frac{1}{157}+\frac{1}{24}}\right] = (4{,}5;\,9{,}9)$$

Los resultados anteriores revelan un efecto muy fuerte y significativo que indica que el uso del teléfono celular mientras se conduce aumenta el riesgo de sufrir un accidente de tránsito. Asimismo, el análisis principal, que ajustó el manejo intermitente de vehículos motorizados, arrojó un **mOR** de 6,5 con un intervalo de confianza de 95% de 4,5 a 9,9.

Análisis de datos de casos y controles equiparados R a 1

En un estudio de casos y controles equiparados de 1969 se consideró la relación hipotética entre los antecedentes de aborto inducido y el resultado de embarazo tubárico en mujeres que habían tenido, al menos, un embarazo anterior. En este estudio se recurrió al **equiparamiento individual 4 a 1 de controles con casos**. Había 18 casos, por ende, 18 conjuntos o estratos equiparados que contenían 5 individuos por estrato. La cantidad total de individuos en el estudio fue, por lo tanto, 18 multiplicado por 5, o sea 90.

Se equiparó a los controles por categorías con cada caso con respecto al tipo de embarazo, la edad de la mujer y del marido. A continuación se muestra la disposición de 2 x 2 de los datos para el estrato **i**, donde $\mathbf{Y_i}$ es igual a 1 si el caso en el estrato **i** tuvo anteriormente un aborto y 0 si no abortó, y donde $\mathbf{X_i}$ denota la cantidad de controles, de los 4 posibles, en el estrato **i** que tuvieron un aborto anterior.

Estrato i:

	E	no E	
D	Y_i	$1-Y_i$	1
no D	X_i	$4-X_i$	4

$Y_i = \begin{cases} 1 & \text{si el caso en el estrato i tuvo anteriormente un aborto} \\ 0 & \text{si no abortó} \end{cases}$

$X_i =$ # de controles en el estrato i que tuvieron un aborto anterior

Aquí se resumen los resultados del estudio:

		# de controles que tuvieron un aborto anteriormente				
		$X_i = 4$	$X_i = 3$	$X_i = 2$	$X_i = 1$	$X_i = 0$
Antecedentes de aborto en los casos	$Y_i = 1$	1	0	3	5	3
	$Y_i = 0$	0	0	0	1	5

Los números en el cuerpo del cuadro representan las cantidades de conjuntos o estratos equiparados. Si suma todos estos números, obtendrá 18, que es la cantidad total de conjuntos equiparados en el estudio, donde cada conjunto equiparado tiene la forma general que se mostró para el estrato **i**. Cada número en el cuerpo del cuadro brinda la cantidad de conjuntos equiparados para los cuales $\mathbf{X_i}$ de los controles tuvieron un aborto anteriormente y para los cuales el caso ha tenido o no un aborto anterior. Por ejemplo, el 1 en la celda de arriba a la izquierda indica que 1 de los 18 conjuntos equiparados comprendía a los 4 controles con abortos anteriores y que el caso también había sufrido un aborto previo. Por el contrario, el 5 en la celda de abajo a la derecha indica que hubo 5 conjuntos equiparados en los que ninguno de los 4 controles ni el caso tuvieron un aborto previo.

¿Cómo analizamos estos datos? La respuesta es que usamos una computadora para realizar un análisis estratificado de los 18 estratos que componen este conjunto de datos para obtener un *odds ratio* de Mantel-Haenszel, una prueba de hipótesis de Mantel-Haenszel y un intervalo de confianza de 95% para el **mOR**. Estos son los resultados calculados.

Análisis estratificado

$\hat{mOR} = 33.00$

prueba MH = 16.0 (P < .0001)

95 % IC para mOR: (2.95, 369.17)

Preguntas de estudio (P15.12)

1. ¿Por qué aquí se emplea el mOR en lugar de un aOR basado en la precisión?
2. Ya que hay solo 5 individuos en cada estrato (es decir, es un conjunto equiparado), ¿cómo puede justificar el uso del estadístico de prueba de MH como un estadístico chi cuadrado de muestra grande?
3. ¿Qué conclusión puede derivar de estos resultados?

Resumen

- Al igual que con los datos apareados, el análisis de **datos equiparados R a 1** es un análisis estratificado, aunque ahora hay **R + 1** individuos en cada estrato.
- El cálculo del **mOR**, la **prueba de MH** y el intervalo de confianza de 95% para el **mOR** se realiza mejor con un programa informático.
- El estrato **i**-avo tiene la disposición de datos que se muestra más abajo, donde $\mathbf{Y_i}$ denota la cantidad de casos (1 ó 0) que están expuestos y $\mathbf{X_i}$ denota la cantidad de controles de un total de **R** que están expuestos en el estrato **i**.

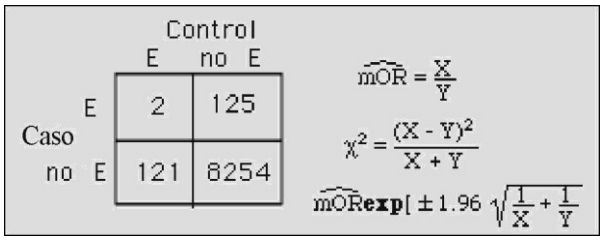


Auto-evaluación(P15.13)

Verdadero o **falso**.

1. El análisis equiparado se puede realizar mediante un análisis estratificado en el que los estratos están formados por el grupo de conjuntos equiparados o por conjuntos equiparados agrupados. **???**
2. En un estudio de casos y controles apareados, el mOR se calcula como: (X - Y)/(X + Y), donde X y Y son frecuencias de pares discordantes. . . . . **???**
3. En un estudio de casos y controles apareados suponga que la cantidad de pares en cada una de las 4 celdas que se emplean para la prueba de McNemar está dada por W = 50, X = 30, Y = 15, 2 = 100. Entonces, el valor calculado del estadístico de la prueba de McNemar es 2. . . **???**
4. El estadístico de la prueba de Mantel-Haenszel no es adecuado cuando existe equiparamiento R a 1 y R es, al menos, igual a 2. . . . . **???**

Considere el siguiente cuadro de McNemar extraído del estudio "Agente naranja" (Donovan et al., Med. J. Aus., 1984). Se trata de un estudio de casos y controles apareados. Los casos son bebés que nacieron con anomalías genéticas y los controles son bebés que nacieron sin dichas anomalías. El factor de exposición es la condición del padre (1 = veterano de la guerra de Vietnam, 0 = no veterano).



5. El mOR a partir de este cuadro es **???**
6. El estadístico de la prueba de Mantel-Haenszel es **???**
7. Un IC de 95% para el mOR está dado por los límites **???**
8. Sobre la base de los datos, ¿los veteranos de la guerra de Vietnam tienen más probabilidades de tener bebés con anomalías genéticas? **???**

Opciones
(0,76; 1,25) **(0,80; 1,32)** **0,07** **0,97** **1,03** **3,98** **no** **sí**

15-4 Equiparamiento (continuación)

Agrupamiento de datos equiparados

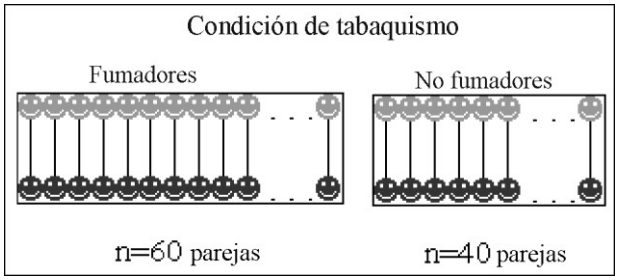
Suponga que la condición de tabaquismo, definida en función de si fumó **alguna vez** versus **nunca** fumó, es la única variable que se equipara en un **estudio de casos y controles apareados** que contiene 100 casos. Presuma, además, que cuando se efectúa el equiparamiento, 60 de los pares contienen 2 fumadores y 40 de los restantes pares tienen 2 no fumadores.

Ahora consideremos 2 pares cualesquiera de los que poseen fumadores, por ejemplo, el par A y el par B. Como el tabaquismo es la única variable que se equipara, el control en el par A cumplía con los requisitos para haber sido elegido como control para el caso en el par B antes del proceso de equiparamiento. De manera similar, el control fumador en el par B podría haber sido elegido para ser el control fumador para el caso en el par A.

Aunque esto no sucedió, en realidad, después del equiparamiento, la **intercambiabilidad** potencial de estos dos controles sugiere que los pares A y B no deberían ser tratados como estratos separados en un análisis equiparado. Los conjuntos equiparados como los pares A y B se denominan **conjuntos equiparados intercambiables**. Para todo el estudio que comprendió a 100 pares equiparados, los 60 pares que incluyen a todos los fumadores son intercambiables, y los restantes 40 pares de no fumadores son intercambiables en forma separada.

Si ignoramos la intercambiabilidad, el análisis típico de estos datos sería un análisis estratificado que trata los 100 pares equiparados como 100 estratos separados. El análisis se podría llevar a cabo, entonces, con la información de los pares discordantes en el cuadro de **McNemar**, como se describió anteriormente. Pero, en verdad, ¿deberíamos ignorar la intercambiabilidad de los conjuntos equiparados? Nuestra respuesta es negativa, principalmente porque tratar estratos intercambiables en forma separada supone artificialmente que dichos estratos son únicos cuando, en realidad, no lo son.

Por lo tanto, ¿cómo debería realizarse el análisis? La respuesta aquí es agrupar conjuntos equiparados intercambiables. En nuestro ejemplo, el agrupamiento significaría que en lugar de analizar 100 estratos distintos con 2 personas por estrato, el análisis consideraría solo dos estratos agrupados, uno agrupando 60 conjuntos equiparados en un estrato de fumadores y el otro agrupando los otros 40 conjuntos equiparados en un estrato de no fumadores.



Preguntas de estudio (P15.14)

Considere un estudio de casos y controles con equiparamiento 2 a 1 por categorías con respecto a grupos etarios (digamos, menos de 60 años versus 60 años o más) y de la condición de tabaquismo (fumó alguna vez versus nunca fumó).

1. ¿Cuáles son los conjuntos equiparados intercambiables en este estudio?

Considere un estudio diferente de casos y controles con equiparamiento 2 a 1 por categorías con respecto al grupo etario, la condición de tabaquismo, el sexo, el nivel de actividad física, el tamaño corporal, la raza y el grupo étnico.

2. ¿El análisis debería requerir el agrupamiento de los conjuntos equiparados?

Resumen

- Dos o más conjuntos equiparados son **intercambiables** si tienen la misma combinación de categorías de equiparamiento.
- Se debe preferir el **análisis agrupado** al **no agrupado** siempre que haya conjuntos equiparados intercambiables.
- El **análisis agrupado** implica combinar conjuntos equiparados intercambiables en un único estrato **agrupado** y luego realizar un análisis estratificado para los estratos agrupados.
- El agrupamiento de conjuntos equiparados intercambiables es adecuado porque mantener estratos intercambiables separados supone artificialmente que dichos estratos son únicos cuando, en realidad, no lo son.
- Si se equiparan diversas variables, los resultados del **análisis agrupado** deben ser apenas diferentes de los resultados del **análisis no agrupado**.

Agrupamiento versus equiparamiento por frecuencias

Aquí abordamos tres preguntas:

1. ¿El **agrupamiento de conjuntos equiparados intercambiables** es equivalente al **equiparamiento por frecuencias**?
2. Si decide agrupar conjuntos equiparados intercambiables, ¿las cantidades y definiciones de los estratos resultantes son idénticos a los estratos que se definirían si se recurre, en su lugar, al equiparamiento por frecuencias?

3. ¿Los resultados que surgen al analizar datos equiparados agrupados siempre serán idénticos a los que se obtienen cuando se usa, en su lugar, el equiparamiento por frecuencias?

La respuesta a cada una de las tres preguntas es **no**. Sin embargo, según ciertos supuestos sobre la forma en que se realiza el equiparamiento, la respuesta a la segunda pregunta podría ser afirmativa.

Respuesta al ítem 1: el agrupamiento requiere **equiparamiento individual**, que es un procedimiento para seleccionar individuos de comparación distintos del equiparamiento por frecuencias. En particular, la cantidad de estratos de equiparamiento resultantes del agrupamiento, la manera en que se definen los estratos y los individuos que se seleccionan dentro de cada estrato pueden no ser idénticos a los elementos correspondientes de un estudio equiparado por frecuencias (véase a continuación).

Agrupamiento versus equiparamiento por frecuencias (continuación)

Respuesta al ítem 2: si una o más de las variables que se equiparan son continuas (por ejemplo, edad), entonces el modo en que se define la intercambiabilidad puede ser distinto del método con el cual podrían establecerse las categorías equiparadas por frecuencias. Por ejemplo, si se equipara individualmente con respecto a la edad, se pueden emplear y considerar intercambiables las categorías de edad más acotadas (por ejemplo, un rango de 5 años) que las elegidas si se utilizara, en su lugar, el equiparamiento por frecuencias (por ejemplo, un rango 15 años). En dicha situación, no solo la cantidad de estratos para un análisis agrupado sería diferente de la cantidad de estratos para los datos equiparados por frecuencias, sino que las definiciones de los estratos también serían distintas.

Por otro lado, si a) se hubiesen definido exactamente las mismas categorías de equiparamiento (por ejemplo, rango de 5 años) tanto para el equiparamiento individual como para el equiparamiento por frecuencias, o b) si la definición de intercambiabilidad se ampliara para que fuera equivalente a las definiciones de las categorías usadas para el equiparamiento por frecuencias, entonces tanto la cantidad como los tipos de estratos serían equivalentes para el análisis agrupado y para el equiparado por frecuencias.

Respuesta el ítem 3: los estratos que resulten de agrupar conjuntos equiparados intercambiables pueden no ser equivalentes a los estratos que se obtengan con el equiparamiento por frecuencias.

Lo ilustramos con el siguiente ejemplo:

Suponga que en un estudio de casos y controles equiparamos solo con respecto a una variable, por ejemplo, la condición de tabaquismo (fumó alguna vez versus nunca fumó). Si nos valemos del **equiparamiento individual** y hay 100 casos, entonces los estratos originales constarán de 100 pares equiparados, donde cada par tendrá un caso y un control con la misma condición de tabaquismo. Ya que solo una variable está involucrada en el equiparamiento, todos los pares equiparados que contienen individuos que fumaron alguna vez son intercambiables y todos los pares equiparados que incluyen individuos que nunca fumaron también son intercambiables (en forma separada). Si, como se recomendó, **agrupamos conjuntos equiparados intercambiables**, entonces terminamos con solo dos estratos: uno para los que fumaron alguna vez y otro para los que nunca fumaron.

Si en contraposición al equiparamiento individual con respecto al tabaquismo, **equiparamos por frecuencias** con respecto al tabaquismo con los mismos 100 casos e igual tamaño muestral total, también terminaríamos con los mismos dos estratos que los obtenidos del agrupamiento, siempre que elijamos los mismos individuos de control que hubiésemos elegido en el equiparamiento por frecuencias.

Sin embargo, aun cuando seleccionemos el mismo número (100) de controles, esto no garantiza que elegiremos a los mismos individuos del grupo de controles posibles que habríamos logrado con el equiparamiento individual. En consecuencia, los dos estratos obtenidos del equiparamiento por frecuencias pueden contener diferentes individuos, si se comparan a los dos estratos generados con el equiparamiento individual. Es probable que el análisis equiparado resultante conduzca a resultados numéricos distintos (aunque no necesariamente a distintas conclusiones de inferencia estadística).

Como segundo ejemplo, suponga que nuevamente tenemos 100 casos para un estudio de casos y controles. Esta vez, **equiparamos individualmente** con respecto a 5 variables, por ejemplo, edad, raza, sexo, tabaquismo,

índice de masa corporal (categorizado en obeso y no obeso) y clase social (en tres categorías). Entonces, dos pares equiparados son intercambiables si tanto el caso como el control están en la misma categoría de edad, raza, sexo, tabaquismo, índice de masa corporal y clase social. Como aquí hay 5 variables involucradas en el equiparamiento, no es probable que haya muchos estratos intercambiables, aunque ciertamente puede haber algunos. De hecho, hasta puede no ser posible encontrar un par para ciertos casos. Asimismo, si **agrupamos estratos intercambiables**, no es probable que la cantidad total de estratos resultantes se reduzca considerablemente en comparación al tamaño original de 100 estratos no agrupados.

Ahora suponga que **equiparamos por frecuencias** los mismos 100 casos, eligiendo 100 controles equiparados con respecto a las mismas 5 variables. Al igual que con el equiparamiento con respecto a solo una variable, incluso si la cantidad de estratos obtenidos del equiparamiento por frecuencias es la misma que la que surja al agrupar estratos equiparados individualmente, los individuos reales pueden ser distintos según los individuos de control que en verdad se seleccionen. En consecuencia, el análisis resultante del equiparamiento por frecuencias y del equiparamiento individual puede dar resultados distintos.

ID	ESTRATO	CASO	EXP	TAB
1	1	1	1	0
2	1	0	1	0
3	2	1	0	0
4	2	0	1	0
5	3	1	1	1
6	3	0	0	1
7	4	1	1	0
8	4	0	0	0
9	5	1	0	1
10	5	0	0	1

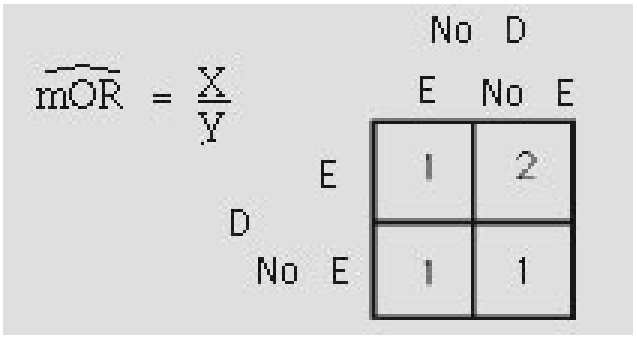
Auto-evaluación(P15.15)

Los datos a la derecha son de un estudio hipotético de casos y controles apareados que incluye 5 pares equiparados, donde la única variable que se equipara es el tabaquismo (TAB). La variable de enfermedad se denomina CASO y la variable de exposición se llama EXP. La variable ID indica que existen 10 combinaciones posibles de estratos en nuestros datos. La variable ESTRATO denota el *número* de estratos formados con la determinada combinación de ID (p. ej., existe 1 estrato con la combinación 1, donde se tabula un caso expuesto sin hábito de tabaquismo; 1 estrato con la combinación 2, donde se tabula un control expuesto sin hábito de tabaquismo; 2 estratos con la combinación 3, donde se tabula un caso no expuesto sin hábito de tabaquismo; etc.)

1. ¿Cuántos pares concordantes tienen ambos miembros expuestos? . . . **???**
2. ¿Cuántos pares concordantes tienen ambos miembros no expuestos? . . . **???**
3. ¿Cuántos pares discordantes hay en los que el caso está expuesto y el control no está expuesto? **???**
4. ¿Cuántos pares discordantes hay en los que el caso no está expuesto y el control está expuesto? **???**

Opciones

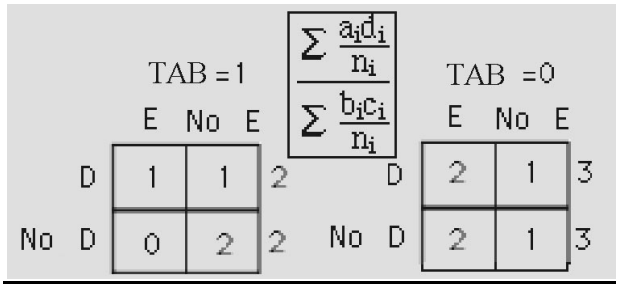
0 **1** **2** **3** **4**



El cuadro resume la información de pares equiparados descrita en las preguntas anteriores.

5. Este cuadro se denomina cuadro **???**
6. ¿Cuál es el mOR calculado para estos datos? . . . . . **???**
7. ¿Qué tipo de análisis equiparado se utiliza con este cuadro? . . . **???**

Opciones
0,5 1 2 de Berkson de Mantel-Haenszel de McNemar agrupado no agrupado



El cuadro agrupa la información de pares equiparados descrita en las preguntas 1-4 en dos estratos de tabaquismo.

8. ¿Cuál es el mOR calculado a partir de estos datos? . . . . . **???**
9. ¿Qué tipo de análisis equiparado se utiliza aquí? . . . . **???**
10. ¿Qué tipo de análisis debería preferirse para estos datos equiparados (donde la condición de tabaquismo es la única variable equiparada)? ¿agrupado o no agrupado? . . **???**

Opciones
1 2 2,5 agrupado indefinido no agrupado

DATOS INTERCAMBIADOS

ID		ESTRATO	CASO	EXP	TAB
1		1	1	1	0
2		1	0	1	0
3		2	1	0	0
4	*	2	0	0	0
5		3	1	1	1
6		3	0	0	1
7		4	1	1	0
8	*	4	0	1	0
9		5	1	0	1
10		5	0	0	1

En los datos arriba se intercambia el control no fumador del estrato 2 con el control no fumador del estrato 4 del conjunto de datos brindado en las preguntas anteriores 1-4.

Consideremos:

W = cantidad de pares concordantes (E = 1, E = 1)
X = cantidad de pares discordantes (E = 1, E = 0)
U = cantidad de pares discordantes (E = 0, E = 1)
W = cantidad de pares concordantes (E = 0, E = 0) para los datos "intercambiados". Entonces:

11. W = . . **???**
12. X = . **???**
13. Y = . . **???**
14. Z = . . **???**
15. mOR (no agrupado) = **???**
16. mOR (agrupado) = **???**

Opciones

0 1 2 2,5 3 4 indefinido

Para los datos apareados que se consideraron en las preguntas 1-10, mOR(no agrupado) = 2 mientras que mOR(agrupado) = 2,5. Para los datos apareados ("intercambiados") que se consideraron en las preguntas 11-16, la mOR(no agrupado) fue indefinida, mientras que mOR(agrupado) = 2,5.



	DATOS PREVIOS					DATOS INTERCAMBIADOS				
ID		ESTRATO	CASO	EXP	TAB	ID	ESTRATO	CASO	EXP	TAB
1		1	1	1	0	1	1	1	1	0
2		1	0	1	0	2	1	0	1	0
3		2	1	0	0	3	2	1	0	0
4	*	2	0	1	0	4	* 2	0	0	0
5		3	1	1	1	5	3	1	1	1
6		3	0	0	1	6	3	0	0	1
7		4	1	1	0	7	4	1	1	0
8	k	4	0	0	0	8	* 4	0	1	0
9		5	1	0	1	9	5	1	0	1
10		5	0	0	1	10	5	0	0	1

17. ¿Cuál de las siguientes opciones ayuda a explicar por qué el mOR agrupado debería preferirse al mOR no agrupado? . . **???**

a. Los mOR agrupados son iguales, mientras que los mOR no agrupados son distintos.
b. Los mOR no agrupados suponen que los pares equiparados intercambiables no son únicos.
c. Los mOR agrupados suponen que los pares equiparados intercambiables son únicos.

Opciones
Todas **Ninguna** **a** **b** **c**

Análisis de datos equiparados por frecuencias

El cuadro que se muestra a continuación muestra el desglose de casos y controles para un **estudio hipotético de casos y controles equiparados** que empleó el **equiparamiento por frecuencias** con respecto a raza y sexo. El estudio tiene el doble de controles que el número casos. Sin embargo, ya que las cantidades de casos y controles difieren en cada uno de los cuatro estratos, la razón de equiparamiento aquí tiene la forma $\mathbf{R_i}$ a $\mathbf{S_i}$, donde $\mathbf{R_i}$ es la cantidad de controles y $\mathbf{S_i}$ es la cantidad de casos en el estrato **i**.

Equiparamiento por frecuencias con respecto a raza y sexo

	Hombre blanco	Mujer blanca	Hombre negro	Mujer negra	
Casos S_i	100 (33%)	100 (33%)	40 (13%)	60 (20%)	300
Controles R_i	200 (33%)	200 (33%)	80 (13%)	120 (20%)	600

Razón de equiparamiento: R_i a S_i

	E	no E	
D	Y_i	$S_i - Y_i$	S_i
no D	X_i	$R_i - X_i$	R_i

Y_i = # de casos en el estrato i expuestos

X_i = # de controles en el estrato i expuestos

Después de este párrafo se encuentran los datos del estudio obtenidos para cada uno de los cuatro estratos. ¿Cómo analizamos estos datos equiparados por frecuencias? La solución es llevar a cabo un **análisis estratificado** de los datos en los 4 estratos para obtener un *odds ratio* basado en la precisión, una prueba de hipótesis de Mantel-Haenszel y un intervalo de confianza de 95% para el **aOR**. Sin embargo, antes de sacar dichos cálculos de

resumen debemos verificar que no haya interacción significativa o importante entre los estratos. Estos son los resultados calculados:

Equiparamiento por frecuencias con respecto a raza y sexo

	Estrato 1 Hombre blanco		Estrato 2 Mujer blanca		Estrato 3 Hombre negro		Estrato 4 Mujer negra		
	E	no E	E	no E	E	no E	E	no E	
Casos S_i	30	70	35	65	11	29	19	41	300
Controles R_i	20	180	22	178	8	72	11	109	600
	$\hat{OR}_1 = 3.86$		$\hat{OR}_2 = 4.36$		$\hat{OR}_2 = 3.41$		$\hat{OR}_2 = 4.59$		

aOR = 4.10
prueba MH = 64.30 (P < .0001)
95 % IC para aOR: (2.86, 5.88)
B-D prueba para interacción: P = .96

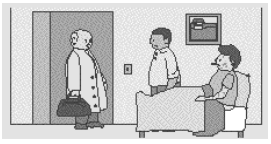
A continuación se incluyen algunas preguntas de estudio sobre la base de estos resultados.

Preguntas de estudio (P15.16)

1. ¿Hay confusión?
2. En función de la comparación de los *odds ratios* específicos para cada estrato, ¿se justifica realizar una evaluación general de la relación entre la exposición y la enfermedad?
3. ¿Por qué se calcula un aOR en lugar de un mOR?
4. Sobre la base de los resultados, ¿a qué conclusión llega sobre la asociación E→D?

Resumen

- Los datos equiparados por frecuencias se pueden examinar mediante un análisis estratificado convencional.
- Para los datos equiparados por frecuencias, los estratos normalmente poseen una razón de equiparamiento variada, con $\mathbf{R_i}$ controles y $\mathbf{S_i}$ casos en el estrato **i**.
- Como los estratos equiparados por frecuencias comúnmente tienen frecuencias de celda de tamaño grande, y las frecuencias de celda cero o vacías son poco usuales, a menudo se usa un aOR basado en la precisión en lugar de un mOR como la estimación puntual general.



Auto-evaluación (P15.17)

Un estudio de casos y controles realizado en 1990 en Puerto Rico tenía como objetivo determinar los factores de riesgo asociados al sarampión grave (D). Los controles y los casos se equipararon por frecuencias según la región de residencia (5 regiones). Los factores potenciales de riesgo que se evaluaron incluyeron: ingreso familiar anual < $5000 (sí o no), enfermedad subyacente (sí o no), madre sin educación de secundaria completa (sí o no) y anemia (sí o no).

1. **Verdadero** o **falso**. Cuando se considera cada factor potencial de riesgo en forma separada como la

variable de exposición, lo apropiado es efectuar un análisis estratificado que incluya 5 estratos. **???**

En el estudio de casos y controles equiparados por frecuencias descrito anteriormente, se compararon 16 casos (es decir, niños con sarampión grave) con controles que consistían en niños seleccionados con sarampión no grave (39 hospitalizados y 38 no hospitalizados); 50% de los casos y 16% de los controles no hospitalizados no padecían de enfermedad subyacente. Al estratificar por región se descubrió que mOR = 5,3 con un IC de 95% para mOR = (1,4; 20,2).

2. ¿Cuál de las siguientes afirmaciones **no** respalda el uso de un mOR para este análisis? **???**
    a. Había celdas con cero o vacías en algunos estratos.
    b. Se recurrió al equiparamiento por frecuencias.
    c. Debía haberse utilizado la regresión logística.

3. **Verdadero** o **falso**. Ya que los controles se equipararon por frecuencias con los casos con respecto a la región de residencia, NO es necesario controlar dicha región en el análisis. . . **???**

4. Si equiparamos por frecuencias y elegimos 3 veces más controles que casos, la razón de equiparamiento es **???**

5. ¿El equiparamiento por frecuencias con respecto a la región de residencia brinda mayor precisión con respecto al *odds ratio* ajustado que si se hubiese empleado el equiparamiento individual? **???**

Opciones

1 a 3 **Siempre** **Falso** **Nunca** **No necesariamente** **R a 1** **R_i a S_i** **Verdadero**

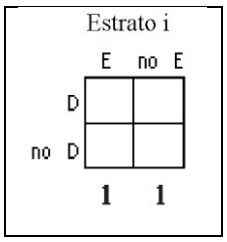
15-5Equiparamiento (continuación)

Análisis de datos de cohorte equiparados

*Aquí describimos cómo analizar datos apareados de un **estudio de cohorte**, donde se forman pares de individuos no expuestos e individuos expuestos con respecto a las variables de equiparamiento seleccionadas.*

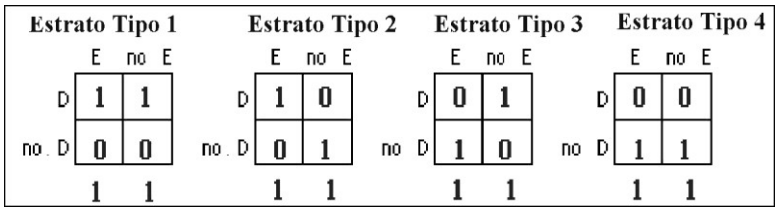
Análisis de datos de cohorte equiparados 1 a 1

Hasta este momento hemos considerado solo datos de casos y controles equiparados. Ahora nos concentramos en el análisis de **datos de cohorte equiparados**. En un estudio de cohorte retrospectivo se empleó la formación de pares para investigar si los turnos rotativos, en comparación con los turnos fijos, tenían relación con el desarrollo de dolor lumbar crónico en un grupo de obreros de Ohio desde 1950 a 1975. El estudio incluía 80 **pares equiparados** de trabajadores expuestos versus no expuestos. A los trabajadors expuestos se les había asignado turnos rotativos, mientras que a los trabajadores no expuestos se les había asignado turnos fijos. El equiparamiento implicó siete variables: 1) año de nacimiento, 2) año de contratación, 3) edad al momento de la contratación, 4) duración del empleo, 5) tipo de trabajo, 6) raza y 7) estado civil.



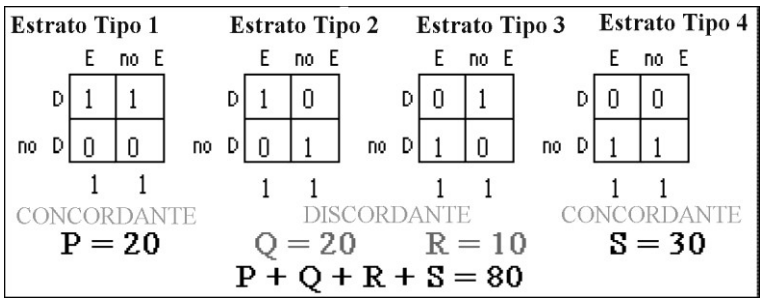
El resultado de salud (dolor lumbar crónico), se determinó a partir de informes obtenidos sobre los trabajadores durante examinaciones médicas bianuales obligatorias realizadas en 1975. Al igual que con los datos de casos y controles equiparados, los 80 pares equiparados en la cohorte representan 80 estratos que contienen 2 personas por estrato. A la derecha se encuentra el cuadro de 2 x 2 para el estrato **i** que relaciona la exposición (asignación a turnos rotativos versus turnos fijos) con el resultado de salud (dolor lumbar crónico). Ya que éste es un estudio de cohorte equiparado con respecto a la exposición, los dos totales de las columnas en el estrato **i** son 1.

Cada uno de los 80 estratos puede adoptar una de las cuatro formas que se muestran a continuación, según el resultado de salud determinado para las personas expuestas y no expuestas en un estrato dado.



El estrato de tipo 1 identifica cualquier par formado donde **tanto** los individuos expuestos **como** los no expuestos tienen el resultado de salud, es decir, ambos poseen dolor lumbar crónico. Esto se denomina **par concordante**. Denotamos la cantidad de pares concordantes de este tipo con **P**. En el estudio se hallaron 20 pares de este tipo. El estrato de tipo 4 identifica aquellos pares donde se determinó que **ni** los individuos expuestos **ni** los no expuestos tienen dolor lumbar crónico. Estos también son **pares concordantes**, ya que tanto los individuos expuestos como los no expuestos presentan el mismo resultado de salud, que esta vez es ausencia de dolor lumbar crónico. Denotamos la cantidad de pares concordantes de este tipo con **S**. En el estudio se encontraron 30 pares de este tipo.

Los otros dos tipos de estratos consideran **pares discordantes**. En el estrato de tipo 2, el individuo expuesto desarrolla el resultado de salud pero el no expuesto no. En el estrato de tipo 3, el individuo expuesto no presenta el resultado de salud, pero el no expuesto sí. Las cantidades de pares discordantes de cada tipo se denominan **Q** y **R**, respectivamente. En el estudio se determinó que **Q** = 20 y **R** = 10. La suma de **P**, **Q**, **R** y **S** = 80, la cantidad de pares en el estudio:



¿Cómo analizamos estos datos? Al igual que con cualquier análisis equiparado, la respuesta general aquí es que se recurre a una computadora para realizar el análisis estratificado de estos 80 estratos a fin de obtener una razón de riesgos de Mantel-Haenszel (**mRR**, una prueba de hipótesis de Mantel-Haenszel y un intervalo de confianza de 95% para la mRR.

Preguntas de estudio (P15.18)

1. ¿Por qué es necesario calcular una mRR en lugar de un mOR o una razón de riesgos ajustada basada en la precisión (es decir, aRR)?

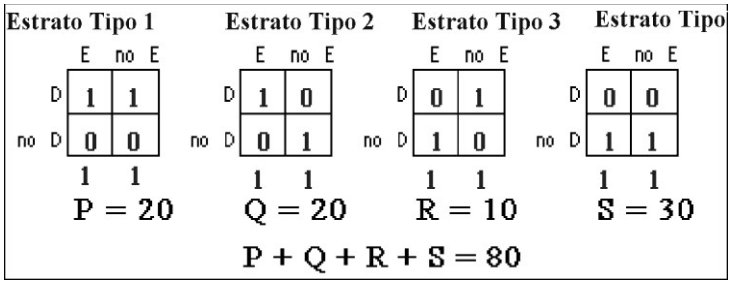
Efectuaremos este análisis en la siguiente actividad.

Resumen

- Ilustramos un análisis equiparado mediante datos de cohorte apareados.
- El estudio incluyó 80 pares o estratos equiparados con dos personas por estrato.
- Había cuatro tipos de estratos: dos incluían **pares concordantes** y dos estaban compuestos por **pares discordantes**.
- **P** = cantidad de pares concordantes: enfermo expuesto y enfermo no expuesto
- **Q** = cantidad de pares discordantes: enfermo expuesto, sano no expuesto
- **R** = cantidad de pares discordantes: enfermo no expuesto, sano expuesto
- **S** = cantidad de pares concordantes: sano expuesto y sano no expuesto
- Los datos se pueden examinar mediante un análisis estratificado para obtener una mRR, una prueba de hipótesis de MH y un intervalo de confianza alrededor de la mRR.

Análisis de datos de cohorte apareados

Continuamos considerando los siguientes datos apareados, o emparejados, de un estudio de cohorte retrospectivo que incluye 80 pares.



Se puede efectuar un análisis estratificado de los 80 estratos con una computadora para obtener una razón de riesgos de Mantel-Haenszel, una prueba de hipótesis de Mantel-Haenszel y un intervalo de confianza de 95% para la mRR. Cada uno de los 80 estratos corresponderá a uno de los cuatro tipos que se presentaron anteriormente, compuesto de pares concordantes o discordantes.

Una forma práctica de realizar este análisis sin usar una computadora es formar el siguiente cuadro con las cantidades de pares concordantes y discordantes **P**, **Q**, **R** y **S**. Este es el **cuadro de McNemar para datos de cohorte apareados o emparejados**. Cada observación en este cuadro representa un par de observaciones, en lugar de una observación individual. Con este cuadro podemos obtener una fórmula para la mRR, el estadístico de prueba de Mantel-Haenszel y un intervalo de confianza de 95% para la mRR.

Cuadro de McNemar

		No expuesto: D	No expuesto: no D
Expuesto	D	**P**	**Q**
	no D	**R**	**S**

$$mRR = \frac{P + Q}{P + R}$$

$$\chi^2_{MHS} = \frac{(Q - R)^2}{Q + R}$$

$$mRR \exp\left[\pm 1.96\sqrt{\frac{Q + R}{(P + Q)(P + R)}}\right]$$

Note que la fórmula para la mRR y su intervalo de confianza involucra la cantidad **P** de pares concordantes con el resultado de salud, además de las cantidades **Q** y **R** de pares discordantes, aunque no se utiliza la información del par concordante de no casos. Al sustituir los valores de **P**, **Q** y **R** en cada una de estas fórmulas obtenemos los resultados que se muestran a continuación:

$$\hat{mRR} = \frac{20 + 20}{20 + 10} = \mathbf{1.33}$$

$$\chi^2_{MHS} = \frac{(20 - 10)^2}{20 + 10} = \mathbf{3.33}$$

95 % IC para mRR: **(0.98, 1.81)**

Preguntas de estudio (P15.19)

1. ¿Cómo interpreta los resultados anteriores en términos de la relación entre los turnos rotativos vs. los fijos (E) y el dolor lumbar crónico (D) que se abordó en el estudio de cohorte apareado?

Resumen

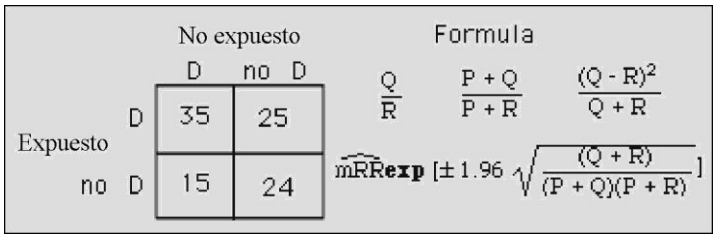
- Una forma práctica de realizar un análisis de datos de cohorte apareados es aplicar el **cuadro de McNemar** que contiene pares concordantes y discordantes.
- Pueden utilizarse fórmulas simples para calcular la mRR, la prueba de hipótesis de MH y el IC del 95% alrededor de la mRR.

Auto-evaluación (P15.20)

Verdadero o **falso**. En un estudio de cohorte apareado:

1. El total de casos en cada conjunto equiparado siempre es igual a 1. . . . **???**
2. La prueba de hipótesis de McNemar incluye solo la información de los pares discordantes. **???**

En base al cuadro de McNemar siguiente:



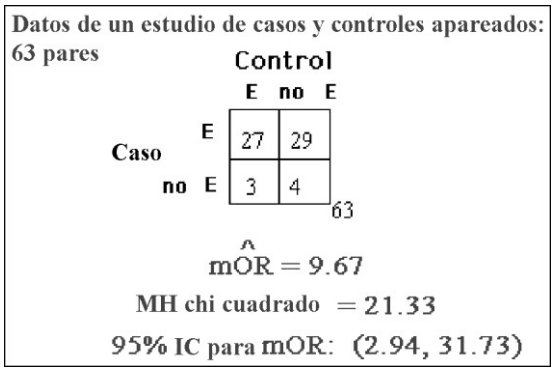
3. La mRR calculada está dada por **???**
4. El estadístico de la prueba de McNemar es **???**
5. El IC de 95% para la mRR posee los siguientes límites **???**

Opciones
(0,96; 1,50) **(1,1; 1,6)** **0,25** **1,20** **1,67** **2,5**

Regresión logística – Covariables equiparadas y no equiparadas para estudios de casos y controles equiparados

¿Por qué el análisis estratificado es inadecuado en los estudios de casos y controles equiparados?

Hemos descrito anteriormente cómo analizar datos de un estudio de **casos y controles apareados** mostrando los resultados para 63 pares mediante el cuadro de McNemar para evaluar la relación entre el uso de estrógenos y el cáncer de endometrio. A continuación se muestran un *odds ratio* de Mantel-Haenszel (mOR), un estadístico de prueba de Mantel-Haenszel y el intervalo de confianza de 95% para el mOR:



El equiparamiento en este estudio se realizó con respecto a tres variables, la edad, el estado civil y la fecha de ingreso a la comunidad para jubilados, y estas son las únicas variables que se controlaron en el análisis.

Sin embargo, había otras variables no equiparadas que también se consideraron como posibles variables de control. Una de ellas fue la condición de enfermedad de la vesícula biliar, que los investigadores tomaron como un factor de riesgo para el cáncer de endometrio que también se debía controlar. ¿Cómo analizamos los datos para dar cuenta no solo de las variables involucradas en el equiparamiento, sino también de la variable que describe la condición de enfermedad de la vesícula biliar? Lamentablemente, el análisis estratificado no es el mejor enfoque para aplicar aquí. Para explicar esto, a continuación se muestran los datos de dos pares de individuos en el conjunto de datos:

PAR	D	E	BILIAR
A	1	0	1
A	0	1	0
B	1	1	0
B	0	1	1

En el par A, el caso padece la enfermedad de la vesícula biliar, pero el control, no. En el par B, el control sufre dicha enfermedad, pero el caso, no. Si queremos estratificar con respecto a la condición de enfermedad de la vesícula biliar y, al mismo tiempo, conservar el equiparamiento, lamentablemente tendremos que dejar de lado estos dos pares del conjunto de datos porque ambos miembros de cada par están en categorías distintas de la variable de condición de enfermedad de la vesícula biliar. Hay una manera de realizar el análisis sin tener que descartar a ningún par formado. El método que se debe emplear aquí es la **regresión logística**, que ejemplificamos en la siguiente actividad.

Resumen

- El análisis estratificado no es el mejor enfoque para examinar los datos equiparados cuando se controlan variables no equiparadas, además de las variables involucradas en el equiparamiento.
- Cuando se estratifica con respecto a variables no equiparadas y equiparadas en un estudio con formación de pares, en el análisis debemos dejar de lado aquellos estratos equiparados en los que los dos individuos del par tienen categorías diferentes de la variable no equiparada.
- La regresión logística puede usarse para controlar tanto las variables equiparadas como las no equiparadas, sin dejar de lado ningún estrato equiparado.

Regresión logística – Covariables equiparadas y no equiparadas en estudios de casos y controles equiparados

Esta actividad describe cómo valerse de la **regresión logística** para controlar tanto las variables equiparadas como las no equiparadas. Nuevamente consideramos los datos de casos y controles apareados que incluyen a 63 pares para evaluar la relación entre el uso de estrógenos y el cáncer de endometrio. Las variables de equiparamiento fueron la edad, el estado civil y la fecha de ingreso a la comunidad para jubilados. La única variable de control no equiparada correspondió a los antecedentes de enfermedad de la vesícula biliar. A continuación se muestra la forma logit de un modelo logístico para estos datos:

$$\text{logit } P = \ln \frac{P}{1-P} = b_0 + \sum_{i=1}^{62} b_i D_i + b_{63} E + b_{64} \text{ BILIAR}$$

D_i = variables ficticias (i parejas equiparadas)
BILIAR = enfermedad de la vesícula biliar (0,1)
E = uso de estrógenos (0,1)

Este modelo contiene 62 variables ficticias para distinguir los 63 pares formados. El modelo también contiene la variable de control no equiparada BILIAR y la variable de exposición E que indica la condición del uso de estrógenos. Este modelo contiene 65 parámetros, de b_0 a b_{64}, que es más de la mitad de la cantidad total de individuos. Cuando, como en este ejemplo, la cantidad de parámetros es grande en relación con la cantidad total de individuos, el modelo logístico se adapta mediante la **estimación de máxima verosimilitud condicional** o método **EMVC**. Para realizar la EMVC, utilice un programa informático.

Cuando se aplica dicho programa para adaptar este modelo, el *odds ratio* calculado que ajusta tanto las variables equiparadas como la variable no equiparada BILIAR está dada por la expresión $e^{\hat{b}_{63}}$, donde $\hat{b}$ es el coeficiente calculado de la variable de exposición. El cálculo resultante es $\hat{OR}(\text{ajustada}) = \exp[2{,}209] = 9{,}11$

Preguntas de estudio (P15.2)

1. ¿Por qué el *odds ratio* obtenido del modelo logístico anterior difiere del valor del mOR de 9,67 que se había calculado anteriormente para estos datos?
2. ¿Cómo expresaría la forma logit de un modelo logístico para el efecto del uso de estrógenos que controle el equiparamiento, pero que no considere la variable BILIAR?
3. El valor de máxima verosimilitud condicional de $\hat{b}_{63}$ para el modelo en la pregunta 2 es 2,269. ¿Cuál es el valor del *odds ratio* ajustado que controla el equiparamiento?

Con la regresión logística, también podemos obtener un estadístico chi cuadrado para probar la hipótesis nula de que el *odds ratio* ajustado es 1, es decir, que no hay un efecto significativo del uso de estrógenos. Una hipótesis nula equivalente es que el coeficiente de la variable de exposición es cero. El estadístico chi cuadrado resultante, que se denomina **estadístico de la razón de verosimilitud** tiene el valor 13,11. Este estadístico es aproximadamente chi cuadrado con 1 gl según la hipótesis nula.

Razón de verosimilitud

$$\chi^2_{1\ df} = 13.11$$

$$H_0: OR_{(ajustado)} = 1 \iff b_{63} = 0$$

Preguntas de estudio (P15.21) (continuación)

4. ¿Cuál es el valor p de la prueba de chi cuadrado? ¿La prueba es significativa?
5. ¿Por qué el estadístico chi cuadrado es distinto del chi cuadrado de McNemar de 21,13 calculado anteriormente para estos datos?

También puede obtenerse un intervalo de confianza de 95% para el *odds ratio* ajustado a partir del resultado del modelo logístico. A continuación se presentan la fórmula para este intervalo de confianza y los

límites de confianza resultantes que se obtuvieron de los valores que arrojó la computadora:

95% IC para OR(ajustado):

$$\exp\left[\hat{b}_{63} \pm 1.96 \text{ s.e.}(\hat{b}_{63})\right] = (2.76, 30.10)$$

Preguntas de estudio (P15.21) (continuación)

6. ¿Cómo interpreta este intervalo de confianza?

Resumen

- La regresión logística puede utilizarse para controlar tanto las variables equiparadas como las no equiparadas.
- Aquí el modelo logístico contiene **G – 1** variables ficticias para **G** estratos equiparados, la variable de exposición (**E**) y cualquier variable no equiparada que también se controlará.
- Para los datos equiparados, el modelo logístico se calcula mediante la **estimación de máxima verosimilitud condicional** (**EMVC**).
- El modelo se calcula mediante la regresión logística condicional cuando hay datos equiparados.
- Si **E** es una variable (0, 1) y ninguna variable predictora es un término de producto, entonces el OR ajustado para el efecto de **E** está dado por **exp(b)**, donde **b** es el coeficiente de la **E**.
- Una prueba de la hipótesis nula de que el OR ajustado = 1 puede obtenerse mediante una **prueba de la razón de verosimilitud** basada en el resultado de la regresión logística.
- Un IC del 95% para el efecto de la exposición está dado por la fórmula siguiente:
  $\exp[b \pm 1{,}96 \text{ s.e.}(b)]$

Métodos de máxima verosimilitud para analizar datos equiparados mediante la regresión logística

El enfoque más popular para calcular los parámetros en un modelo de regresión logística es la estimación de **máxima verosimilitud (MV)**. En verdad, existen dos enfoques de MV alternativos: el **método no condicional (EMVNC) y el método condicional (EMVC)**. Ambos requieren diferentes programas informáticos.

De estos dos métodos, el enfoque **EMVC** es el que normalmente se emplea para analizar datos equiparados. Se puede recurrir a un programa informático de análisis de supervivencia para aplicar el método **EMVC**, como se brinda en los paquetes de programas informáticos SAS y SPSS.

La distinción clave entre los enfoques de MV no condicional y condicional se refiere a la **cantidad de parámetros (p)** en el modelo que se utiliza en relación con la **cantidad total de individuos** (**n**) en el estudio. Si **p** es pequeña en relación con **n**, entonces normalmente se utiliza la **EMVNC**. Si **p** es grande en relación con **n**, entonces se prefiere la **EMVC**.

Sin embargo, la estadística nunca ha determinado con precisión *qué se entiende por grande* exactamente. No obstante, por lo general se sostiene que el modelo logístico para datos equiparados a menudo se encuentra en la categoría "grande".

Por ejemplo, en un estudio de casos y controles que incluyó 100 pares formados con equiparamiento con respecto a la EDAD, la RAZA y el SEXO, un modelo logístico (suponiendo que no hay agrupamiento y que no hay términos de interacción) tendría la siguiente forma:

$$\text{logit } P(X) = b_0 + \sum_{i=1}^{99} b_i D_i + b_{100} E$$

donde las **D** denotan 99 variables ficticias para distinguir los 100 estratos equiparados y **E** denota una variable de exposición dicotómica.

La cantidad de parámetros en este modelo es p = 101, mientras que la cantidad total de individuos es n = 100 x 2 = 200. Aquí, p comúnmente se consideraría grande en relación con n, por lo que se recomendaría el enfoque de EMVC. De hecho, para los datos apareados, puede demostrarse matemáticamente que el *odds ratio* obtenido al utilizar la EMVC es el cuadrado del *odds ratio* que se habría obtenido si, en su lugar, se hubiese utilizado la EMVNC en ese mismo modelo, es decir:

$$\hat{OR}_{EMVNC} = (\hat{OR}_{EMVC})^2$$

Se ha demostrado que el *odds ratio* para el método de **EMVC** es el *odds ratio* correcto (es decir, **no sesgado**), lo que significa que (el cuadrado) de este *odds ratio* obtenido a partir del método **EMVNC** está **sesgado**. En especial, si el *odds ratio* correcto (es decir, mediante **EMVC**) resulta ser 3, entonces el *odds ratio* sesgado (es decir, mediante **EMVNC**) sería 9, lo que representa un sesgo muy grande.

Por el contrario, si no se hubiese utilizado el equiparamiento en un estudio de casos y controles con 200 individuos, un modelo logístico para evaluar el efecto de la exposición E controlando, por ejemplo, la EDAD, la RAZA y el SEXO adoptaría la siguiente forma:

$$\text{logit } P(X) = b_0 + b_1 EDAD + b_2 RAZA + b_{3SEXO} + b_4 E$$

La cantidad de parámetros en este modelo es **p** = 5, mientras que la cantidad total de individuos es **n** = 200. Aquí, a diferencia del modelo con equiparamiento, **p** es claramente pequeña en relación con **n**, por lo que se recomendaría el enfoque de **EMVNC**.

En general, el método **EMVC** siempre da la respuesta (por ejemplo, el valor del *odds ratio*) correcta es decir, no sesgada). Sin embargo, cuando **p** es pequeña en relación con **n**, se prefiere el método **EMVNC** porque no solo dará básicamente la misma respuesta que el método **EMVC**, sino que el método **EMVNC** normalmente brindará un intervalo de confianza más preciso alrededor del valor puntual que el obtenido mediante el método **EMVC**.

El modelo logístico EVW para datos equiparados

Una expresión general para el modelo logístico que permite el análisis de datos equiparados puede expresarse de la siguiente manera:

$$\text{logit } P(X) = \alpha + \sum_{i=1}^{p_1} \gamma_{1i} V_{1i} + \sum_{i=1}^{p_2} \gamma_{2i} V_{2i} + \beta E + \sum_{j=1}^{p_3} \delta_j E W_j$$

Las **α**, **γ**, **β** y **δ** de este modelo (previamente denotadas con **b** en la actividad sobre modelación logística) representan coeficientes de regresión desconocidos que se deben calcular (normalmente) mediante la estimación de máxima verosimilitud condicional (es decir, **EMVC**). Las **E**, **V** y **W** representan las variables predictoras en este modelo. Asimismo, las **EW** representan términos de producto de **E** con una de las **W**. Más específicamente,

la variable predictora **E** denota una única **variable de exposición** de interés, el primer grupo de las **V** denota **variables ficticias** para identificar $p_1 + 1$ estratos equiparados, el segundo grupo de las **V** denota las variables de control (es decir, **factores potenciales de confusión**) que no están involucrados en el equiparamiento y las **W** denotan **modificadores de efecto** potenciales que entran en el modelo como términos de producto con la variable de exposición **E**.

Damos un ejemplo de este modelo para un estudio sobre la relación entre el uso de estrógenos (**E = EST**, 1 indica "sí", 0 denota "no") y el cáncer de endometrio (**D = CEN**) que incluye $p_1 + 1 = 63$ pares de mujeres en una comunidad para jubilados de Los Ángeles. Las variables con respecto a las cuales se equiparó fueron la edad, el estado civil y la fecha de ingreso a la comunidad.

Otra variable, no equiparada, es la presencia o ausencia de enfermedad de la vesícula biliar (**BILIAR** = 1 si está presente, 0 si está ausente). El siguiente es un modelo logístico sin interacción para estos datos basado en la fórmula general anterior de **EVW**:

$$\text{logit } P(X) = \alpha + \sum_{i=1}^{62} \gamma_{1i} V_{1i} + \gamma_{21} \text{BILIAR} + \beta \text{EST}$$

Las $\mathbf{V_{1i}}$ en este modelo denotan $\mathbf{p_1} = 62$ variables ficticias para los 63 pares formados. Ya que solo hay una variable no equiparada en el modelo, es decir, **BILIAR**, solo hay un término $\mathbf{V_{2i}}$ en la segunda suma de la fórmula general de **EVW**, es decir, $\mathbf{p_2} = 1$.

Cuando se usa la **EMVC** para calcular los parámetros en el modelo, los únicos parámetros que, en verdad, se calculan son los coeficientes de EST y BILIAR, ya que **α** y las **γ** se dejan de lado en la función de verosimilitud condicional que se maximiza.

La fórmula del *odds ratio* ajustado del efecto del uso de estrógenos, ajustado por las variables de equiparamiento y por BILIAR es, entonces, igual a:

$$\hat{OR}(\text{adj}) = \exp(b) \qquad \text{donde} \quad b = \hat{\beta}$$

Un modelo logístico alternativo para estos datos que permite la posibilidad de interacción entre el uso de estrógenos y la condición de enfermedad de la vesícula biliar está dado por:

$$\text{logit } P(X) = \alpha + \sum_{i=1}^{62} \gamma_{1i} V_{1i} + \gamma_{21} \text{BILIAR} + \beta \text{EST} + \delta_1 \text{EST} \times \text{BILIAR}$$

Para este modelo con interacción, la fórmula del *odds ratio* ajustado está dada por:

$$\hat{OR}(\text{ajustada}) = \exp(b + \hat{\delta}_1 \text{BILIAR})$$

de forma que el valor del *odds ratio* "con interacción" cambiará (es decir, se modificará) en función de los valores del modificador potencial de efecto BILIAR.

Auto-evaluación (P15.22)

Complete los espacios vacíos.

Se realizó un estudio de casos y controles equiparados sobre el cáncer cervical en Sydney, Australia (Brock et. al., J. Nat. Cancer Inst., 1988) que incluyó a 313 mujeres. La variable de resultado fue la condición de cáncer cervical. Las variables con respecto a las cuales se equiparó fueron la edad y la condición socioeconómica (CSE).

Otras variables adicionales no involucradas en el equiparamiento fueron la condición de tabaquismo (TAB) (0, 1), la cantidad de parejas sexuales a lo largo de la vida (QS) y la edad de la primera relación sexual (ES).

1. ¿Cuál de los modelos logísticos que se muestran a continuación sería apropiado para el análisis de estos datos? . . . . . **???**
2. ¿Qué método de cálculo debería emplearse para adaptar el modelo adecuado para estos datos? . . . . . **???**

Opciones
EMV condicional **Mínimos cuadrados** **EMV no condicional** **a** **b** **c**

a. $\text{logit}\, P = b_0 + b_1 TAB + b_2 QS + b_3 ES$

b. $\text{logit}\, P = b_0 + b_1 TAB + b_2 QS + b_3 ES + \sum_{j=1}^{k-1} d_j D_j$

donde las D_j son variables ficticias para k estratos

c. $\text{logit}\, P = b_0 + b_1 TAB + b_2 QS + b_3 ES + b_3 TAB + b_3 CSE$

Los resultados que se muestran a continuación se obtuvieron al adaptar un modelo de regresión logística a los datos sobre cáncer cervical. Suponiendo que se utilizó el modelo correcto, conteste las siguientes preguntas:

Variable	b	E.E.	Valor-P
TAB	1.4361	0.3167	.0000
QS	0.9598	0.3057	.0017
ES	- 0.6064	0.3341	.0695

3. ¿Cuál es el *odds ratio* ajustado para el efecto de TAB en el resultado de cáncer cervical? **???**
4. ¿La cantidad de parejas sexuales a lo largo de la vida es un predictor significativo del resultado de cáncer cervical si controlamos las variables que se equiparan y las otras que no se equiparan? **???**
5. ¿Se están controlando las variables involucradas en el equiparamiento? . . . **???**
6. Un IC de 95% para el efecto (ajustado) del TAB está dado por la expresión **???**

Opciones
1,4361 **1,4361±1,96(0,3167)** **1,4361/0,3167** **exp(1,4361)** **exp[1,4361±0,3167)**
exp[1,4361±1,96(0,3167)] **no** **sí**

Referencias

Referencias sobre equiparamiento

Breslow NE, Day NE. *Statistical Methods in Cancer Research*. *Volume 1. The Analysis of Case-Control Studies*. International Agency for Research in Cancer, Lyon, 1980.

Brock KE, Berry G, Mock PA, MacLennan R, Truswell AS, Brinton LA. *Nutrients in diet and plasma and risk of in situ cervical cancer*. J Natl Cancer Inst, 15 de junio de 1988; 80(8):580-5.

Díaz T, Núñez JC, Rullan JW, Markowitz LE, Barker ND, Horan J. *Risk factors associated with severe measles in Puerto Rico*. Pediatr Infect Dis J 1992; 11(10):836-40. (Ejemplo de equiparamiento por frecuencias)

Donovan JW, MacLennan R, Adena M. *Vietnam service and the risk of congenital anomalies. A case-control study*. Med J; agosto de 1984; 140(7):394-7.

Halpin TJ, Holtzhauer FJ, Campbell RJ, Hall LJ, Correa-Villasenor A, Lanese R, Rice J, Hurwitz ES. *Reye's syndrome and medication use*. JAMA 1982; 248(6):687-91.

Kleinbaum DG, Klein M. *Logistic Regression: A Self-Learning Text*, 2º edición, Springer Verlag Publishers, 2002.

Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic Research: Principles and Quantitative Methods*. John Wiley and Sons Publishers, Nueva York, 1982.

Kupper LL, Karon JM, Kleinbaum DG, Morgenstern H, Lewis DK, *Matching in epidemiologic studies: validity and efficiency considerations*. Biometrics 1981; 37(2):271-291.

McNeil D. *Epidemiologic Research Methods*. John Wiley and Sons Publishers, 1996. (Ejemplo de formación de pares en un estudio de cohorte)

Miettinen OS. *Individual matching with multiple controls in the case of all-or-none responses*. Biometrics 1969; 25(2):339-55.

Miettinen OS. *Estimation of relative risk from individually matched series*. Biometrics 1970; 26(1):75-86.

Ury HK. *Efficiency of case-control studies with multiple controls per case: continuous or dichotomous data*. Biometrics 1975; 31(3):643-9.

Diseño de casos cruzados

Maclure M, Mittleman MA. *Should we use a case-crossover design?* Annu Rev Public Health 2000; 21:193-221.

Maclure M. *The case-crossover design: a method for studying transient effects on the risk of acute events*. Am J Epidemiol 1991; 133(2):144-53.

Redelmeier DA, Tibshirani RJ. *Association between cellular-telephone calls and motor vehicle collisions*. N Engl J Med 1997; 336(7):453-8.

Ejercicios

EPI-1. Tipos de equiparamiento

Suponga que utiliza un estudio de casos y controles para investigar si una variable de exposición dicotómica, E, está asociada a una variable de enfermedad dicotómica, D, y decide equiparar con respecto a grupos etarios (3 categorías) y el sexo. Responda a las siguientes preguntas presumiendo que todas las condiciones *a priori* para la confusión están satisfechas por los grupos etarios y el sexo.

a. Si elige los controles mediante el **equiparamiento por frecuencias** con respecto a los grupos etarios y el sexo, ¿esto significa que puede ignorar el control de los grupos etarios y el sexo en su análisis?
b. Si **equipara individualmente** con respecto a los grupos etarios y el sexo, ¿esto implica que puede ignorar el control de los grupos etarios y el sexo en el análisis?
c. Si **equipara individualmente** con respecto a los grupos etarios y el sexo, ¿debería controlar estas dos variables mediante un análisis estratificado usando conjuntos equiparados individualmente como sus estratos? (Suponga que no se controlan otras variables distintas de los factores que se equiparan).

d. ¿El agrupamiento de conjuntos equiparados individualmente es equivalente al equiparamiento por frecuencias?
e. Si **equipara por frecuencias** con respecto a los grupos etarios y el sexo, ¿puede considerar agrupar con respecto a una variable (por ejemplo, los grupos etarios) sin hacerlo con respecto a la otra (por ejemplo, el sexo)?
f. Si **equipara por frecuencias** con respecto a los grupos etarios y el sexo¿obtendrá una mayor precisión (del *odds ratio* ajustado) que si equipara individualmente?
g. Si **equipara individualmente** con respecto a los grupos etarios y el sexo, ¿logrará una mayor precisión (del *odds ratio* ajustado) al agrupar conjuntos equiparados individualmente (intercambiables) en lugar de emplear conjuntos equiparados individualmente como los estratos?
h. Si **equipara individualmente** con respecto a los grupos etarios y el sexo, ¿obtendrá el mismo *odds ratio* ajustado independientemente de si agrupa o no conjuntos equiparados intercambiables o si utiliza conjuntos equiparados individualmente como los estratos?
i. ¿Esperaría que el equiparamiento por frecuencias arroje aproximadamente el mismo *odds ratio* que el obtenido al agrupar estratos equiparados en forma individual?
j. Cuando se recurre al **equiparamiento individual** con diversas variables de equiparamiento, incluso variables continuas, ¿será ventajoso realizar un análisis estratificado usando conjuntos equiparados agrupados como sus estratos?

EPI-2. Análisis de datos equiparados

Un investigador planea llevar a cabo un estudio de casos y controles para examinar la relación entre una variable de exposición, E, y cierta variable de enfermedad, D, controlando los efectos potenciales de confusión de cierta variable externa, F. El siguiente cuadro describe la población (de origen) total en estudio.

F sí

	E	**no E**
Casos	120	40
Controles	1080	1080
	1200	1020

F no

	E	**no E**
Casos	40	40
Controles	360	1080
	400	1020

a. Sobre la base de los datos anteriores, ¿la variable F parece ser un "factor de riesgo" para la enfermedad?
b. Desarrolle su respuesta. Suponga que se toma una muestra aleatoria de 240 controles de una población de origen y que se estudiarán los 240 casos. Calcule las "frecuencias esperadas" para los dos cuadros que resultarían del muestreo aleatorio.
c. ¿La variable F es un factor de confusión para los datos (esperados) que se tomaron de una muestra aleatoria? Desarrolle su respuesta.
d. ¿Cuál es el método adecuado de análisis para los datos (esperados) que se obtuvieron de una muestra aleatoria: estratificado o no estratificado? Desarrolle su respuesta.
e. Suponga que se usan los 240 casos y que se hallan 240 controles mediante equiparamiento por categorías. Calcule las frecuencias esperadas para los dos cuadros que resultarían de este diseño equiparado.
f. ¿Cuál es el método apropiado de análisis para los datos equiparados del ítem e: estratificado o no estratificado? Desarrolle su respuesta.
g. Realice un análisis estratificado (equiparado) de los datos de frecuencias esperadas. Asegúrese de calcular una estimación puntual de efecto, una prueba de significación y un IC de 95%.

h. Para la población dada, ¿qué enfoque esperaría que diera resultados más precisos (en el cálculo del *odds ratio* ajustado): el equiparamiento o el muestreo aleatorio? Desarrolle su respuesta.
i. En general, ¿cómo decidiría si equiparar o no con respecto a variables externas cuando efectúa un estudio de casos y controles?

EPI-3. Equiparamiento: estudios de seguimiento versus de casos y controles

Este ejemplo ficticio ilustra y contrasta un estudio de seguimiento equiparado con un estudio de casos y controles equiparado con respecto al control de la confusión mediante un factor equiparado (la edad). Considere una cohorte subyacente de 2.000.000 de personas en la que 50% son jóvenes y 50% son individuos de edad avanzada. Suponga que 50% de los jóvenes y 10% de las personas mayores están expuestos. Además, suponga que la edad es un factor de riesgo para la enfermedad. El riesgo de enfermedad para cada grupo etario expuesto, durante un período dado de tiempo, se resume a continuación:

Grupo	Cantidad	Riesgo	Casos
Jóvenes expuestos	500.000	0,005	2500
Jóvenes no expuestos	500.000	0,001	500
Personas mayores expuestas	100.000	0,030	3000
Personas mayores no expuestas	900.000	0,006	5400

a. ¿Cuál es la razón de riesgos para los expuestos vs. los no expuestos entre los jóvenes y entre las personas mayores? ¿La edad es un modificador de efecto?
b. Elabore el cuadro crudo de 2 x 2 para la exposición y la enfermedad. ¿Cuál es la razón de riesgos cruda para la exposición en esta cohorte? ¿La edad es un factor de confusión en esta cohorte?
c. Suponga que llevó a cabo un estudio de seguimiento equiparado a partir de esta cohorte de la siguiente manera: se incluyó 1% de los individuos expuestos (6000 individuos). También participaron 6000 individuos no expuestos que fueron equiparados con los individuos expuestos con respecto al grupo etario. Elabore el cuadro de 2 x 2 (no estratificado) que resultaría si este estudio se hubiera efectuado correctamente y no tuviese error aleatorio (es decir, brinde las cantidades esperadas). Calcule la razón de riesgos. ¿Hay confusión en la razón de riesgos?
d. Suponga que realizó un estudio de casos y controles equiparados con esta cohorte de la siguiente manera: participó 1% de los casos (esto debería sumar 114 casos). También se incluyeron 5 controles para cada caso, lo que implica 570 controles seleccionados de manera tal que tuvieran la misma distribución de edad que los casos (es decir, los controles se equipararon con respecto a la edad). Elabore el cuadro de 2 x 2 resultante (no estratificado) suponiendo que no hay error aleatorio. Elija los controles a partir de la cohorte completa (no solo de la población sana; los números se calculan más fácilmente). Calcule el *odds ratio* de exposición. ¿El *odds ratio* está sesgado?
e. Ajuste el sesgo del ítem d) (si fuese necesario) estratificando por edad en el análisis del estudio de casos y controles (o sea, calcule *odds ratios* específicos para cada estrato).
f. En el ítem d), los controles se eligieron de toda la cohorte, en lugar de seleccionarse sólo de la proporción de individuos sanos de la misma. ¿Qué medida de asociación de la enfermedad calcula el *odds ratio* de exposición ajustado cuando se escogen los controles de esta forma? ¿Por qué en este estudio hay muy poca diferencia si los controles se seleccionaron de toda la cohorte o sólo de la proporción de individuos sanos de la misma?
g. Suponga que realizó un estudio de control de casos y controles en el que los controles no se equipararon con los casos. Como en el ítem d), se incluyeron 114 casos y los controles se eligieron de toda la cohorte. A diferencia del ítem d), los controles no están equiparados con respecto a la edad y se escogieron 600 controles. Elabore el cuadro de 2 x 2 resultante suponiendo que no hay error aleatorio. Calcule el *odds ratio* de exposición. ¿El *odds ratio* está sesgado?

h. Ajuste el sesgo del ítem g) (si fuese necesario) estratificando por edad en el estudio de casos y controles (es decir, calcule *odds ratios* específicos para cada estrato).
i. ¿Hay alguna ventaja de realizar el estudio de casos y controles equiparados (como en los ítems d y e), en contraposición al estudio de casos y controles no equiparados (como en los ítems g y h)? Responda esta pregunta simplemente en términos de estos estudios en particular e ignore el hecho de que, para cada estudio, se seleccionó una cantidad de controles apenas distinta (570 vs. 600). Se tomaron estas cantidades para hacer más simples los cálculos.
j. Considere un estudio de casos y controles en el que se desea controlar el barrio de residencia y cada caso reside en un barrio distinto. ¿Hay alguna ventaja al efectuar un estudio de casos y controles equiparados (equiparando con respecto al barrio) en lugar de un estudio de casos y controles no equiparados? Desarrolle su respuesta.
k. Compare el estudio de seguimiento equiparado del ítem c) con el estudio de casos y controles equiparados del ítem d) con respecto al control de la confusión por la edad (el factor equiparado).
l. A continuación encontrará un extracto del libro *Modern Epidemiology* [Epidemiología moderna] (primera edición) de Kenneth Rothman. La cita corresponde al primer párrafo de sus principios de la sección sobre equiparamiento tomado de la página 237. A pesar del estilo literario, hay un punto importante que se plantea con respecto a los principios de equiparamiento que se intentó ilustrar en toda esta pregunta. Explique lo que plantea Rothman.

Rothman escribió:
El tema del equiparamiento en epidemiología es cautivante. Lo que al principio parece claro es seductoramente engañoso. Mientras que la claridad de un análisis en el que se ha evitado firmemente la confusión mediante el perfecto equiparamiento de las series comparadas pareciera ser, sin dudas, imposible de malinterpretar, el fundamento intuitivo de esta contundencia lograda por el equiparamiento es una estructura sorprendentemente inestable que no siempre respalda las conclusiones que se pueden sacar. La dificultad yace en que nuestra intuición sobre el equiparamiento surge del conocimiento de experimentos o estudios de seguimiento, mientras que el equiparamiento se aplica mayormente en estudios de casos y controles, que difieren lo suficiente de los estudios de seguimiento como para hacer que las implicaciones del equiparamiento sean distintas y contrarias a lo que la intuición nos llevaría a creer.

Respuestas a las preguntas de estudio y auto-evaluaciones

P15.1

1. El síndrome de Reye se asoció solo a ciertos tipos de virus, por lo cual el "tipo de virus" era un factor de riesgo importante para el resultado.
2. Los niños más grandes tenían menos probabilidad de desarrollar el síndrome de Reye, es decir, la edad era un factor de riesgo importante.

P15.2

1. Hay que categorizar las variables continuas edad, presión arterial y tamaño corporal.
2. Se eligen uno o más controles que estén en la misma categoría de edad, sexo, condición de tabaquismo, presión arterial y tamaño corporal que el caso.
3. No. El caso y el control deben estar en la misma categoría de cada una de las variables que se equiparan. La elección del control en la pregunta no es adecuada, porque tanto la categoría de tamaño corporal como la de presión arterial del control difieren de las observadas en el caso.
4. {PA alta, edad $\geq$ 55} = 120; {PA alta, edad < 55} = 180; {PA normal, edad $\geq$ 55} = 150; {PA normal, edad < 55} = 150.
5. {PA alta, edad $\geq$ 55} = 20%; {PA alta, edad < 55} = 30%; {PA normal, edad $\geq$ 55} = 25%; {PA normal, edad < 55} = 25%.
6. Equiparamiento por frecuencias, porque debería ser logísticamente más fácil y menos costoso encontrar grupos de controles especialmente por hospital, que encontrar controles para un caso por vez.
7. Equiparamiento individual, porque hay diversas variables con respecto a las cuales equiparar, muchas de las cuales están bastante individualizadas.

P15.3

1. 30 casos y 60 controles
2. 70 casos y 140 controles
3. R_i a S_i, ya que $R_1 = 60$, $S_1 = 30$ y $R_2 = 140$, $S_2 = 70$
4. Cuatro
5. Razón de equiparamiento R_i a S_i. Aunque, en general, hay el doble de controles que de casos, la cantidad de casos y controles varía dentro de cada estrato.
6. Hombre blanco: R = 200, S = 100; mujer blanca: R = 200, S = 100; hombre negro: R = 80, S = 40; mujer negra: R = 120, S = 60.

P15.4

1. Para una cantidad fija de individuos índice, el tamaño muestral total para el estudio aumenta a medida que R aumenta.
2. Hay un incremento de 33% al pasar de R = 1 a R = 2. Esto indica que se ganará una precisión considerable si se recurre al equiparamiento 2 a 1, en lugar del equiparamiento 1 a 1.

P15.5

1. Falso. En un estudio de casos y controles, el grupo índice está compuesto por casos.
2. Falso. En un estudio de cohorte, el grupo de comparación está formado por individuos no expuestos.
3. Falso. El control de un caso dado en esta situación tendría que ser igual con respecto a los dos factores de equiparamiento.
4. Verdadero
5. Verdadero. El equiparamiento 5 a 1 daría como resultado un aumento de la precisión de 4,3%, comparado con el equiparamiento 4 a 1.
6. Verdadero
7. Verdadero
8. Falso. Elegir 2 controles por caso versus 1 control por caso aumentará la precisión equivalente a 33%. Este incremento de la precisión continuará con cada control que se agregue por caso. Sin embargo, el cuadro de la eficiencia de Pittman, valores calculados para valores distintos de R (cantidad de controles por caso), indica un menor retorno con respecto a la eficiencia una vez que R supera 4.

P15.6

1. No. El color de cabello no tiene una relación conocida con el cáncer de vejiga.
2. Como el color de cabello no tiene una relación conocida con el cáncer de vejiga y no es un factor de riesgo que se debe controlar, es probable que el equiparamiento con respecto al color de cabello no tenga ningún efecto en la precisión del efecto calculado de la exposición en la enfermedad, es decir, aunque el equiparamiento con respecto al color de cabello hará que los casos y los controles estén "equilibrados" en términos del color de cabello, no es probable que el efecto calculado sea más preciso de lo que sería a partir de los datos "desequilibrados" que se obtendrían si no se equiparara.
3. Los productos de comidas rápidas tienden a tener muchas grasas saturadas. Entonces, si equipara con respecto a la cantidad de productos de comida rápida en su dieta, efectivamente estará equiparando con respecto al nivel de colesterol.
4. Tendrá dificultad para encontrar combinaciones para algunos individuos índice y, en consecuencia, es probable que obtenga un tamaño muestral mucho más pequeño que lo planeado originalmente.

P15.7

1. No hay una razón firme a favor del equiparamiento y los motivos en contra superan las causas a favor.
2. Podría esperar ganar precisión del equiparamiento, pero también debe sopesar las otras razones enumeradas, especialmente el costo en tiempo y dinero que implica realizar el equiparamiento.
3. Nuevamente, su decisión depende de cómo sopese todas las razones a favor y en contra del equiparamiento. El equiparamiento con respecto al barrio de residencia puede ser una forma práctica de controlar los factores sociales y ambientales que son difíciles de medir. Sin embargo, si la variable de exposición tiene un componente conductual, debe tener cuidado de no sobrequiparar en esta situación.

P15.8

1. Falso. Una razón para optar por equiparar en un estudio de casos y controles es obtener un cálculo más preciso del *odds ratio* de interés.
2. Falso. La reducción del tamaño muestral a raíz de la imposibilidad de encontrar una combinación adecuada sería una desventaja del equiparamiento.
3. Verdadero
4. a, d y e
5. Falso. Aún debe preocuparse por el equiparamiento con respecto a factores de riesgo débiles y el sobre-equiparamiento.
6. Depende. Quedará supeditado al costo que implique ganar precisión. Es importante considerar otros factores como el costo, el dinero, el tiempo, qué tan fuertes o débiles son los factores de riesgo, el sobrequiparamiento, etc.

P15.9

1. Logit $P(X) = b_0 + b_1(D1) + b_2(D2) + \ldots + b_{99}(D99) + b_{100}(E)$ donde de D1 a D99 son variables ficticias que distinguen los 100 pares formados. En especial, los D_i pueden definirse de la siguiente manera: $D_i = 1$ para un individuo en el par i-avo y $D_i = 0$ para un individuo que no se encuentre en el i-avo par. Por lo tanto, para cada uno de los dos individuos en el primer par, $D1 = 1$, $D2 = D3 = \ldots = D99 = 0$ y para cada individuo en el 100-avo par, $D1 = D2 = \ldots D99 = 0$.
2. Logit $P(X) = b_0 + b_1(D1) + b_2(D2) + \ldots + b_{99}(D99) + b_{100}(PAS) + b_{101}(COL) + b_{102}(E)$ donde de D1 a D99 son 99 variables ficticias que distinguen los 100 pares formados.
3. En dicho análisis estratificado, se deberá dejar de lado cualquier par en el que el caso esté en una categoría distinta de PAS o COL que su correspondiente control. Los únicos pares que se mantendrán en el análisis serán aquellos en que, tanto el caso como el control, se encuentren en las mismas categorías de PAS y COL.

P15.10

1. Todos los estratos tienen celdas con cero, de

forma que no es posible calcular *odds ratios* específicos para cada estrato.

P15.11

1. El valor de 9,67 del mOR indica una relación muy fuerte entre el uso de estrógenos y el cáncer de endometrio, controlando las variables que se equiparan. La prueba de MH tiene un valor p igual a cero con cuatro decimales. Por lo tanto, la estimación puntual es muy significativa. El IC de 95% es bastante amplio, por ende, hay una imprecisión considerable en la estimación puntual. No obstante, los resultados sugieren, en general, un fuerte efecto del uso de estrógenos en el desarrollo de cáncer de endometrio.

P15.12

1. Todos los estratos tienen, al menos, una celda con cero, ya que el caso en cada estrato está expuesto o no expuesto.
2. Aunque solo hay 5 individuos por estrato, el tamaño total del estudio es igual a 90. La propiedad de muestra grande de la prueba de MH depende del tamaño muestral total, no del tamaño muestral específico para cada estrato.
3. El aborto anterior tiene un efecto muy fuerte y significativo (mOR = 33, P < 0,0001) en el desarrollo de embarazo tubárico para las mujeres con, al menos, un embarazo previo. Sin embargo, hay falta de precisión en el cálculo, probablemente debido a que la cantidad de casos (18) y el tamaño muestral total (90) son relativamente pequeños.

P15.13

1. Verdadero
2. Falso. En un estudio de casos y controles apareados, el mOR se calcula como X/Y donde X y Y son frecuencias de pares discordantes.
3. Falso. El valor calculado del estadístico de la prueba de McNemar es $(X - Y)^2 / (X + Y) = 5$.
4. Falso. El estadístico de la prueba de MH es adecuado para el equiparamiento R a 1.
5. 1,03
6. 0,07
7. (0,80; 1,32)
8. no

P15.14

1. Cuatro conjuntos intercambiables: fumadores menores de 60 años, fumadores de 60 años o más, no fumadores menores de 60 años y no fumadores de 60 años o más.
2. "¡Depende!" Si hay algún conjunto equiparado intercambiable, ciertamente el agrupamiento es adecuado. Sin embargo, hay tantas variables con respecto a las cuales se está equiparando que probablemente los datos del estudio contendrán solo unos pocos conjuntos de estratos intercambiables. En consecuencia, es posible que el uso de un análisis agrupado tenga un efecto insignificante en los cálculos del *odds ratio* y en la precisión en torno a dicho valor, cuando se compara con un análisis equiparado no agrupado.

P15.15

1. 1
2. 1
3. 2
4. 1
5. de McNemar
6. 2
7. no agrupado
8. 2,5
9. agrupado
10. agrupado
11. 2
12. 1
13. 0
14. 2
15. indefinido
16. 2,5
17. a

P15.16

1. Hay pocas pruebas de confusión, ya que los *odds ratios* crudo (4,09) y ajustado (4,10) son apenas distintos.
2. Hay pocas pruebas de interacción, puesto que

los *odds ratios* específicos para cada estrato son muy cercanos unos de otros (alrededor de 4). En consecuencia, una evaluación general de la relación entre la exposición y la enfermedad parece justificable.

3. Hay grandes cantidades en cada uno de los cuatro estratos y no hay celdas con cero en ninguno de ellos. En consecuencia, no se necesita un mOR (para solucionar el problema de las celdas con cero) y, por lo tanto, puede calcularse un aOR.
4. Existe aproximadamente el cuádruple de efecto de la exposición en la enfermedad. Este efecto es significativo (sobre la base de la prueba de MH) y el intervalo de confianza obtenido muestra un cálculo de efecto razonablemente preciso.

P15.17

1. Verdadero
2. c: tener celdas con cero en algunos estratos respalda el uso de un mOR para este o cualquier análisis estratificado. El equiparamiento por frecuencias no indica normalmente la necesidad de utilizar un mOR, pero, en este ejemplo, el uso del equiparamiento por frecuencias requiere un mOR a causa de, al menos, una frecuencia cero en el estrato 1 y probablemente otras frecuencias cero en otros estratos. Si se recurre a la regresión logística, entonces no se puede usar un mOR para calcular el *odds ratio*.
3. Falso. Si ignoramos el control de la variable que se equipara en un estudio de casos y controles, nuestro cálculo está sesgado hacia el valor nulo.
4. R_i a S_i: mediante el equiparamiento por frecuencias, la razón de equiparamiento de los controles con los casos generalmente será distinta para estratos diferentes.
5. No necesariamente. Debería esperar mayor precisión del equiparamiento por frecuencias en contraposición a lo que sucedería si no se efectúa el equiparamiento, pero no se garantiza de que el equiparamiento por frecuencias dará mayor precisión que el equiparamiento individual y viceversa.

P15.18

1. Se trata de un estudio de cohorte, donde normalmente utilizamos una razón de riesgos (RR) como la medida de efecto. Asimismo, ya que todos los estratos tienen celdas con cero, no es posible calcular *odds ratios* o razones de riesgos específicas para cada estrato, las cuales se deben promediar para obtener una aRR.

P15.19

1. El valor de 1,33 de la mRR indica una relación relativamente débil entre el tipo de turno de trabajo y el dolor lumbar crónico, controlando las variables que se equiparan. La prueba de MH tiene un valor p igual a 0,0680. En consecuencia, la estimación puntual tiene una significación límite y no es significativa al nivel de 0,05. El intervalo de confianza de 95% se encuentra entre un límite inferior (0,98) apenas debajo de 1 y un límite superior de tamaño moderado (1,81), apenas arriba de 1; por lo tanto, hay una imprecisión considerable en la estimación puntual. En general, los resultados sugieren que los turnos rotativos tienen un efecto relativamente débil.

P15.20

1. Falso. La cantidad total de expuestos en cada conjunto equiparado siempre es igual a 1.
2. Verdadero
3. 1,2
4. 2,5
5. (0,96; 1,50)

P15.21

1. El valor del mOR controla solo la variable que se equipara, es decir, no controla la variable BILIAR.
2. $$\text{logit } P(X) = b_0 + \sum_{i=1}^{62} b_i D_i + b_{63} E$$
3. El *odds ratio* ajustado es exp(2,269) = 9,67, que es el mismo valor que el mOR.
4. P = 0,0001, que es extremadamente pequeño,

por lo que rechazamos la hipótesis nula a los niveles de 0,05 y llegamos a la conclusión de que el uso de estrógenos es un predictor significativo de la condición de cáncer de endometrio, si controlamos las variables que se equiparan y la variable BILIAR no equiparada.

5. El estadístico chi cuadrado de McNemar ignora el control de la variable BILIAR, mientras que el estadístico chi cuadrado aquí controla dicha variable BILIAR.
6. El intervalo de confianza es bastante amplio, lo que indica la existencia de una imprecisión considerable en el *odds ratio* calculado.

P15.22

1. b
2. EMV condicional
3. exp(1,436)
4. sí
5. sí
6. exp[1,4361 $\pm$ 1,96(0,3167)]

ÍNDICE

1 A 1, *569*, *VEA EQUIPARAMIENTO*
ABANDONO DEL ESTUDIO, *225*
AGRUPAMIENTO, 586, 587, 588
AGRUPAMIENTO DE DATOS EQUIPARADOS, 586
AJUSTE POR EDAD, 98
ALEATORIZACIÓN, 30, 461, 466
ALEATORIZACIÓN, 31
ANÁLISIS AGRUPADO, 579, 587, 588, 611
ANÁLISIS BRUTO, 19
ANÁLISIS DE DATOS, 10
ANÁLISIS DE DATOS, 18
ANÁLISIS DE DATOS DE COHORTE EQUIPARADOS, 595
ANÁLISIS DE DATOS EQUIPARADOS – OPCIONES Y PRINCIPIOS GENERALES, 576
ANÁLISIS EQUIPARADO, 572, 577, 578, 579, 580, 581, 583, 585, 586, 588, 590, 596, 611
ANÁLISIS ESTRATIFICADO, 19, 25, 163, 214, 336, 337, 381, 382, 462, 464, 470, 471, 472, 490, 493, 494, 497, 499, 502, 505, 507, 509, 512, 520, 536, 544, 545, 557, 558, 572, 576, 577, 578, 579, 580, 581, 582, 584, 585, 586, 587, 592, 593, 594, 596, 597, 598, 599, 605, 606, 610, 612
ANÁLISIS NO AGRUPADO, 587
ANÁLISIS POR INTENCIÓN DE TRATAR, 36, 37, 222
ANÁLISIS SIMPLE, 385, 456
AOR, 526, 592
APAREADOS, 580, 581, 582, 585, 586, 589, 591, 595, 596, 597, 598, 599, 601, 611
ASOCIACIÓN, 17
BASADO EN LA PRECISIÓN, 522
BASADOS EN LA PRECISIÓN, 521
BASE DEL ESTUDIO, 199
BILATERAL, 395
BINOMIOS, 387
BROTES, 23, 116
CÁLCULO AJUSTADO, 245, 247, 252, 288, 299, 333, 334, 335, 336, 337, 340, 341, 342, 344, 360, 361, 364, 365, 366, 367, 369, 373, 380, 383, 384, 459, 462, 463, 471, 493, 497, 498, 499, 500, 516, 520, 521, 522, 531, 532, 536, 538, 546, 560, 561
CÁLCULO AJUSTADO DE MANTEL-HAENSZEL, 521
CÁLCULO CRUDO, 333, 365
CÁLCULO CRUDO DE EFECTO, 333, 334, 335, 341, 493
CÁLCULO DE MÁXIMA VEROSIMILITUD, 520
CÁLCULO VÁLIDO, 195
CÁLCULOS AJUSTADOS DE MANTEL-HAENSZEL, 530
CÁLCULOS AJUSTADOS MEDIANTE EL USO DE LA PONDERACIÓN POR LA PRECISIÓN, 522
CASOS CRUZADOS, 53, 58
CASOS INCIDENTES, 132
CASOS PREVALENTES, 132
CASOS Y COHORTE, 132
CASOS Y CONTROLES, 15, 23, 24, 33, 34, 35, 38, 43, 44, 45, 46, 47, 48, 49, 50, 53, 54, 55, 56, 57, 58, 62, 64, 65, 66, 67, 69, 112, 113, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 131, 132, 133, 134, 135, 136, 138, 140, 141, 146, 147, 154, 155, 161, 171, 180, 182, 183, 184, 187, 191, 195, 196, 202, 213, 216, 217, 221, 223, 224, 225, 236, 237, 238, 239, 240, 249, 252, 257, 258, 262, 263, 268, 269, 275, 279, 285, 291, 299, 316, 332, 334, 337, 339, 340, 341, 342, 343, 347, 353, 355, 356, 359, 363, 373, 375, 380, 381, 385, 386, 423, 424, 425, 426, 427, 428, 430, 431, 432, 433, 434, 443, 449, 450, 451, 452, 457, 459, 461, 464, 465, 466, 469, 470, 494, 501, 510, 512, 518, 521, 526, 528, 530, 532, 533, 534, 539, 543, 553, 555, 556, 559, 565, 566, 567, 568, 569, 570, 571, 572, 573, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 592, 593, 594, 595, 598, 599, 601, 603, 605, 606, 607, 608, 609, 610, 611, 612
CASOS Y CONTROLES ANIDADOS, 53, 55, 56, 249
CAUSALIDAD, 31, 34
CELDA A, 411, 414, 427
CELDAS CON CERO, 493, 520, 521, 530, 531, 533, 534, 561, 594, 610, 612
CHI CUADRADO, 386, 409, 509
CLASIFICACIÓN ERRÓNEA, 264
CLASIFICACIÓN ERRÓNEA DIFERENCIA, 299
CLASIFICACIÓN ERRÓNEA DIFERENCIAL, 279, 318, 355
CLASIFICACIÓN ERRÓNEA NO DIFERENCIAL, 276, 277, 278, 279, 280, 283, 286, 288, 290, 291, 293, 294, 295, 296, 297, 298, 299, 300, 313, 314, 324, 325, 326
CLASIFICACIÓN INDEPENDIENTE. *VER CLASIFICACIÓN ERRÓNEA NO DIFERENCIAL*

COHORTE, 14, 38, 222
COHORTE *CAMBIADA*, 76
CONCORDANTES, 581, 582, 595
CONFUSIÓN, 17, 18, 328, 329, 330, 332, 333, 362, 364, 458, 467, 572
CONFUSIÓN, 343, 370
CONFUSIÓN BASADA EN LOS DATOS, 329, 338, 339, 341, 342, 357, 360, 365, 366, 369, 578
CONFUSIÓN CONJUNTA BASADA EN LOS DATOS, 365
CONFUSIÓN CONJUNTA VERSUS MARGINAL. *CONFUSIÓN*
CONFUSIÓN MARGINAL BASADA EN LOS DATOS, 365
CONJUNTOS EQUIPARADOS INTERCAMBIABLES, 586
CONTROL, 362, 370, 456, 458, 460, 572
CONTROLES, 45, 239
CONTROLES DE BASE HOSPITALARIA, 45
CONTROLES DE BASE POBLACIONAL, 45
COVARIABLES, 11
CRITERIOS PARA EL SESGO DE CONFUSIÓN, 338
CUESTIONES ÉTICAS, 36
DENSIDAD DE INCIDENCIA, 80, 82
DESVENTAJA. *VEA EQUIPARAMIENTO*
DICOTÓMICA, 477
DIFERENCIA DE PREVALENCIA, 154
DIFERENCIA DE RIESGOS, 150, 154, 155
DIFERENCIA DE TASAS, 154
DIRECCIÓN DEL SESGO, 209, 217
DIRECCIONALIDAD, 33, 34, 38, 39, 42, 44, 54, 57, 66
DISCORDANTES, 595
DISEÑO DE CASOS CRUZADOS, 58
DISEÑO DE CASOS Y CONTROLES ANIDADOS, 55
DISEÑO DE ESTUDIO, 9, 14, 38, 43
DISEÑOS DE ESTUDIOS, 340
DISEÑOS DE ESTUDIOS OBSERVACIONALES, 38
DISEÑOS HÍBRIDOS, 53
DISEÑOS INCOMPLETOS, 58
DISTRIBUCIÓN, 392
DISTRIBUCIÓN HIPERGEOMÉTRICA, 412, 413, 428, 510, 512
DISTRIBUCIÓN MUESTRAL, 393, 395
DISTRIBUCIÓN NORMAL, 391, 393, 395, 396, 400, 406, 407, 408, 409, 411, 418, 419, 424, 426, 431, 437, 441, 452, 453, 499, 536
DISTRIBUCIÓN NORMAL ESTÁNDAR, 393
DOS COLAS, 391, 393, 395, 396, 397, 398, 399, 406, 408, 409, 410, 416, 421, 426, 437, 452, 453, 509, 511, 512, 517, 564
ECOLÓGICOS, 38, 58, 59, 61, 67
EFECTO, 18
EFECTO AJUSTADO, 348, 467, 499, 500, 501, 509
EFECTO DEL TRABAJADOR SANO, 228
EFECTO DOSIS-RESPUESTA, 19, 31, 356
EFICACIA, 30, 36
EFICIENCIA DE PITTMAN, 570
ENFERMEDAD POCO FRECUENTE, 126, 132, 134, 234
ENMASCARAMIENTO, 36
ENSAYO CLÍNICO, 36
ENSAYOS DE INTERVENCIÓN COMUNITARIA, 30, 31
EOR, 123, 124, 424
EPIDEMIOLOGÍA, 7
EQUIPARAMIENTO, 461, 462, 465, 469, 470, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 584, 585, 586, 587, 588, 592, 593, 594, 595, 598, 599, 600, 601, 602, 604, 605, 606, 607, 608, 609, 610, 611, 612
EQUIPARAMIENTO 1 A 1, 570, 609
EQUIPARAMIENTO INDIVIDUAL, 469, 470, 566, 568, 569, 570, 573, 588, 594, 612
EQUIPARAMIENTO POR CATEGORÍAS, 566
EQUIPARAMIENTO POR FRECUENCIAS, 469, 470, 566, 568, 569, 570, 588, 612
ERROR ALEATORIO, 196, 198, 205, 206, 208, 209, 216, 218, 219, 239, 259, 348, 386, 459, 607
ERROR ESTÁNDAR, 400
ERROR SISTEMÁTICO, 196, 197, 198, 205, 206, 208, 216, 218, 219, 221, 223, 264, 348
ESPECIFICIDAD, 271, 272, 273, 274, 276, 278, 280, 282, 283, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 299, 301, 304, 305, 307, 308, 309, 310, 312, 317, 318, 320, 321, 322, 323, 324, 325, 326, 327, 359
ESTADÍSTICA DE LA RAZÓN DE VEROSIMILITUD, 552, 600
ESTADÍSTICO, 386, 387, 390, 391
ESTADÍSTICO DE PRUEBA, 392
ESTADÍSTICO DE PRUEBA T, 394
ESTADO ESTACIONARIO, 70, 71, 138, 139, 141
ESTÁNDAR DE ORO, 302, 303
ESTIMACIÓN, 385, 386
ESTIMACIÓN DE INTERVALOS, 386, 387, 389, 390, 399, 453
ESTIMACIÓN DE MÁXIMA VEROSIMILITUD, 482, 485, 600, 601, 602

ESTIMACIÓN PUNTUAL, 386, 389, 399, 449, 454, 471, 472, 493, 494, 497, 499, 593, 606, 611, 612
ESTRATIFICACIÓN, 10, 18, 19, 20, 31, 279, 325, 348, 462, 465, 466, 493, 498, 499, 504, 509, 561
ESTRATO EQUIPARADO, 569
ESTUDIO DE CASO-COHORTE, 54
ESTUDIO DE CASOS Y CONTROLES
DESVENTAJAS, 44
VENTAJAS, 44
ESTUDIO DE CASOS Y CONTROLES ANIDADOS, 56
ESTUDIO DE CASOS Y CONTROLES CON MUESTREO POR DENSIDAD, 55
ESTUDIO DE CASOS Y CONTROLES EQUIPARADOS, 58, 583, 608
ESTUDIO DE COHORTE, 38, 39, *VEA PERSONA-TIEMPO*
DESVENTAJAS, 40
VENTAJAS, 39
ESTUDIO DE FRAMINGHAM, 38
ESTUDIO DE MORBILIDAD PROPORCIONAL, 59
ESTUDIO DE MORTALIDAD PROPORCIONAL, 60
ESTUDIO ECOLÓGICO, 58, 59, 61, 67
ESTUDIO OBSERVACIONAL, 30
ESTUDIO TRANSVERSAL, 15, 43, 50, 51, 52, 53, 67, 90, 123, 124, 125, 198, 202, 218, 223, 241, 242, 243, 341, 349, 421, 501, 502
ESTUDIOS ANALÍTICOS, 30
ESTUDIOS DE COHORTE, 39
ESTUDIOS DE PRUEBAS DE TAMIZAJE, 304
ESTUDIOS DESCRIPTIVOS, 30
ESTUDIOS EXPERIMENTALES, 30, 31, 466, 467, 469, 470
ESTUDIOS OBSERVACIONALES, 30, 31, 38, 249, 258, 466, 467
ESTUDIOS PROPORCIONALES, 59
EXCESO DE RIESGO, 150, 151, 152, 153, 155, 159, 164, 528
EXPOSICIÓN, 8- 66
FACTOR DE CONFUSIÓN, 338
FACTOR DE RIESGO, 338
FACTOR POTENCIAL DE CONFUSIÓN, 348
FACTORES DE CONFUSIÓN, 99, 354, 356, 372, 375, 376, 380, 383, 458, 470, 575
FACTORES DE RIESGO, 338, 339
FALACIA ECOLÓGICA, 59
FALTA DE RESPUESTA, 17, 45, 214, 222, 223, 225, 226, 227, 241, 242, 243, 262
FÓRMULA BASADA EN LA PRECISIÓN, 524
FRACCIÓN ETIOLÓGICA, 164
FRACCIÓN ETIOLÓGICA ENTRE LOS EXPUESTOS, 169, 174, 182, 193
FRACCIÓN PREVENIDA, 162, 174
FRACCIÓN PREVENIDA ENTRE LOS EXPUESTOS, 177
FRECUENCIA DE ENFERMEDAD, 9, 15, 68
FRECUENCIAS DE CELDA CORREGIDAS, 286
FUNCIÓN DE RIESGO, 481
FUNCIÓN DE RIESGO, 479
GRUPO DE COMPARACIÓN. *VEA GRUPO ÍNDICE*
GRUPO DE CONTROL, 237
GRUPO ÍNDICE, 565
HACIA ADELANTE, 33
HACIA ATRÁS, 33, 44
HACIA EL VALOR NULO, 209, 240, *VEA DIRECCIÓN DEL SESGO*
HIPÓTESIS ALTERNATIVA, 388, 391, 392, 393, 394, 395, 396, 397, 398, 399, 405, 406, 408, 409, 410, 415, 421, 423, 425, 426, 436, 437, 452, 453, 509, 511, 512, 517, 548, 552, 562
HIPÓTESIS CONCEPTUAL, 30
HIPÓTESIS NULA, 64, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 405, 406, 407, 410, 412, 413, 414, 415, 421, 423, 424, 425, 426, 427, 428, 436, 437, 438, 450, 452, 453, 454, 482, 491, 492, 501, 504, 505, 509, 510, 512, 513, 517, 518, 519, 548, 549, 552, 553, 556, 557, 558, 559, 560, 561, 562, 564, 600, 601, 613
HIPÓTESIS OPERACIONAL, 30
IC. *VEA INTERVALO DE CONFIANZA*
IMPACTO POTENCIAL, 162, 163
INCIDENCIA, 68, 69, 71
INCIDENCIA ACUMULADA, 76, 77, 93, 94, 406
INCIDENCIA ACUMULADA, 74
INCIDENCIA ACUMULADA SIMPLE, 74, 75
INFERENCIA ESTADÍSTICA, 143, 198, 385, 386, 387, 388, 390, 588
INFERENCIAS CAUSALES. *VEA CAUSALIDAD*
INTERACCIÓN, 343, 116–37, 362, 458, 472, 500
INTERCAMBIABILIDAD, 586
INTERVALO DE CONFIANZA, 399, 401, 499
LA PRUEBA DE MANTEL-HAENSZEL, 509, *VEA ANÁLISIS ESTRATIFICADO*
LATENCIA, 40
LETALIDAD, 74, 94
LIBRO DE ESTUDIO, 5
LOGARITMO NATURAL, 418
LOGIT, 478

MÁRGENES FIJOS, 427, 512
MÁXIMA VEROSIMILITUD, 520
MÁXIMA VEROSIMILITUD CONDICIONAL, 600
MEDIDA AJUSTADA DE EFECTO, 286
MEDIDA DE EFECTO, 9, 15, 17, 112, 113, 132, 195, 264, 385, 520
MEDIDA DE EFECTO CORREGIDA, 290
MEDIDA DE EFECTO DE DIFERENCIAS, 520
MEDIDA DE EFECTO DE RAZONES, 520
MEDIDA DE FRECUENCIA DE ENFERMEDAD, 9, 15, 68
MEDIDA DE IMPACTO POTENCIAL, 162
MÉTODO DE DOBLE CIEGO, 37
MÉTODO DIRECTO, 98
MÉTODO DIRECTO DE AJUSTE. *VEA TASA*
MÉTODOS DE MANTEL-HAENSZEL, 576
MODELACIÓN MATEMÁTICA, 10, 18, 214, 472
MODELO MATEMÁTICO, 347, 463, 472, 475
MODIFICACIÓN DE EFECTO, 343, 347, 348, 349, 459
MODIFICADOR DE EFECTO, 347
MOR, 532, 579, 592, *VEA ODDS RATIO DE MANTEL-HAENSZEL*
MORTALIDAD, 94
MORTALIDAD POR CUALQUIER CAUSA, 94
MORTALIDAD POR ENFERMEDAD ESPECÍFICA, 94
MRDI, 534, 535
MRR, 534, 535, 596
MUESTRA, 199, 386
MUESTRA GRANDE, 407, 536
MUESTREO POR DENSIDAD, 55
NIVEL DE SIGNIFICACIÓN, 392
NNT, 156, *VEA NÚMERO NECESARIO A TRATAR*
NO DIFERENCIAL. *VEA CLASIFICACIÓN ERRÓNEA*
NO DIRECCIONAL, 34
NÚMERO NECESARIO A TRATAR, 155, 156
ODDS, 117
ODDS RATIO, 112, 125, 116–37, 231, 423
ODDS RATIO AJUSTADO, 248, 599
ODDS RATIO AJUSTADO BASADO EN LA PRECISIÓN, 526
ODDS RATIO BASADO EN LA PRECISIÓN, 531
ODDS RATIO DE EXPOSICIÓN, 258
ODDS RATIO DE EXPOSICIÓN AJUSTADO, 341
ODDS RATIO DE MANTEL-HAENSZEL, 531, 532, 539, 584, 598
ODDS RATIO DE PREVALENCIA, 123, 526
ODDS RATIO DE PREVALENCIA AJUSTADO, 341
ODDS RATIO DE RIESGO, 123, 124, 125, 126, 127, 134, 136, 526
OPCIONES DE ANÁLISIS, 462, *VEA ANÁLISIS ESTRATIFICADO*
OR, 116–37
PAR CONCORDANTE, 581, 595
PAR DISCORDANTE, 581, 595
PARADOJA DE SIMPSON, 328–30
PARÁMETRO, 386, 391
PARÁMETRO DE SELECCIÓN, 244
PARÁMETRO OBJETIVO, 206
PÉRDIDA DE SEGUIMIENTO, 225
PÉRDIDAS DE SEGUIMIENTO, 89
PERÍODO DE RIESGO, 73
PERSONAS-AÑOS, 81
PERSONA-TIEMPO, 71, 82, 85, 137–41
PLAN DE MUESTREO, 9, 14
POBLACIÓN, 386
POBLACIÓN CANDIDATA, 59
POBLACIÓN DE ESTUDIO, 199, 206, 237
POBLACIÓN DINÁMICA ESTABLE, 83
POBLACIÓN ESTABLE, 241
POBLACIÓN ESTÁNDAR, 98
POBLACIÓN EXTERNA, 202
POBLACIÓN FUENTE, 43, 45, 55
POBLACIÓN OBJETIVO, 199
PONDERACIÓN LINEAL, 521, 536
PONDERACIÓN LOGARÍTMICA LINEAL, 521, 536, 538
POR, 123
PORCENTAJE DE RIESGO ATRIBUIBLE DE LA POBLACIÓN, 164
PORCENTAJE DE RIESGO ATRIBUIBLE ENTRE LOS EXPU*ESTOS*, 169
PORCENTAJE DE RIESGO PREVENIDO DE LA POBLACIÓN, 174
PORCENTAJE DE RIESGO PREVENIDO ENTRE LOS EXPUESTOS, 178
PRECISIÓN, 195, 196, 375, 376, 386, 458, 521, 569, 572, 575
PREGUNTA DE ESTUDIO, 10, 13
PREVALENCIA, 68, 70, 71, 90, 154, 304
PREVALENCIA DE PERÍODO, 90, 93
PREVALENCIA PUNTUAL, 90
PROBABILIDAD CONDICIONAL, 134
PROBABILIDAD DE SELECCIÓN, 229
PROBABILIDAD POSPRUEBA, 306
PROBABILIDAD PROMEDIO PREVIA A LA PRUEBA, 306

PROMEDIO PONDERADO, 335, 499, 520
PROPORCIONAL, 38
PROSPECTIVA. *VEA SECUENCIA TEMPORAL*
PROSPECTIVO. *VEA DISEÑO DE ESTUDIO*, *VEA DISEÑO DE ESTUDIO*
PRUEBA DE BRESLOW DAY, 348, 504, 505
PRUEBA DE CHI CUADRADO, 409
PRUEBA DE CHI CUADRADO DE MANTEL, 497
PRUEBA DE CHI CUADRADO DE MANTEL-HAENSZEL, 471
PRUEBA DE DETECCIÓN, 305
PRUEBA DE DIAGNÓSTICO, 302
PRUEBA DE HIPÓTESIS, 388, 391, 405, 411, 437, 498
PRUEBA DE HIPÓTESIS DE MH, 580, 582
PRUEBA DE MANTEL-HAENSZEL, 414, 509, 585, 597, 598
PRUEBA DE RAZÓN DE VEROSIMILITUD, 552
PRUEBA DE SIGNIFICACIÓN. *VEA PRUEBA DE HIPÓTESIS*
PRUEBA DE TENDENCIA, 546, 554
PRUEBA EXACTA DE FISHER, 411, 412, 413, 414, 427, 428
R A 1, 569, *VEA RAZÓN DE EQUIPARAMIENTO*
RAZÓN DE DENSIDADES DE INCIDENCIA, 137, 138, 435
RAZÓN DE EQUIPARAMIENTO, 569
RAZÓN DE PRODUCTOS CRUZADOS, 121, 234
RAZÓN DE RIESGOS, 17, 19, 112, 113, 125, 150, 234, 247, 405, 417
RAZÓN DE RIESGOS AJUSTADA, 341, 522
RAZÓN DE RIESGOS CRUDA, 333, 341, 367
RAZÓN DE RIESGOS INSTANTÁNEOS, 137
RAZÓN DE TASAS, 137, 138, 137–41, 435
RAZONES DE SELECCIÓN, 229
RDI, 137, 142, 386, 435
REGIÓN DE ACEPTACIÓN, 393
REGIÓN DE RECHAZO, 393
REGLA DE DECISIÓN, 393
REGRESIÓN LOGÍSTICA, 576, 598
RELACIÓN DE DOSIS-RESPUESTA, 26, 545, 546
RESTRICCIÓN, 461, 467
RESULTADO DE SALUD, 7, 10, 19, 23, 30, 330
RETROSPECTIVA. *VEA SECUENCIA TEMPORAL*
RETROSPECTIVO, 43, *VEA SECUENCIA TEMPORAL*
R_I A 1. *VEA RAZÓN DE EQUIPARAMIENTO*
RI A SI. *VEA RAZÓN DE EQUIPARAMIENTO*, *VEA RAZÓN DE EQUIPARAMIENTO*
RIESGO, 16, 71, 72, 74, 89, 92, 94, 150
RIESGO ATRIBUIBLE, 151
RIESGO ATRIBUIBLE DE LA POBLACIÓN, 164
RIESGO DE MORTALIDAD, 94
RIESGO INSTANTÁNEO, 80
RIESGO POBLACIONAL PREVENIDO, 174
ROR, 123, 124, 125
RR, 117, 125, 137, 386
SECUENCIA TEMPORAL, 34
SEGUIMIENTO, 38, 121
SENSIBILIDAD, 271, 272, 273, 274, 276, 278, 280, 282, 283, 285, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 299, 300, 301, 304, 305, 307, 308, 309, 310, 312, 317, 318, 320, 321, 322, 323, 324, 325, 326, 327, 359
SESGO, 8, 17, 39, 43, 50, 132, 195, 206, 213, 237, 264, 328
SESGO DE BERKSON, 195, 239, 240, 257
SESGO DE CONFUSIÓN, 17, 196, 213
SESGO DE INFORMACIÓN, 17, 195, 208, 264, 286
SESGO DE SELECCIÓN, 17, 208, 221
SESGO DE TRANSICIÓN, 210
SISTEMÁTICAMENTE, 207
SOBRE-EQUIPARAMIENTO, 573
SUBCONJUNTO CANDIDATO DE FACTORES DE CONFUSIÓN, 370
SUPERVIVENCIA SELECTIVA, 223
SUPUESTO DE ENFERMEDAD POCO FRECUENTE, 126
T DE STUDENT, 386
TABLA DE MCNEMAR, 582, 597
TASA, 80, 89, 94, 137, 138, 155
TASA AJUSTADA POR EDAD, 97
TASA ATRIBUIBLE, 155
TASA DE INCIDENCIA, 82
TASA DE MORTALIDAD, 74, 95
TASA INSTANTÁNEA, 80
TASA PROMEDIO, 80, 82, 83
TENDENCIA LINEAL, 548
TEOREMA DE BAYES, 135
THETA, 520
TITANIC, 390
UNA COLA, 391, 395, 397, 398, 399, 406, 408, 409, 410, 412, 415, 416, 421, 425, 426, 437, 449, 452, 453, 454, 509, 511, 512, 517, 519, 554, 562
UNILATERAL, 395, 409
VALIDEZ, 195, 196, 363, 375
VALIDEZ EXTERNA, 203
VALIDEZ INTERNA, 203
VALOR ESPERADO, 475

VALOR NULO, 114, 151
VALOR P, 395, 409, 498
VALOR PREDICTIVO, 304, 305, 306, 307, 308, 309, 310, 311, 312, 321, 322, 326, 327
VALOR PREDICTIVO NEGATIVO, 305
VALOR PREDICTIVO POSITIVO, 305
VARIABILIDAD, 18
VARIABILIDAD MUESTRAL, 386
VARIABLE, 13
VARIABLE ALEATORIA, 402
VARIABLE CONTÍNUA, 264
VARIABLE DE CONTROL, 9, 19, 20, 23, 343, 456
VARIABLE DE EXPOSICIÓN. *VEA EXPOSICIÓN*
VARIABLE DEPENDIENTE, 463
VARIABLE FICTICIA, 551
VARIABLE INDEPENDIENTE, 463
VARIABLE INTERVINIENTE, 338
VARIABLES, 7, 8
VARIABLES DE CONTROL, 8, 10, 11, 12, 14, 21, 24, 25, 213, 332, 347, 348, 351, 365, 366, 381, 382, 458, 459, 461, 462, 463, 464, 465, 467, 468, 469, 471, 472, 475, 478, 482, 485, 487, 560, 572, 573, 599, 602
VARIABLES DE EXPOSICIÓN, 23
VARIABLES EXTERNAS, 456, 457, 458, 462, 466, 490, 493, 572, 607
VARIANZA, 431, 537
VENTAJA. *VEA EQUIPARAMIENTO*
VP-. *VEA VALOR PREDICTIVO NEGATIVO*
VP+. *VEA VALOR PREDICTIVO POSITIVO*