



Artículo especial

## ¿Cómo mejorar los intervalos de confianza?

### How can we improve the confidence intervals?

Miguel Ángel Martínez-González<sup>a,\*</sup>, María Seguí-Gómez<sup>a,b</sup> y Miguel Delgado-Rodríguez<sup>a,c</sup>

<sup>a</sup> Departamento de Medicina Preventiva y Salud Pública, Facultad de Medicina Clínica, Universidad de Navarra, Pamplona, España

<sup>b</sup> European Center for Injury Prevention, Universidad de Navarra, Pamplona, España

<sup>c</sup> Medicina Preventiva y Salud Pública, Universidad de Jaén, Jaén, España

#### INFORMACIÓN DEL ARTÍCULO

*Historia del artículo:*

Recibido el 10 de marzo de 2009

Aceptado el 22 de octubre de 2009

On-line el 24 de diciembre de 2009

El propósito de este artículo es defender el uso de los intervalos de confianza (IC) en las publicaciones médicas y explicar un método bayesiano de cálculo simple y aproximado para mejorar los IC. Este método se puede aplicar usando los programas estadísticos convencionales.

#### La superioridad del intervalo de confianza

Los valores *p* de significación estadística (contraste de hipótesis) de los que se suele abusar en las publicaciones médicas se han criticado ampliamente<sup>1-6</sup> y parece lógico dar mayor importancia a los IC<sup>2-4,6</sup>, que son una alternativa mucho más directa y fácil de entender. Las normas STROBE («Strengthening the reporting of observational studies in epidemiology»)<sup>7</sup> indican explícitamente que, al escribir artículos sobre estudios epidemiológicos observacionales, se presenten las medidas de asociación ajustadas y acompañadas de sus IC. Las normas CONSORT, elaboradas por los editores de revistas biomédicas, anteponen el uso del IC a los valores *p* para presentar los resultados de ensayos clínicos<sup>8</sup>.

Los IC consisten simplemente en presentar los resultados e indicar un rango para el resultado que es creíble a partir de los datos recogidos en el estudio. La amplitud de un IC y los valores de sus límites facilitan la interpretación de los datos. Supóngase que se trabaja con el riesgo relativo (RR). El RR es el número de veces que es más probable que se produzca una enfermedad en un grupo expuesto a un factor frente a un grupo no expuesto. Si estudiamos el efecto de la actividad física como factor determinante de enfermedad mental y el RR es de 0,70, esto indica que hacer ejercicio protege (es menor el riesgo de enfermedad en los que

practican actividad física que en los que no la practican). Para su correcta interpretación se requiere conocer el IC. Si el IC oscila entre 0,40 y 1,2,3 este intervalo incluye el RR=1 (ausencia de relación entre actividad física y enfermedad mental) y no permite concluir nada sobre la relación entre ejercicio físico y enfermedad mental. Por el contrario, si el IC oscila entre 0,5 y 0,9, que excluye el valor de RR=1, se saca la conclusión de que hacer ejercicio protege de la enfermedad mental y se puede hacer una recomendación a la población. Está claro que los IC ayudan a fundamentar decisiones prácticas en salud<sup>9</sup>.

Los resultados presentados en la tabla 1 proceden de una investigación epidemiológica prospectiva española. Se trata del proyecto Seguimiento Universidad de Navarra (cohorte SUN) que valoró la relación entre la actividad física en el tiempo libre y el riesgo de desarrollar enfermedad mental<sup>10</sup>. En esta tabla se explica cómo calcular el IC para la *odds ratio*<sup>11,12</sup> (OR). La OR es análoga al RR, y se interpreta también como el número de veces que es más probable que un efecto (la enfermedad mental) sea más frecuente en un grupo de expuestos a un factor (hacer ejercicio) frente a un grupo de no expuestos (no hacer ejercicio). La OR estimada da un valor de 0,7, lo que indica que el ejercicio protege frente a la enfermedad mental. Los límites del IC para esa OR son 0,57 y 0,87, lo que indica que confiamos en que el valor real de la OR se encuentre entre 0,57 y 0,87. Este intervalo excluye la OR=1, por lo que tenemos cierta confianza de que la OR real no será 1. Por tanto, podemos afirmar que hacer ejercicio no es igual que no hacerlo en cuanto a presentar una enfermedad mental en el futuro.

#### El pensamiento bayesiano

Los cálculos del IC que se han ofrecido en la tabla 1 se han realizado según el concepto frecuentista de probabilidad, que es el habitualmente utilizado. Pero en realidad existen 2 corrientes

\* Autor para correspondencia.

Correo electrónico: mamartinez@unav.es (M.A. Martínez-González).

Tabla 1

Odds ratio de enfermedad mental en niveles extremos (quintil 5 comparado con quintil 1) de actividad física en el tiempo libre. Proyecto SUN (1999-2005)<sup>10</sup>. Cálculo del intervalo de confianza para la odds ratio

	Enfermedad mental (casos incidentes)	
	Sí	No
Actividad física en tiempo libre		
Alta ( $\geq 32,95$ METS-h/sem)	156 (a)	1.929 (b)
Baja ( $\leq 3,27$ METS-h/sem)	235 (b)	2.048 (d)
OR	$\frac{a \times d}{b \times c} = \frac{156 \times 2.048}{1.929 \times 235} = 0,70$	
Error estándar de OR	$\sqrt{1/a+1/b+1/c+1/d} = \sqrt{1/156+1/1.929+1/235+1/2.048} = 0,11$	
Intervalo de confianza del 95% para la OR	$0,70 \times \text{EXP}(\pm 1,96 \times 0,11) = 0,57 \text{ a } 0,87$	

METS-h/sem: equivalentes metabólicos por horas a la semana; OR: odds ratio.

que definen formalmente la probabilidad; la frecuentista y la bayesiana. La definición frecuentista mantiene que las probabilidades son los límites a los que tiende la proporción con la que sucede un acontecimiento. Por ejemplo, si se lanza una moneda un número muy grande de veces, la proporción con la que saldrá cara tenderá a ser del 50% (probabilidad=0,5). Así, el pensamiento frecuentista define que la probabilidad de que ocurra un suceso sería la proporción que se obtendría si el número de repeticiones fuese infinito. Este método, al llevar implícita la necesidad de repeticiones, no puede estimar la probabilidad individual de un suceso único.

Hay preguntas para las que no existen posibilidades de repetir la experiencia. Por ejemplo, ¿cuál es la probabilidad de que un paciente concreto adquiera hoy una infección hospitalaria? Solo habrá una ocasión, que son las próximas 24 h. Aunque haya posibilidad de repetición, por ejemplo al día siguiente o a los 2 días, esas repeticiones son respuestas a otras preguntas distintas. Este problema vendría a ser la versión cuantitativa y formalizada del viejo aforismo clínico de que «no hay enfermedades, sino enfermos»<sup>8</sup>. La respuesta frecuentista es aplicarle a cualquier paciente lo que ha ocurrido previamente con pacientes similares en días previos. Esta extrapolación, que es en principio válida, aunque vaya acompañada de un cierto grado de incertidumbre, prescinde del hecho de que no hay 2 seres humanos iguales.

El otro enfoque del pensamiento en estadística es la filosofía bayesiana<sup>9</sup>, que maneja la probabilidad como un concepto «subjetivo». Este método añade a la interpretación frecuentista el grado de certeza previa del investigador antes de conocer los resultados. Sería como el grado de credibilidad que se tiene de antemano sobre un fenómeno o hasta qué punto se está dispuesto a predecir cada posible valor del resultado.

Frecuentistas y bayesianos mantienen actualmente una cordial disputa en el mundo de la estadística. Hay más frecuentistas que bayesianos<sup>13</sup>, aunque los bayesianos parecen ir ganando terreno en medicina<sup>9</sup>. De hecho, varias revistas médicas generales han revelado decididamente su predisposición a publicar preferentemente artículos que usen métodos bayesianos<sup>14</sup>. A los bayesianos se les denuncia cierto carácter arbitrario o irracional en sus probabilidades subjetivas (que se suelen llamar «a priori»), ya que estas probabilidades no son directamente medibles. Por el contrario, se les podría criticar a los frecuentistas que la verificación frecuentista de la probabilidad empírica nunca se ha hecho. Nadie ha lanzado una moneda al aire infinitas veces.

Una interpretación habitual del valor p frecuentista, que no sólo es cándida, sino también muy equivocada, es creer que los valores p habitualmente presentados en las publicaciones médi-

cas representan la probabilidad de que la hipótesis nula ( $H_0$ ) sea cierta. De ninguna manera es así. Si la  $H_0$  es ausencia de relación y la hipótesis alternativa ( $H_1$ ) es presencia de relación, el valor p de significación estadística es la probabilidad de obtener, sólo por azar, en una muestra los datos observados o unos todavía más alejados de la  $H_0$ , pero no es la probabilidad de que tal  $H_0$  sea cierta. La probabilidad de que una hipótesis sea cierta no es calculable con los métodos frecuentistas ni con ningún otro procedimiento porque, por desgracia, al trabajar con muestras se desconoce la realidad a nivel poblacional, sólo se estima.

Sin embargo, en muchas situaciones podría calcularse una probabilidad bayesiana o una credibilidad para la  $H_1$  que se pone a prueba. Se verá luego cómo se hace algo similar para estimar IC.

Los métodos bayesianos se han descrito como procedimientos para mostrar cómo las creencias previas se alteran una vez conocidos los resultados de un estudio<sup>9</sup>.

Siguiendo con el ejemplo anterior, uno podría preguntarse antes de ver algún dato, ¿cuál es la creencia previa en que una vida físicamente activa pueda prevenir la enfermedad mental? Ésta sería la probabilidad a priori de la hipótesis. El investigador ha de anticipar una predicción, que depende de lo que él piensa o conoce hasta ese momento.

Esta creencia es, en cierto modo, subjetiva. Por ejemplo, si se otorgase una probabilidad del 50% de que la  $H_0$  sea cierta (ausencia de relación entre ejercicio y enfermedad mental), existiría también una probabilidad del 50% de que la  $H_1$  fuese cierta (hay una relación entre ejercicio y enfermedad mental). Una creencia del 50% en cada hipótesis indica la duda completa del investigador sobre la existencia de relación alguna entre ambas variables. El cociente de ambas probabilidades sería 1 (50%/50%=1) y recibe el nombre de «odds a priori». A continuación se toma la muestra y se calcula una cantidad, llamada factor Bayes<sup>14</sup>, que mide cuántas veces es más fuerte en los datos recogidos la prueba a favor de la  $H_0$  que la de la  $H_1$ . El producto de la odds a priori por el factor Bayes es la «odds a posteriori». Esta odds a posteriori sí es traducible directamente como una probabilidad de que la hipótesis sea cierta, algo que no sucede con los métodos frecuentistas (tabla 2). En este planteamiento están algunas de las ventajas de los métodos bayesianos, pero también las razones por las que hay investigadores que los contemplan con sospecha, ya que les parece que lo científico sería solamente considerar lo que los datos dicen y no cómo cambian las creencias tras observar los datos.

El uso del factor Bayes<sup>14</sup> puede superar las limitaciones en la interpretación del valor p frecuentista. Se ha demostrado que hallazgos «significativos» apoyados por valores p significativos ( $p < 0,05$ ) pero cercanos a 0,05 tienden a no replicarse en estudios futuros<sup>15</sup>. Además, y de modo coherente, los factores Bayes interpretan estos valores p próximos a 0,05 como una prueba sólo relativamente débil contra la  $H_0$ . Los valores p siguen siendo populares porque cuando son muy pequeños ( $< 0,01$ ) no suelen conducir a conclusiones equivocadas y, sobre todo, porque no necesitan hacer explícitas las creencias previas (odds a priori) ni

<sup>8</sup> Gregorio Marañón, gran médico humanista, no atacaba la inferencia estadística (que necesita definir grupos homogéneos de pacientes para aplicar a unos de ellos los resultados observados en otros), sino que insistía en la existencia de personas, de seres humanos, detrás de las patologías que éstos presentaban.

**Tabla 2**  
Esquema del pensamiento bayesiano

Creencia de que la H <sub>0</sub> es correcta (H <sub>0</sub> =ausencia de relación entre exposición y efecto)	p
Creencia de que la H <sub>1</sub> es correcta (H <sub>1</sub> =existencia de relación entre exposición y efecto)	q = 1 - p
Odds a priori de la creencia en que H <sub>0</sub> es cierta	p/q
Factor Bayes	$\frac{p(\text{datos} H_0)}{p(\text{datos} H_1)}$
Odds a posteriori de H <sub>0</sub>	odds a priori de H <sub>0</sub> × Factor Bayes = (p/q) $\frac{p(\text{datos} H_0)}{p(\text{datos} H_1)}$

H<sub>0</sub>: hipótesis nula; H<sub>1</sub>: hipótesis alternativa.

**Tabla 3**  
Odds ratio posterior de enfermedad mental en niveles extremos de actividad física en el tiempo libre. Proyecto SUN (1999-2005)<sup>10</sup>. Cálculo del intervalo de confianza según una aproximación bayesiana simple<sup>17</sup>

Concepto	Ejemplo	Interpretación/resolución
<b>A) A priori</b>		
Apuesta 1:	¿Para qué valor «x» se considera «a priori» que es equiprobable que OR > x y OR ≤ x?	OR <sub>priori(mediana)</sub> = 0,9 Se cree tan probable que OR > 0,9 como que OR ≤ 0,9
Apuesta 2:	¿Entre qué dos valores (LI, LS) se considera «a priori» que hay una probabilidad del 95% de que esté la verdadera OR?	LI (OR <sub>priori</sub> ) = 0,27* LS (OR <sub>priori</sub> ) = 3,00 (debe cumplirse que OR <sup>2</sup> = LI • LS) Se está el 95% seguro de que la verdadera OR estará entre 0,27 y 3,0 (0,9 <sup>2</sup> /3 = 0,27)
Obtener la variancia a priori	$\text{Variancia} = \left( \frac{\ln \frac{LS}{LI}}{2z} \right)^2$	$\text{Variancia} = \left( \frac{\ln \frac{3,0}{0,27}}{2 \times 1,96} \right)^2 = 0,38$
Obtener la IP	siendo z el valor de la distribución normal, para una p=0,05; z=1,96 IP = 1/variancia	IP = 1/0,38 = 2,6
<b>B) Estudio</b>		
Lo encontrado en la muestra	OR <sub>muestral</sub>	OR <sub>muestral</sub> = 0,70
Obtención de la variancia muestral	IC del 95% <sub>muestral</sub> = LI a LS $\text{Variancia muestral} = \left( \frac{\ln \frac{LS}{LI}}{2z} \right)^2$	IC del 95% <sub>muestral</sub> = 0,57 a 0,87 $\text{Variancia} = \left( \frac{\ln \frac{0,87}{0,57}}{2 \times 1,96} \right)^2 = 0,0116$
Obtención de la IM	IM = 1/variancia muestral	IM = 1/0,0116 = 86
<b>C) Posterior</b>		
OR posterior	$\text{OR}_{\text{post}} = \text{EXP} \left( \frac{\text{IP} \times \ln(\text{OR}_{\text{priori}}) + \text{IM} \times \ln(\text{OR}_{\text{muestral}})}{\text{IP} + \text{IM}} \right)$	$\text{OR}_{\text{post}} = \text{EXP} \left( \frac{2,6 \times \ln(0,9) + 86 \times \ln(0,7)}{2,65 + 86} \right) = 0,71$
Obtención del error estándar posterior	EE <sub>post</sub> = 1/(IP+IM) <sup>0,5</sup>	EE <sub>post</sub> = raíz cuadrada [1/(2,6+86)] = 0,11
IC del 95% posterior	OR <sub>post</sub> × EXP ( ± 1,96 × EE <sub>post</sub> )	0,71 × EXP ( ± 1,96 × 0,11 ) = 0,57 a 0,87

EE: error estándar; EXP: exponencial (e elevado a la cantidad indicada); IC: intervalo de confianza; IM: información muestral; IP: información a priori; LI: límite inferior; LN: logaritmo neperiano; LS: límite superior; OR: odds ratio; z: valor de la distribución normal.

\* Se han buscado estos límites sólo a modo de ejemplo, pero teniendo en cuenta que la escala es multiplicativa, de modo que 0,9/0,27 = 3,0/0,9.

definir de modo formal las posibles H<sub>1</sub>. Todo esto simplifica las tareas de cálculo. Aceptar valores p próximos a 0,05 como «significativos» puede incrementar el riesgo de conclusiones espurias.

Habría que ser más escéptico frente a este umbral convencional, sobre todo cuando la creencia previa en la H<sub>1</sub> concreta que se está valorando sea muy baja. El valor p=0,05 sigue siendo mítico, como si constituyese una frontera claramente delimitada. En cambio, lo más sensato sería que a los metaanálisis y a los estudios que se pretende que sean tan convincentes como para cambiar la práctica clínica<sup>16</sup> se les exigiese un valor p mucho más bajo que el convencional límite de 0,05.

como toda una gama o distribución de posibles valores (distribución de probabilidad) por los que el investigador apuesta, ya que cree que entre ellos estará el verdadero valor.

Lo sorprendente es que esto debe fijarse antes de saber nada de los resultados del estudio. Así, en el ejemplo de actividad física y riesgo de enfermedad mental, se le podría haber preguntado al investigador antes de que viese los resultados mostrados en la tabla 1 cuál es la probabilidad que otorgaba subjetivamente a cada uno de los posibles valores de la OR para relacionar actividad física y enfermedad mental.

Supóngase que el investigador anticipa 2 predicciones subjetivas, antes de conocer lo que indican los datos reales (tabla 3). En la primera predicción, arriesga sobre cuál será el valor real de la OR para actividad física y enfermedad mental. Al valorar subjetivamente si será mayor de 0,9, el investigador otorga una probabilidad del 50% a que así sea. Lógicamente, eso implica que para él la probabilidad de que sea menor o igual a 0,9 es también del 50%. Esto equivaldría a afirmar que el investigador pronostica según sus creencias que lo más probable es que la actividad física

**Intervalos de confianza con métodos bayesianos simples y aproximados**

En la teoría bayesiana, se habla de un «priori» o «a priori» para un parámetro. Este prior no se entiende como un solo valor, sino

otorgue una protección débil frente al riesgo de enfermedad mental, ya que la reducción relativa del riesgo sería del 10%.

Luego hace una segunda predicción que ahora no es ya sobre el valor concreto que tendrá la OR, sino sobre cuál será su IC. Se pregunta, ¿cuál será la probabilidad de que un intervalo entre 0,27 y 3,0 incluya el valor real de la OR? Contesta que un 95% (tabla 3). El investigador afirma que él subjetivamente solo da a priori un 5% de probabilidad a que los valores de la OR estén por debajo de 0,27 o por encima de 3.

Una vez formalizadas estas 2 creencias o predicciones subjetivas, se puede proceder como muestra la tabla 3 para calcular un IC bayesiano de manera simple y aproximada. Este método es el recomendado por uno de los autores que más han contribuido a avanzar la metodología epidemiológica, Sander Greenland<sup>17</sup>. Al realizar los cálculos explicados en la tabla 3 se obtendrá un IC a posteriori que vale de 0,57 a 0,87; es decir, la información a priori prácticamente no ha modificado lo obtenido en el estudio. Esto es así porque la información a priori representa aquí un peso muy pequeño en relación con el peso que tiene la información obtenida por la muestra. Esto no es de extrañar. Se debe a la falta de creencias previas firmes y a la indefinición que de antemano tenía el investigador en este ejemplo sobre cuál sería el resultado esperable. Sus pronósticos previos eran muy poco atrevidos, pues creía a priori en un IC amplio y en una OR próxima a la unidad.

¿Qué pasaría si las creencias previas del investigador fuesen más precisas? Supóngase ahora que el investigador tiene una credibilidad previa del 50% de que la verdadera OR estará alrededor del valor nulo (1=ausencia de relación), y que en su 2.º pronóstico (sobre el IC) el investigador anticipase que la OR se va a mover en un margen muy estrecho, entre 0,8 y 1,25 (con el 95% de probabilidad). En estas circunstancias, el investigador da muy poco margen de credibilidad a que la verdadera OR sea menor de 0,8 o mayor de 1,25. Eso significa que su información previa debe tener mucho peso. Ahora, los pesos de la información previa y muestral serían similares (77 y 86, respectivamente, como se puede calcular usando la información de la tabla 3). Como los pesos ahora son muy parecidos, la OR posterior vendría a equivaler a un punto casi equidistante entre ambas estimaciones (la subjetiva a priori y la encontrada en la muestra). El resultado con estas nuevas creencias del investigador sería el siguiente:  $OR_{posterior}=0,83$ ; IC del 95%: 0,71-0,97. Sigue siendo una OR estadísticamente significativa al excluir la unidad, pero ya está más próxima al resultado nulo.

Se puede usar una simple hoja de cálculo programada en Excel para obtener estos cálculos con diversas suposiciones a priori. Se encuentra disponible una hoja de cálculo de este tipo en [www.unav.es/preventiva](http://www.unav.es/preventiva), dentro del apartado «Recursos bioestadística».

El procedimiento descrito en la tabla 3 se puede aplicar sin dificultad para otros parámetros (como el RR, la *hazard ratio*, o la diferencia de riesgos), incluso si se han obtenido por métodos multivariados, tras ajustar por factores de confusión, ya que lo único que se necesita por parte de los datos de la muestra son los límites del IC; los demás datos se derivan de hacer explícitas las creencias previas del investigador.

Se recomienda hacer un examen más detallado para ver cómo cambian los resultados según cambien las creencias a priori. Hacer esto con un amplio elenco de suposiciones o posibles valores a priori proporciona una magnífica idea de la robustez de los resultados. Por ejemplo, Naska et al<sup>18</sup> analizaron el riesgo de mortalidad coronaria asociado a dormir la siesta en una cohorte de 23.681 hombres y mujeres griegos seguidos durante una media de 6,3 años. El RR (*hazard ratio* de un modelo de Cox) tras el ajuste multivariado para quienes dormían la siesta fue de 0,66 (IC del 95%: 0,45-0,97) al compararlos con quienes nunca

Tabla 4

Ejemplo de análisis de sensibilidad: cambio del efecto de la siesta en función de que cambien las suposiciones o creencias a priori

Creencias a priori: RR (límites de probabilidad al 95%)	RR posterior (IC del 95%)
Ninguna (aproximación frecuentista)	0,66 (0,45-0,97)
1,00 (0,25-4,00)	0,68 (0,47-0,98)
1,00 (0,33-3,00)	0,69 (0,48-0,99)
1,00 (0,50-2,00)	0,73 (0,52-1,02)
1,50 (0,45-5,00)	0,71 (0,49-1,03)
1,50 (0,56-4,00)	0,74 (0,51-1,05)
1,50 (0,75-3,00)	0,80 (0,57-1,12)
2,00 (1,00-4,00)	0,86 (0,61-1,20)
2,00 (1,33-3,00)	1,11 (0,84-1,47)
2,00 (1,45-2,75)	1,27 (1,00-1,63)

IC: intervalo de confianza; RR: riesgo relativo.

dormían la siesta. La siesta, por tanto, resultó protectora. Se presentan en la tabla 4 diferentes estimaciones de cómo cambiarían las conclusiones si variasen las creencias a priori acerca del efecto de la siesta sobre la mortalidad coronaria. Se asumen 3 supuestos a priori: RR=1, RR=1,5 y RR=2. Una creencia de que la siesta no protege ni perjudica (RR=1) que tuviese amplios márgenes de posibles valores admisibles para el RR (0,25 a 4,00 o 0,33 a 3,00) no modificaría sustancialmente las estimaciones. Pero si las creencias a priori consistiesen en que la siesta duplica el riesgo de mortalidad coronaria (RR=2), los resultados del estudio griego no proporcionarían una prueba de que la siesta es protectora (tabla 4).

Una posibilidad interesante de los planteamientos bayesianos es que su aplicación implica ampliar el error estándar. Así, los IC bayesianos acaban siendo más amplios, como puede comprobarse en los ejemplos anteriores. Esta característica otorga una ventaja, ya que, de alguna manera, lleva a considerar el sesgo impredecible motivado por la falta de aleatorización o de comparabilidad exacta de los grupos expuesto y no expuesto. Solo cuando hay ya una gran cantidad de información a priori aportada por estudios con gran rigor metodológico (por ejemplo, los que procedan de estudios experimentales con reparto aleatorizado) se reducirán los márgenes de credibilidad a priori y los IC posteriores no se ampliarán tanto.

En el fondo, una vez que se comprende el pensamiento bayesiano, se entiende también que la aproximación frecuentista supone simplemente que se le está dando a toda la información previa un peso equivalente a 0. Pero esto no es muy sostenible. Debe admitirse que, antes de conocer los datos, suele existir siempre cierta información para mantener algún tipo de creencia previa o idea aproximada de por dónde van a ir los resultados que se encuentren. Por ejemplo, sólo muy rara vez el investigador esperará encontrar una  $OR > 10$  o una  $OR < 0,1$ . Esto marca unos límites lógicos y permite integrar conocimientos previos y poner en su contexto cada estudio particular. Como para tantos asuntos en la vida, no parece aceptable prescindir de todo contexto previo y creer que los únicos datos relevantes son los que se acaban de obtener.

Esta aproximación bayesiana sencilla y aproximada, que no requiere ningún programa especial de ordenador, ayuda a integrar la investigación propia con la ajena y a poder matizar más adecuadamente las conclusiones de un estudio.

#### Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

## Agradecimientos

Agradecemos a María Teresa Barrio (Cardiología, Clínica Universidad de Navarra) y a Jon Toledo (Neurociencias, CIMA, Universidad de Navarra) sus comentarios, que nos ayudaron a mejorar una versión previa de este manuscrito.

## Bibliografía

1. Argimón JM. El intervalo de confianza: algo más que un valor de significación estadística. *Med Clin (Barc)* 2002;118:382-4.
2. Rothman KJ. A show of confidence. *N Engl J Med* 1978;299:1362-3.
3. Gardner MJ, Altman DG. Confidence intervals rather than P values: Estimation rather than hypothesis testing. *BMJ* 1986;292:746-50.
4. Martínez-González MA, De Irala J. Intervalos de confianza y contraste de hipótesis. En: Martínez-González MA, Sánchez-Villegas A, Faulín Fajardo FJ, editores. *Bioestadística amigable*, 2 ed. Madrid: Díaz de Santos; 2006. p. 155-233.
5. Argimón JM. La ausencia de significación estadística en un ensayo clínico no significa equivalencia terapéutica. *Med Clin (Barc)* 2002;118:701-3.
6. Rothman KJ, Greenland S, Lash T. Precision and statistics in epidemiologic studies. En: Rothman KJ, Greenland S, Lash T, editores. *Modern Epidemiology*, 3 ed. Philadelphia, PA: Lippincott Williams and Wilkins; 2008. p. 148-67.
7. Fernández E. Estudios epidemiológicos (STROBE). *Med Clin (Barc)* 2005;125:43-48.
8. Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. for the CONSORT group. The revised CONSORT statement for reporting randomised trials: Explanation and elaboration. *Ann Intern Med* 2001;134:663-94.
9. Bayarri MJ, Cobo E. Una oportunidad para Bayes. *Med Clin (Barc)* 2002;119:252-253.
10. Sánchez-Villegas A, Ara I, Guillén-Grima F, Bes-Rastrollo M, Varo-Cenarruzabeitia JJ, Martínez-González MA. Physical activity, sedentary index and mental disorders in the SUN cohort study. *Med Sci Sports Exerc* 2008;40:827-834.
11. Martínez-González MA, De Irala J, Guillén-Grima F. Qué es una *odds ratio*? *Med Clin (Barc)* 1999;112:416-22.
12. Martínez-González MA, Alonso A, López-Fidalgo J. ¿Qué es una *hazard ratio*? (nociones de análisis de supervivencia). *Med Clin (Barc)* 2008;131:65-72.
13. Bland JM, Altman DG. Bayesians and frequentists. *BMJ* 1998;317:1151.
14. Goodman SN. Toward evidence-based medical statistics. 2: The Bayes factor. *Ann Intern Med* 1999;130:1005-13.
15. Ioannidis JPA. Effect of formal statistical significance on the credibility of observational associations. *Am J Epidemiol* 2008;168:374-83.
16. Katki HA. Evidence-based evaluation of p values and Bayes factors. *Am J Epidemiol* 2008;168:384-8.
17. Greenland S. Introduction to bayesian statistics. En: Rothman KJ, Greenland S, Lash T, editores. *Modern Epidemiology*, 3 ed. Philadelphia, PA: Lippincott Williams and Wilkins; 2008. p. 334-5.
18. Naska A, Oikonomou E, Trichopoulou A, Psaltopoulou T, Trichopoulos D. Siesta in healthy adults and coronary mortality in the general population. *Arch Intern Med* 2007;167:296-301.

