

## Sobre cómo analizar la credibilidad de un ensayo clínico o metaanálisis cuyo resultado principal se ofrezca en *odds ratio*, riesgo relativo o *hazard ratio*

Javier Escrig-Sos

Cirugía General y Digestiva. Hospital General de Castellón. Castellón. España.

### Resumen

Más allá del tipo de diseño y del análisis estadístico aplicado, la credibilidad de un trabajo de investigación radica en la compatibilidad de su resultado con la intensidad que aceptemos que pueda tener el fenómeno estudiado desde el punto de vista biológico. Esto requiere, en última instancia, un juicio de valor. En el presente artículo se describe un procedimiento para tratar de aproximarnos objetivamente a los límites de la intensidad que, de acuerdo con los datos que se manejen, debiera tener dicho fenómeno biológico para que, sobre esta base, nuestro juicio, derivado de los conocimientos disponibles sobre el problema, le pueda adjudicar la etiqueta de credibilidad. El procedimiento es válido cuando los resultados de la investigación se den en forma de *odds ratio*, riesgo relativo o *hazard ratio*, que a pesar de su difícil interpretación, son los estadísticos probablemente más utilizados en ensayos clínicos y metaanálisis, es decir, en los estudios con mayor nivel de evidencia en cuanto a su diseño metodológico.

**Palabras clave:** Credibilidad. Ensayo clínico. Metaanálisis. Intervalo crítico.

### ON HOW TO ANALYZE THE CREDIBILITY OF A CLINICAL TRIAL OR META-ANALYSIS WHOSE MAIN RESULT IS EXPRESSED IN ODDS RATIO, RELATIVE RISK OR HAZARD RATIO

Beyond the type of design or the statistical analysis applied, the credibility of a research study lies in the compatibility of its results with the intensity that the reader could accept that the phenomenon studied might have from a biological point of view. Ultimately this requires a value judgment. The present article describes a procedure that can be used to objectively approach the limits of intensity that that a biological phenomenon could have, according to the data presented, so that, based on the reader's judgment derived from the available knowledge of the problem, the study can be deemed credible. The procedure is valid when the results of the study are expressed in odds ratio, relative risk or hazard ratio. Although these statistics are difficult to interpret, they are probably the most widely used in clinical trials and meta-analyses, that is, in studies whose methodological designs provide the highest level of evidence.

**Key words:** Credibility. Clinical trial. Meta-analysis. Critical interval.

### Introducción

Estamos, sin duda, en un tiempo en el que la avalancha de información a la que tiene acceso el clínico es ingente y desproporcionada para lo que abarca ya no sólo su capacidad de análisis, sino también el tiempo disponi-

ble de lectura. El primer problema derivado de ello es poder separar el grano de la paja, y hacerlo con fundamento, tarea no siempre fácil, pues como glosó Smith<sup>1,2</sup>, exdotor del *British Medical Journal*, el clínico no es un héroe que deba dominar a fondo las herramientas estadísticas y metodológicas de la investigación para las que no ha sido específicamente preparado, igual que pueda dominar la materia que es objeto de su trabajo diario. El segundo problema es escoger el grano que sea de calidad, que pueda verdaderamente aplicarse en su desarrollo profesional, y que sea susceptible de producir un cambio positivo en éste. Para que esto ocurra es necesario que creamos como posible aquello que se nos dice en la bibliografía. Nadie puede negar que al acabar de leer un

Correspondencia: Dr. J. Escrig Sos.  
Cirugía General y Digestiva. Hospital General de Castellón.  
Avda. Benicasim, s/n. 12004 Castellón. España.  
Correo electrónico: escrig\_vicsos@gva.es

Manuscrito recibido el 6-4-2005 y aceptado el 16-6-2005.

trabajo de investigación, en última instancia, lo que se hace es precisamente un ejercicio de credibilidad.

La medicina basada en la evidencia, con sus postulados, junto con la expansión y la accesibilidad de ciertas herramientas estadísticas, han contribuido enormemente a allanar este camino, pero al final es el propio clínico quien decide si lo que lee es realmente creíble, es decir, si es aceptable según lo que él conoce acerca del fenómeno de que se trate, y según la idea que posee sobre hasta qué punto o con qué intensidad dicho fenómeno pueda ocurrir en la naturaleza. Se trata de un paso que sobresale de lo que es un análisis estadístico pertinente, una interpretación correcta de éste, y un oportuno diseño del estudio que se analiza, etapas éstas imprescindibles pero no suficientes por sí mismas para que un resultado sea creíble. La credibilidad se basa, pues, en un juicio cualitativo, pero lleva implícito el conocimiento aproximado o la intuición, en cuanto a un resultado, de unos límites de naturaleza cuantitativa que no se pueden traspasar. La lectura crítica de un artículo se limita muchas veces a la búsqueda de fallos groseros en éste, mientras que el aspecto de su credibilidad se suele dar por resuelto tan sólo según el valor de la *p* o según el propio diseño del trabajo –lo que es un abordaje preliminar del problema, aunque es también insuficiente–, o ni se repara en él, puesto que la mayor parte de las veces ni se le hace mención en la discusión. Es preciso tener, pues, grandes conocimientos metodológicos y del problema investigado, o de lo contrario recurrir a comentarios que aparecen tras la publicación del estudio en cuestión para conformar una idea más concreta de su credibilidad. De hecho, las revistas médicas deberían aumentar más de lo que lo hacen ya la proporción de artículos de revisión y divulgación<sup>2</sup>.

En cirugía, afortunadamente, en los últimos años se ha multiplicado la aparición de ensayos clínicos y metaanálisis, cuya escasez años atrás se nos echaba en cara por parte de otros colegas. Posiblemente la aparición de la cirugía laparoscópica haya sido el gran estímulo. En bastantes de ellos se desarrolla una parafernalia estadística ciertamente sofisticada; muchos presentan un diseño impecable, pero como en nuestro campo se manejan sobre todo variables cualitativas y predominan ciertas pruebas estadísticas, es frecuente expresar los resultados en forma de razón (*ratio*), medidas que son de interpretación bastante críptica y muy poco intuitiva o inmediata<sup>3</sup>. En este artículo se intentará dar unas orientaciones acerca de cómo interpretar de forma rápida, correcta y con pocas matemáticas, unas medidas de carácter cualitativo muy frecuentes en bibliografía médica, sobre todo en estudios de alto nivel, como ensayos clínicos y metaanálisis, como son las expresadas en forma de razones o ratios: *odds ratio* (OR), *riesgo relativo* (RR), o lo que es igual cuando se trata de estudios de supervivencia, *hazard ratio* (HR), para después mostrar un sistema relativamente sencillo de aproximarse a los límites cuantitativos que debe presentar un fenómeno, para que, sobre la base de lo expresado en el estudio que estamos leyendo, podamos juzgar si éste es creíble en cuanto a sus resultados, dando por sentado que el diseño de tal estudio sea correcto y las herramientas estadísticas usadas sean las apropiadas y muestren un valor de *p* suficientemente

pequeño que permita aceptar que el efecto que se comunica existe realmente y en la dirección en que se informa. Todo ello se apoyará sobre ejemplos de estudios, alguno muy conocido, de reciente aparición en el campo de la cirugía general.

### *Cómo se lee una OR y medidas similares*

Aunque en su fórmula matemática difieran, pues ésta debe corresponderse con las distintas situaciones en que se utilizan, la lectura de una OR, un RR y un HR es idéntica. Basta aplicar las palabras “veces más” o “por cada 1”. Si en la comparación en cuanto a curaciones de un novedoso tratamiento A frente a un antiguo tratamiento B se produce una OR de 1,3, significa que A cura 1,3 veces más que B, o en A se producen 1,3 curaciones por cada 1 curación que en B. Por tanto, A cura más que B. Siempre hay un grupo al que se le endosa la *ratio*, que es el grupo que se considera experimental (en este ejemplo es A), y otro grupo que sirve de referencia o de grupo control, al que se le endosa el 1, cifra que ni siquiera se cita en el resultado, que en este ejemplo es el grupo B. Si se tratara de un estudio de supervivencia, como en una regresión de Cox, el HR (algunos aquí también lo llaman RR) indicaría que el riesgo de eventos terminales que se podrían esperar en A durante el tiempo de seguimiento es 1,3 veces más alta que en B; por tanto, la supervivencia sería 1,3 veces peor en A que en B.

Si, a pesar de que no se exprese, el riesgo del grupo control se enrasa a 1, ello quiere decir que este 1 es el cero de la escala. Así, una OR, un RR o un HR de 1 significa que en A se produce el mismo efecto que en B, pues 1 evento en A por cada 1 en B supone un empate. Como ya habrán observado en la bibliografía, este tipo de estadístico se suele reportar, cuando se refiere al resultado principal del estudio, acompañado del llamado intervalo de confianza (IC), que es expresión de lo que la variabilidad debida al azar puede hacerlo oscilar en la población, o en la realidad. De esta forma, lo que en un estudio se calcula y se ofrece como un resultado puntual, no tiene por qué corresponderse fielmente con el verdadero valor en la población, valor que siempre será desconocido; pero el IC, a fin de cuentas y a pesar de que matemáticamente no sea exactamente así, expresa los límites que con una cierta seguridad contendrán ese verdadero valor. Generalmente los IC se calculan con una seguridad del 95%, lo que quiere constatar que, a la postre, sólo deja una probabilidad de alrededor de 0,05 de que el verdadero valor del OR no se halle en este intervalo. Esto es fácil de traducir aproximadamente a lo que es el valor *p*. Si el IC del 95% de una *ratio* no contiene el 1, el grado de evidencia de que los datos del estudio sean compatibles con la hipótesis nula –que es expresión de igualdad entre lo que se compara– ha de ser menor de 0,05, y viceversa. Los IC, de este modo, nos proporcionan la misma información cualitativa que el valor de *p*, pero además nos dan también una idea inicial de la intensidad posible del fenómeno estudiado a través de sus propios límites, siempre según los datos de la muestra empleada, idea que tiene que ver con lo que se podría llamar importancia práctica de un resultado, lo cual sin

Download English Version:

<https://daneshyari.com/en/article/9394613>

Download Persian Version:

<https://daneshyari.com/article/9394613>

[Daneshyari.com](https://daneshyari.com)