

Medidas del efecto de un tratamiento (II): *odds ratio* y número necesario para tratar

V. Abraira

Unidad de Bioestadística Clínica. Hospital Ramón y Cajal. Madrid.

En una nota previa¹ se comentaron algunos índices de asociación entre variables que también se usan en los ensayos clínicos para expresar el efecto de un tratamiento, cuando el resultado en cada paciente se representa mediante una variable binaria; al final de ella se anunciaba que había otros índices y a estos otros se va a dedicar esta segunda parte.

En primer lugar, el *odds ratio* (OR) que, a pesar de las frecuentes quejas^{2,3} sobre lo poco intuitivo que es y en consecuencia lo proclive que resulta a posibles malas interpretaciones, es uno de los índices más usados. Por ejemplo, en los originales publicados en esta revista en 2001 (hasta abril) es el único índice de asociación que aparece⁴. En la nota previa¹ se decía que cuando el resultado en cada paciente se representa mediante una variable binaria, es decir un episodio que puede ocurrir o no, el resultado para un grupo de pacientes es la proporción de pacientes en los que el episodio ocurre. Esta proporción estima la probabilidad, o riesgo, del episodio en ese grupo. Todos los índices revisados entonces se basan en comparar el riesgo en el grupo tratado con el riesgo en el grupo control.

Sin embargo, hay otra manera de representar la probabilidad, que proviene del mundo del juego. Por ejemplo, la probabilidad de que al tirar un dado salga un dos es 1/6: el dado tiene seis caras, todas con igual probabilidad de salir, y sólo una de ellas tiene el dos; obviamente la probabilidad de que no salga es 5/6, en consecuencia el dos saldrá en una proporción de $(1/6)/(5/6) = 1/5$ o, dicho de otra manera, la probabilidad de que salga un dos es 1/5 de la probabilidad de que no salga. Cuando los apostadores de las películas dicen, por ejemplo, que las apuestas a tal caballo están 1 a 5 se refieren a este cociente (quieren decir que una sexta parte de los jugadores apuestan por ese ca-

ballo mientras que las otras cinco sextas partes no apuestan por él), que en inglés se denomina *odds* y para el que no hay una traducción española comúnmente aceptada; aunque los apostadores en el juego de la pelota lo llaman momio, el diccionario de la RAE no recoge esta acepción que, además, es casi desconocida fuera de los frontones. Es decir, el *odds* es el cociente entre la probabilidad de que el episodio de interés ocurra y la probabilidad de que no ocurra y habitualmente se estima por el cociente entre el número de veces que ha ocurrido el acontecimiento y el número de veces que no ha ocurrido.

Volvamos al ejemplo de la nota anterior¹ cuyos resultados figuran en la tabla 1. A partir de la tabla, podemos estimar la probabilidad de que un paciente del grupo tratado muera por el cociente $20/200 = 0,10$ (10%) y el *odds* correspondiente por el cociente $20/180$. Del mismo modo, la probabilidad de que un paciente del grupo control muera se puede estimar por el cociente $15/100 = 0,15$ (15%) y el *odds* por $15/85$. Si queremos comparar ambos grupos (tratado y control), lo podemos hacer con el cociente entre ambas probabilidades, o riesgo relativo (RR) ($RR = 0,10/0,15 = 0,667$), o por el cociente entre ambos *odds*, u *odds ratio* (momio relativo, si aceptáramos la terminología del frontón):

$$OR = \frac{20/180}{15/85} = \frac{20 \times 85}{15 \times 180} = 0,63$$

Si no hay efecto del tratamiento, la probabilidad de que el acontecimiento ocurra es la misma en ambos grupos, el *odds* también es igual y en consecuencia el OR es 1. Cuanto mayor sea el efecto del tratamiento, más distintos son los *odds* y, por tanto, más se alejará el OR de 1 (en el sentido de mayor o menor que 1). En este sentido, se comporta del mismo modo que el RR. El RR, sin embargo, es más simple que el OR, ¿por qué, entonces, la popularidad de éste? Hay tres motivos para ello, en primer lugar el OR se puede estimar también en los estudios de casos y controles en los que no es posible estimar el RR, en segundo lugar, el OR permite, mediante la regresión logística⁵, ajustar la relación en estudio por el efecto de otras variables

Correspondencia: Dr. V. Abraira.
Unidad de Bioestadística Clínica.
Hospital Ramón y Cajal.
Ctra. Colmenar, km 9,100. 28034 Madrid.
Correo electrónico: victor.abraira@hrc.es

(ésta es la razón por la que Sánchez-Seco et al⁴ lo usan) y, en último lugar, en aquellas situaciones, muy frecuentes en epidemiología comunitaria pero bastante menos en epidemiología clínica, en las que el riesgo es muy pequeño en ambos grupos, el OR y el RR son muy parecidos.

El último índice que vamos a revisar es el *número necesario para tratar* (NNT), que es también el último índice propuesto⁶, con un uso creciente pero del que también aparecen quejas sobre su mal uso⁷. Representa el número de pacientes que deberían recibir el tratamiento experimental, en lugar del tratamiento control, para que un paciente adicional obtenga el beneficio (o, de forma equivalente, evite el perjuicio si el episodio en estudio es adverso). Hay que notar la importancia de la palabra “adicional” en la definición. El NNT se calcula como el inverso de la reducción absoluta del riesgo (RAR). En el ejemplo, el riesgo en el grupo control es 0,15; en el grupo tratado es 0,10; por tanto, la reducción absoluta del riesgo es 0,05, lo que quiere decir que por cada 100 pacientes que sigan ese tratamiento se evitarán 5 muertes con respecto a las que se habrían producido si se hubiera empleado el tratamiento del grupo control, por tanto, para evitar una muerte se necesitarán tratar 100/5 o equivalentemente $NNT = 1/0,05 = 20$. Obviamente, cuanto mayor sea el efecto del tratamiento menor será el NNT, un tratamiento con el que necesitamos tratar a 5 pacientes para salvar una vida es, en principio, mejor que otro con el que necesitemos tratar a 20. La principal ventaja de este índice es que expresa el efecto del tratamiento en unos términos que permiten fácilmente comparar sus ventajas con sus inconvenientes (efectos adversos, costes, etc.), de un modo, por tanto, muy cercano a la toma de decisiones clínicas. Hay que tener en cuenta que si el tratamiento no tuviera efecto $RAR = 0$ entonces NNT sería infinito, lo que daría lugar a algún problema⁸, sobre todo en el cálculo e interpretación de su intervalo de confianza.

Del mismo modo, para evaluar efectos adversos se puede definir el número necesario para perjudicar, representado por sus siglas en inglés NNH, aunque ya se ha señalado⁸ que esta terminología es confusa y que sería preferible usar el número necesario a tratar para producir un beneficio (NNTB) y número necesario a tratar para producir un perjuicio (NNTH).

Tabla 1. Resultados de un hipotético ensayo clínico

Grupo	Fallecidos		Total
	Si	No	
Tratado	20	180	200
Control	15	85	100
Total	35	265	300

Tabla 2. Resultados de otro hipotético ensayo clínico

Grupo	Fallecidos		Total
	Si	No	
Tratado	20	480	500
Control	30	470	500
Total	50	950	1.000

Puntos clave

- El *odds ratio* es uno de los índices más usados, a pesar de ser el menos intuitivo. Su principal ventaja es que permite, mediante la regresión logística, ajustar la relación en estudio por el efecto de otras variables.
- El *número necesario a tratar* expresa el efecto del tratamiento de un modo que permite fácilmente comparar sus ventajas con sus inconvenientes (efectos adversos, costes, etc.), muy cercano por lo tanto a la toma de decisiones clínicas.
- Existe evidencia empírica que sugiere que los índices relativos (*odds ratio* y *riesgo relativo*) tienden a ser más parecidos entre diferentes ensayos del mismo tratamiento que los índices absolutos (*reducción absoluta del riesgo* y *número necesario a tratar*).

En la primera columna de la tabla 3 figuran todos los índices calculados para los datos de la tabla 1. Se insiste en que todos ellos están calculados con los mismos datos, por tanto, todos reflejan el mismo efecto del tratamiento, aunque como unos están basados en diferencias y otros en cocientes, la relación entre ellos no se mantiene entre distintos estudios, es decir dos ensayos diferentes pueden dar lugar al mismo RR pero diferente NNT. Sirvan de ejemplo los resultados representados en la tabla 2, en los que los riesgos son menores en ambos grupos (0,06 en el control y 0,04 en el tratado), los índices calculados para estos datos figuran en la columna 2 de la tabla 3. Obsérvese que aunque los RR son iguales para los datos de ambas tablas, los demás son distintos, y los NNT, muy distintos.

Existe evidencia empírica proveniente de cientos de metaanálisis que sugiere que los índices relativos (OR y RR) tienden a ser más parecidos entre diferentes ensayos del mismo tratamiento que los índices absolutos (RAR y NNT)⁹ o, dicho de otra manera, los índices relativos serían característicos del tratamiento, mientras que los abso-

Tabla 3. Índices del efecto del tratamiento calculados para los datos de las tablas 1 y 2

	Tabla 1	Tabla 2
RAR	0,05	0,02
RRR	0,333	0,333
RR	0,667	0,667
OR	0,63	0,653
NNT	20	50

RAR: reducción absoluta del riesgo; RRR: reducción relativa del riesgo; RR: riesgo relativo; OR: *odds ratio*; NNT: número necesario para tratar.

lutos dependerían del tratamiento y también del riesgo basal. Para decidir si el resultado de un ensayo, expresado con un índice absoluto como el NNT, es aplicable en nuestro medio, hay que evaluar si el riesgo basal del ensayo es similar al de nuestro medio y, en caso contrario, adaptar el índice a nuestro riesgo basal. Del mismo modo, en un metaanálisis, el cálculo de un NNT global puede ser poco informativo⁷ si en los ensayos individuales hay heterogeneidad en los riesgos basales.

BIBLIOGRAFÍA

1. Abaira V. Medidas del efecto de un tratamiento (I): reducción absoluta del riesgo, reducción relativa del riesgo y riesgo relativo. SEMERGEN 2000; 26: 535-536.
2. Sackett DL, Deeks JJ, Altman DG. Down with odds ratios! Evidence-Based Med 1996; 1: 164-166.
3. Altman DG, Deeks JJ, Sackett DL. Odds ratios should be avoided when events are common [carta]. Br Med J 1998; 317: 1318.
4. Sánchez-Seco Higuera P, Delgado Delgado P, Retuerta Ameztoy P, López-Doriga de la Vega A, Viejo Martínez S, Díaz García RM. Prevalencia de factores de riesgo cardiovascular en la población anciana de la zona básica de salud de Horche (Guadalajara). SEMERGEN 2001; 27: 118-122.
5. Abaira V, Pérez de Vargas A. Métodos multivariantes en Bioestadística. Madrid: Centro de Estudios Ramón Areces, 1996.
6. Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. N Engl J Med 1988; 318: 1728-1733.
7. Smeeth L, Haines A, Ebrahim S. Numbers needed to treat derived from meta-analyses-sometimes informative, usually misleading. Br Med J 1999; 318: 1548-1551.
8. Altman DG. Confidence intervals for the number needed to treat. Br Med J 1998; 317: 1309-1312.
9. Deeks JJ, Altman DG, Dooley G, Sackett DL. Choosing an appropriate dichotomous effect measure for meta-analysis: empirical evidence of the appropriateness of the odds ratio and relative risk [resumen]. Control Clin Trials 1997; 18: S84-S85.