

EVIDENCIAS EN PEDIATRÍA

Toma de decisiones clínicas basadas en las mejores pruebas científicas

www.evidenciasenpediatria.es

Fundamentos de medicina basada en la evidencia

Ensayo clínico (IV). Resultados. Variables y medidas de efecto e impacto

Molina Arias M¹, Ochoa Sangrador C²

¹Servicio de Gastroenterología. Hospital Infantil Universitario La Paz. Madrid (España).

²Servicio de Pediatría. Hospital Virgen de la Concha. Zamora (España).

Correspondencia: Manuel Molina Arias, mma1961@gmail.com

Palabras clave en inglés: clinical trial; absolute risk reduction; number needed to treat.

Palabras clave en español: ensayo clínico; reducción absoluta del riesgo; número necesario a tratar.

Fecha de recepción: 21 de abril de 2015 • **Fecha de aceptación:** 29 de abril de 2015

Fecha de publicación del artículo: 6 de mayo de 2015

Evid Pediatr.2015;11:33.

CÓMO CITAR ESTE ARTÍCULO

Molina Arias M, Ochoa Sangrador C. Ensayo clínico (IV). Resultados. Variables y medidas de efecto e impacto. Evid Pediatr. 2015;11:33.

Para recibir Evidencias en Pediatría en su correo electrónico debe darse de alta en nuestro boletín de novedades en <http://www.evidenciasenpediatria.es>

Este artículo está disponible en: <http://www.evidenciasenpediatria.es/EnlaceArticulo?ref=2015;11:33>

©2005-15 • ISSN: 1885-7388

Ensayo clínico (IV). Resultados. Variables y medidas de efecto e impacto

Molina Arias M¹, Ochoa Sangrador C²

¹Servicio de Gastroenterología. Hospital Infantil Universitario La Paz. Madrid (España).

²Servicio de Pediatría. Hospital Virgen de la Concha. Zamora (España).

Correspondencia: Manuel Molina Arias, mma1961@gmail.com

Una vez analizados en entradas anteriores algunos aspectos claves del ensayo clínico, como pueden ser la aleatorización o el enmascaramiento¹, vamos a centrarnos en los tipos de resultados que nos ofrecen estos estudios. Describiremos en primer lugar los tipos de variables que podemos encontrar al analizar los resultados de un ensayo, centrándonos posteriormente en la valoración de las medidas de efecto e impacto.

Las medidas de asociación, que son aquellas que miden la fuerza de la asociación o riesgo entre dos variables, fueron descritas previamente al tratar los estudios observacionales^{2,3}, por lo que no se tratarán en este artículo.

VARIABLES DE RESULTADO

Una variable de resultado de un ensayo clínico es toda característica medida en los sujetos de estudio que nos permita diferenciar el efecto encontrado en los grupos comparados y plantear el contraste de hipótesis⁴. De manera habitual, la hipótesis nula de un ensayo clínico afirma que no existe diferencia de efecto entre las intervenciones comparadas con respecto a la variable de resultado elegida. Esta variable seguirá una distribución de probabilidad conocida que nos permitirá calcular su probabilidad de existir y que nos ayudará a decidir si rechazamos la hipótesis nula y damos como bueno el efecto detectado en la variable de resultado.

Las variables de resultado se clasifican en primarias y secundarias⁴. Las primarias son las que ayudan a contestar la pregunta o preguntas principales que el ensayo trata de dilucidar y son seleccionadas por su importancia clínica. Este tipo de variables condiciona el tamaño de la muestra del ensayo clínico, que deberá ser lo suficientemente grande para que el contraste de hipótesis tenga la potencia necesaria y la diferencia esperada resulte estadísticamente significativa.

Las variables secundarias responden a otras preguntas importantes del estudio. Por ejemplo, si ensayamos el efecto de un nuevo fármaco para el tratamiento de la cardiopatía isquémica, la variable principal puede ser la reducción de mortalidad, mientras que una variable secundaria podría ser la disminución de costes con el nuevo tratamiento. En ocasiones, la variable secundaria puede ser también muy importante y

habrá que tenerla en cuenta durante la fase de diseño con vistas a que el estudio tenga la potencia suficiente.

TIPOS DE VARIABLE

En cualquier ensayo clínico se recoge información variable acerca de las características de los sujetos que se aleatorizan, de los efectos de la intervención, de su eficacia, etc. Estos datos constituyen las variables del estudio, que podrán ser de diferente tipo en función de la naturaleza de los datos que las constituyen⁵ (**tabla 1**).

En primer lugar, podemos dividir las variables en cualitativas y cuantitativas. Las variables cualitativas son aquellas en las que cada valor corresponde a una característica o cualidad del sujeto al que pertenece. Como norma, los posibles valores que puede adoptar una variable cualitativa son excluyentes entre sí (hombre/mujer, vacunado/no vacunado, grupo de gravedad de una enfermedad a la que pertenece cada sujeto, etc.) y exhaustivos (cubren todos los posibles valores).

Las variables cualitativas pueden clasificarse en nominales y ordinales. Las nominales son aquellas en las que la escala de medida se compone de palabras que permiten únicamente clasificar a los sujetos, sin posibilidad de ordenar o jerarquizar sus valores. Un ejemplo sería la variable color de pelo, clasificada en tres categorías como rubio, moreno o pelirrojo. Las variables cualitativas nominales pueden ser dicotómicas o binarias cuando solo admiten dos categorías (hombre o mujer, vacunado o no, vivo o muerto, etc.) o politómicas, cuando admiten más de dos categorías (por ejemplo, grupo sanguíneo A, B, AB y 0).

Tabla 1. CLASIFICACIÓN DE LAS VARIABLES SEGÚN SU NATURALEZA

Cualitativas	Según el número de valores posibles: dicotómicas/politómicas Según la escala de medida: nominales/ordinales
Cuantitativas	Según el número de valores posibles: continuas/discretas Según la escala de medida: ordinales/de razón/de intervalo

Por su parte, las variables cualitativas ordinales sí permiten que sus valores estén jerarquizados. Por ejemplo, la variable "fumar" podría definirse como cualitativa nominal dicotómica (fuma/no fuma), con valores mutuamente excluyentes, o como cualitativa ordinal (no fuma/fuma poco/fumador moderado/muy fumador), con valores jerarquizados. Hay que tener en cuenta que, aunque es frecuente asignar a las diferentes opciones valores numéricos para su codificación (0: no fuma; 1: fuma poco; 2: fumador moderado; 3: muy fumador), estos no permiten realizar operaciones matemáticas, ya que la característica subyacente hace que carezca de sentido. Por ejemplo, el que fuma poco está menos expuesto al tabaco que el muy fumador, pero no puede decirse con rigor que esté expuesto la tercera parte.

Por último, las variables cuantitativas son aquellas en las que los datos se expresan realmente con números por lo que, además de operaciones de orden, permiten operaciones algebraicas. Este tipo de variables se dividen, a su vez, en discretas y continuas.

Una variable cuantitativa discreta es aquella que solo puede tomar cierto número de valores aislados. Los posibles valores serán finitos, generalmente representados por números enteros. Ejemplos de este tipo de variables pueden ser el número de hijos, el número de cigarrillos fumados al día, el número de episodios de una enfermedad, etc. Cuando las unidades de medida son variables (por ejemplo, la diferente longitud de cigarrillo consumido, la diferente intensidad o duración de episodio) estas variables suelen adoptar escalas de medida ordinales.

Por su parte, las variables cuantitativas continuas pueden tomar cualquier valor intermedio entre los números enteros. Idealmente, los posibles valores son infinitos, pudiendo añadir decimales sin fin. En la práctica, esto se ve limitado por la precisión de los aparatos de medida. Ejemplos típicos de este tipo de variables son el peso y la talla.

Algunos autores dividen las variables cuantitativas continuas en variables de intervalo y variables de razón, según el tipo de escala que utilicemos para medirlas. Las escalas de intervalo son aquellas que no proporcionan información sobre la magnitud absoluta del atributo que se mide. El ejemplo clásico sería el de la temperatura. Una escala en grados Celsius permite ordenar jerárquicamente las temperaturas de varios objetos e incluso realizar operaciones algebraicas con los diferentes valores, pero en sentido estricto no se puede afirmar que un objeto que está a 20 grados esté el doble de caliente que uno que esté a 10. Esto es así porque el cero de la escala no indica la ausencia de atributo.

Sin embargo, en las escalas de razón el cero sí indica la ausencia de atributo, por lo que la razón existente entre dos valores sí indica la relación entre los dos objetos medidos. Ejemplo de escala de razón sería la concentración de un tóxico en una muestra biológica. Existe el valor cero (0 g/ml) y muestras con 10 g/ml tienen el doble de tóxico que muestras con 5 g/ml.

MEDIDAS DE EFECTO O IMPACTO

Las medidas de impacto reflejan el efecto esperado por la intervención en estudio, ya sea como efecto beneficioso o como acción preventiva, en el caso de estar evaluando factores de riesgo. Las fundamentales son las fracciones atribuibles, las medidas de diferencia, las reducciones de riesgo (relativa y absoluta) y el número necesario a tratar (NNT). Las fracciones atribuibles fueron ya tratadas al hablar de los estudios observacionales², por lo que nos centraremos en las últimas referidas.

I. Medidas de diferencia

Como su propio nombre indica, estas medidas reflejan la diferencia existente en una misma magnitud de una variable entre los grupos comparados en un ensayo clínico.

La medida que se utilice dependerá del tipo de variable que estemos analizando. En el caso de variables cuantitativas, lo habitual es calcular la diferencia de medias o la diferencia de medianas, dependiendo de cómo se distribuya la variable de estudio. En el caso de variables cualitativas lo habitual será calcular la diferencia de proporciones entre los dos grupos. Un caso particular de diferencia sería la diferencia absoluta de riesgo o reducción absoluta de riesgo (RAR), que será tratada en el siguiente epígrafe.

En general, la diferencia calculada indica la contribución de una intervención en la producción de un efecto entre los expuestos a dicha intervención. Esto se basa en la presunción de que, de no existir la exposición, el riesgo de presentar el efecto sería igual que en los no expuestos.

La diferencia entre expuestos y no expuestos puede dar un resultado positivo, negativo o nulo. Un resultado positivo indica que la exposición favorece el efecto, uno negativo indicará que se comporta como factor protector y un valor nulo, ausencia de relación entre exposición y efecto. La cuantía del resultado dará una indicación del impacto de la exposición sobre la producción del efecto.

2. Reducciones de riesgo

Contemplaremos en este apartado la reducción relativa del riesgo (RRR) y la RAR.

Imaginemos un ensayo en el que el efecto esperado es la necesidad de cirugía. En los grupos intervención y control la proporción de sujetos que precisan cirugía fue, respectivamente, del 25% y del 50% (0,25 y 0,50). La RAR es la diferencia de riesgo entre los dos grupos: 25% ($0,50 - 0,25 = 0,25$), y se interpreta como que el riesgo en los tratados es un 25% menor que en los controles. La RRR es la diferencia de riesgos entre los dos grupos (RAR) respecto del control, y se estima dividiendo la RAR por el riesgo en el grupo control

($0,25/0,50 = 0,50$). En nuestro ensayo la RRR es 0,50 (50%) y se interpreta como que el riesgo en los tratados equivale al 50% del riesgo en los controles.

El problema de la RRR es que, aunque técnicamente correcta, tiende a magnificar el efecto y no cuantifica claramente el esfuerzo que debe hacerse para obtener un resultado determinado. Veámoslo con un ejemplo. Supongamos un ensayo en el que la mortalidad es del 10% en el grupo de intervención y del 20% en el control. La RRR será del 50%, mientras que la RAR será del 10%. Ahora supongamos otro estudio con 0,1% de mortalidad en el grupo de intervención y 0,2% en el control. El valor de la RRR será igualmente del 50%, mientras que la RAR será de 0,1%. ¿En cuál de los dos estudios es la intervención más eficaz? Claramente en el primero, con una RAR del 10%, que es la medida que mejor refleja el impacto de la exposición sobre el efecto. Habiendo tanta diferencia en la medida absoluta de riesgo entre los dos estudios, vemos como la medida relativa de reducción de riesgo es la misma en ambos. Por este motivo, siempre será preferible el uso de medidas absolutas de impacto frente a las relativas, especialmente si estamos estudiando hechos poco frecuentes, en los que pequeños cambios del riesgo basal absoluto pueden dar lugar a grandes cambios en los efectos relativos. Como veremos a continuación, es el efecto absoluto el que tendremos que considerar para estimar.

3. Número necesario a tratar

Es, probablemente, la medida de impacto que debe utilizarse de forma preferente al valorar los resultados de un ensayo clínico con variable de resultado cualitativa⁷.

La forma de calcularlo es habitualmente sencilla, basta realizar el inverso de la RAR ($1/RAR$), aunque existen métodos más complejos para su cálculo en otras situaciones⁸.

El NNT cuantifica el número de sujetos que debemos someter a la exposición o tratamiento para conseguir el efecto deseado. Siguiendo el ejemplo anterior, a una RAR del 10% (0,1) le corresponde un NNT de 10 ($1/0,10$), mientras que a una RAR de 0,1% (0,001) un NNT de 1000 ($1/0,001$). Vemos de nuevo como en el segundo estudio es necesario hacer un esfuerzo considerablemente superior para obtener el resultado buscado en un sujeto, a pesar de que la RRR del riesgo es la misma en ambos estudios (50%).

Un equivalente al NNT sería el número necesario para dañar (NND), que reflejaría el número de sujetos que deben exponerse para que se produzca un evento adverso. De manera similar, el número necesario para diagnosticar indica el número de pacientes que deben someterse a una prueba diagnóstica por cada caso que diagnosticuemos⁹.

Por último, debemos referir que es siempre deseable que los autores del ensayo proporcionen el NNT con su correspondiente intervalo de confianza, que nos dará una idea de la precisión de los resultados del estudio. Un intervalo estrecho nos indicará una gran precisión, mientras que un intervalo más amplio nos dará menor información por su menor precisión. Esta recomendación realizada para el NNT es también aplicable a las otras medidas de efecto aquí presentadas.

BIBLIOGRAFÍA

1. Molina Arias M, Ochoa Sangrador C. Ensayo clínico (III). Aleatorización. Enmascaramiento. *Evid Pediatr.* 2015;11:15.
2. Molina Arias M, Ochoa Sangrador C. Estudios observacionales (II). Estudios de cohortes. *Evid Pediatr.* 2014;10:14.
3. Molina Arias M, Ochoa Sangrador C. Estudios observacionales (III). Estudios de casos y controles. *Evid Pediatr.* 2014;10:33.
4. Bakhai A, Chhabra A, Wang D. Endpoints. En: Wang D, Bakhai A (eds.). *Clinical trials. A practical guide to design, analysis, and reporting.* Chicago: Remedica; 2006. p. 37-46.
5. Martínez-González MA, Gea A, Sayón-Orea C. Procedimientos descriptivos. En: Martínez-González MA, Sánchez-Villegas A, Toledo EA, Faulín J (eds.). *Bioestadística amigable.* 3.ª edición. Barcelona: Elsevier; 2014. p. 13-64.
6. Molina Arias M. Cálculo de la reducción del riesgo y del número de pacientes necesarios a tratar. *Rev Pediatr Aten Primaria.* 2012;14:369-72.
7. Aplicabilidad práctica de los resultados. En: Argimón JM, Jiménez J (eds.). *Métodos de investigación clínica y epidemiológica.* 3.ª edición. Madrid: Elsevier; 2004. p. 296-306.
8. Gómez-Acebo I, Dierssen-Sotos T, Llorca J. Número necesario de tratamientos: interpretación y estimación en análisis multivariados y con datos censurados. *Med Clin (Barc).* 2014;142:451-6.
9. Gerke O, Vach W. Number allowed to diagnose (letter). *Epidemiology.* 2014;25:158-9.