

Análisis Crítico de la Literatura

Medwave. Año XII, No. 3, Marzo-Abril 2012. Open Access, Creative Commons.

Criterios generales de validez de los estudios clínicos

Autor: Miguel Araujo Alonso⁽¹⁾

Filiación: ⁽¹⁾Hospital de Los Andes, Chile; Medwave, Santiago, Chile

Correspondencia: maraujo@medave.cl

doi: <http://dx.doi.org/10.5867/medwave.2012.03.5338>

Ficha del Artículo

Citación: Araujo M. Criterios generales de validez de los estudios clínicos. *Medwave* 2012 Mar/Abr;12(3) doi: 10.5867/medwave.2012.03.5338

Fecha de envío: 26/2/2012

Fecha de aceptación: 27/2/2012

Fecha de publicación: 1/3/2012

Origen: solicitado

Tipo de revisión: sin revisión por pares

Concepto de sesgo (error sistemático)

A grandes rasgos, existen dos variedades de error que pueden afectar a los resultados de un estudio clínico, los **errores por azar** y los **sistemáticos**. Estos errores producen resultados que se desvían de la verdadera asociación entre las variables. Por ejemplo, por efecto de un sesgo, el estudio puede mostrar:

1. Que el fármaco tiene un efecto beneficioso, cuando en realidad no es superior al placebo, o bien
2. Que cierto factor aumenta 5 veces la probabilidad de cáncer, cuando en realidad aumenta el riesgo sólo al doble.

A los **errores aleatorios**, producto del azar, nos referiremos más adelante en esta serie, al tratar sobre la significación estadística, el valor P y los intervalos de confianza.

Un **error sistemático** es un fenómeno que afecta el estudio en cualquiera de sus etapas, y que tiende a producir resultados que se desvían de los valores reales en un sentido determinado. De allí lo de "sistemático", porque el fenómeno actúa en un sentido predominante. A este tipo de error se le conoce como "sesgo".

En entregas posteriores analizaremos sesgos específicos de las principales variedades de estudios clínicos existentes. Aquí nos ocuparemos del concepto de manera general.

¿Cuándo se produce sesgo en un estudio?

Para que ocurra sesgo, debe producirse un fenómeno que afecte de manera diferenciada al grupo expuesto y al control, por ejemplo:

1. Que la susceptibilidad de desarrollar el desenlace sea mayor o menor en uno de los grupos, por razones distintas al factor estudiado (sesgo de selección).

Esta diferencia de riesgo puede estar representada por factores conocidos (edad, proporción de diabéticos, *score* de gravedad, etcétera), o desconocidos. En el caso de los factores que han sido medidos durante el estudio, su efecto pudiera ser ajustado -corregido mediante técnicas estadísticas- durante el análisis. Si no han sido medidos no será posible corregir los resultados en función de esas variables.

Muchas veces se puede sospechar la presencia de factores, no necesariamente medibles, que pueden estar sesgando los resultados. Por ejemplo, hay grupos de sujetos a los que se atribuye un mejor o peor pronóstico por el solo hecho de pertenecer a ese grupo, aunque sus restantes características sean similares:

- Los pacientes hospitalizados respecto de sujetos no hospitalizados
- Los sujetos de distinta condición socioeconómica
- Los sujetos que participan voluntariamente en un estudio, respecto de los que no acceden a participar, o
- Los trabajadores respecto de la población general (*healthy worker effect*).

2. Que el desenlace sea medido con distinta intensidad en unos y otros.

Por ejemplo, imagine que se está comparando dos grupos correspondientes a períodos distintos, por ejemplo un grupo atendido entre 2001 y 2005, y otro entre 2006 y 2010. Es posible que se introduzca un sesgo si los métodos o criterios de diagnóstico aplicados en ambos períodos no fueron enteramente equivalentes.

3. Que la exposición sea medida o expresada con distinta intensidad en unos y otros.

El ejemplo más evidente de este sesgo se da en los estudios de casos y controles, y está representado por el "sesgo de recuerdo" y sus diversas variantes. Nos referiremos a él en una serie próxima.

Sesgo y error de clasificación

Cuando medimos la exposición o los desenlaces de un paciente podemos equivocarnos al calificar su condición. Por ejemplo, si para identificar a los pacientes como hipertensos/no hipertensos usamos solamente un cuestionario, en el que preguntamos por el antecedente de hipertensión al propio sujeto, podríamos estar clasificando como no hipertensos a pacientes que sí lo son, pero no lo saben.

Cuando la exposición o el desenlace son medidos en forma errónea con similar intensidad en ambos grupos, es decir en forma aleatoria, se produce lo que denominamos **error de clasificación no diferencial**. Este tipo de error no modifica la magnitud de la asociación en términos relativos (que es la forma habitual de expresar los resultados), pero sí la altera en términos absolutos.

Siguiendo con el ejemplo inicial, imaginemos que el estudio evalúa la asociación entre el consumo de sal y la hipertensión:

- Los expuestos son sujetos con alto consumo de sal, de los cuales un 20% desarrolla hipertensión; y
- Los controles son sujetos con bajo consumo de sal, de los cuales un 10% desarrolla hipertensión.

El Riesgo Relativo (RR) del estudio es 2, es decir, los sujetos con alto consumo de sal tienen el doble de riesgo de desarrollar hipertensión. En términos absolutos, la diferencia de riesgo es de un 10%.

Por el error de clasificación mencionado, que afectó por igual a ambos grupos, hubo un 50% de subdiagnóstico, es decir, los hipertensos en el grupo expuesto son en realidad 30%, y en los no expuestos 15%. ¿Cómo se modifican los resultados? El RR sigue siendo el mismo, pero la diferencia en términos absolutos sube a 15%.

Por el contrario, cuando la clasificación inadecuada de la condición del paciente afecta de manera distinta a los grupos, hablamos de un **error de clasificación diferencial**. Este tipo de error da lugar a lo que normalmente conocemos como sesgo de medición.

Por cierto, en todo estudio es esperable algún grado de error de clasificación, que podremos minimizar mediante el uso de estrategias e instrumentos lo más válidos y acuciosos posibles para recoger los datos.

Mientras menos evidente el desenlace (imagine por ejemplo que el resultado a medir es la aparición de hepatocarcinoma), mayor esfuerzo debemos hacer para detectarlo. Mientras más subjetivo o menos acuerdo exista sobre la definición del desenlace, más importante será que el instrumento para evaluar el desenlace se encuentre estandarizado, los criterios se apliquen por igual a todos los pacientes, y si es un cuestionario o escala, que se encuentre validado en el país.

Cuando para verificar la exposición dependemos de la memoria del paciente, la información sobre exposiciones más recientes será por lo general más confiable. Existe

además la posibilidad de limitar el error de clasificación usando varias fuentes de información complementarias (por ejemplo, agregar a la entrevista que se hace al paciente a algunos de sus familiares directos, revisar los antecedentes disponibles en la ficha clínica o en bases de datos administrativas, etcétera).

De lo ya expuesto podemos inferir que existen dos grandes categorías de sesgo en los estudios clínicos:

- Sesgos de selección: se refieren a cualquier distorsión en la estimación del efecto, que resulta de la manera en que los sujetos son seleccionados para la población del estudio.
- Sesgos de medición o información: se refieren a cualquier distorsión en la estimación del efecto, que resulta de un error de medición o de clasificación de los sujetos en una o más variables.

Los sesgos pueden originarse por debilidades del diseño del estudio, o aparecer después durante la ejecución. Una vez producidos los sesgos son difícilmente corregibles, por lo que el esfuerzo por evitarlos debe ser llevado a cabo al diseñar el estudio, previniéndolos y tomando medidas apropiadas para minimizarlos.

Dirección de los sesgos

También se pueden clasificar los sesgos según la dirección de su efecto:

- Hay sesgos que producen una sobreestimación de la asociación entre las variables, también conocidos como sesgos positivos o sesgos "en contra de la hipótesis nula".
- Otros producen disminución en la magnitud de la asociación entre las variables, es decir, son sesgos negativos o a favor de la hipótesis nula (por ejemplo, determinan que el riesgo relativo del estudio sea más próximo a 1 que el real).
- Un sesgo extremo puede en teoría invertir la dirección de la asociación verdadera, llevando por ejemplo a que un factor de riesgo de enfermedad aparezca como protector (*switchover bias*).

Usted encontrará en la literatura epidemiológica una amplia variedad de sesgos descritos. Si bien es importante ir conociendo poco a poco las variedades principales o más frecuentes, no espere asimilarlos todos de una vez. Frente a eso, razone al leer un estudio en los siguientes términos:

¿Hay motivos para pensar que los pacientes son significativamente distintos entre los grupos, más allá de lo que pueda uno esperar por azar? Y si lo son... ¿fue suficiente el ajuste realizado en sus características basales durante el diseño -por ejemplo el pareo realizado en un estudio de casos y controles- o durante el análisis -por ejemplo el análisis multivariado realizado a la cohorte- para hacer comparables los grupos? (sesgo de selección).

¿Hay motivos para pensar que la exposición y resultados no fueron medidos de la misma forma entre los grupos? Si lo fueron... ¿se conocen todos los resultados o hay una

proporción demasiado grande de pacientes cuyos resultados no conocemos y que podrían ser distintas entre los grupos (pérdidas de seguimiento)? (potencial de sesgo de medición).



Esta obra de Medwave está bajo una licencia Creative Commons Atribución-NoComercial 3.0 Unported. Esta licencia permite el uso, distribución y reproducción del artículo en cualquier medio, siempre y cuando se otorgue el crédito correspondiente al autor del artículo y al medio en que se publica, en este caso, Medwave.