

Análisis Crítico de la Literatura

Medwave. Año XII, No. 1, Enero 2012. Open Access, Creative Commons.

Fundamentos del análisis crítico: concepto de validez y condiciones básicas para el análisis

Autor: Miguel Araujo Alonso⁽¹⁾

Filiación:

⁽¹⁾Hospital de Los Andes, Medwave

Correspondencia: maraujo@medwave.cl

doi: <http://dx.doi.org/10.5867/medwave.2012.01.5293>

Ficha del Artículo

Citación: Araujo M. Fundamentos del análisis crítico: concepto de validez y condiciones básicas para el análisis. Medwave 2012 Ene;12(1) doi: 10.5867/medwave.2012.01.5293

Fecha de envío: 26/12/2011

Fecha de aceptación: 26/12/2011

Fecha de publicación: 1/1/2012

Origen: solicitado

Tipo de revisión: sin revisión por pares

Resumen

El análisis crítico de la literatura es un proceso de evaluación que permite al lector formarse una idea del potencial de error en los resultados de un estudio, ya sea por sesgo o confusión. El análisis crítico trata en gran medida de verificar si el estudio cumple ciertos criterios o condiciones metodológicas deseables. Existen múltiples listas de chequeo que suelen utilizarse para guiar el análisis, pero completar una lista no es sinónimo de haber analizado críticamente un estudio. La validez interna de un estudio ha sido definida como la capacidad que tienen los hallazgos de la investigación de representar la verdadera relación entre una exposición y un desenlace, bajo las circunstancias particulares en que fue realizado. Para evaluar estos aspectos al analizar un estudio, se debe poner atención en los criterios de inclusión y exclusión utilizados, en los métodos de muestreo, en las características basales de los pacientes enrolados en el estudio, y en la descripción de las variables del estudio. La validez externa se refiere a las inferencias que podemos hacer más allá de la población muestral o aquella que concierne al estudio. Dentro de esta se distingue la validez poblacional y la validez ecológica. Por último, se discuten las amenazas a la validez externa que se deben considerar al evaluar un estudio.

Abstract

Critical analysis of literature is an assessment process that allows the reader to get an idea of potential error in the results of a study, errors arising either from bias or confusion. Critical analysis attempts to establish whether the study meets expected criteria or methodological conditions. There are many checklists available that are commonly used to guide this analysis, but filling out a checklist is not tantamount to critical appraisal. Internal validity is defined as the extent to which a research finding actually represents the true relationship between exposure and outcome, considering the unique conditions in which the study was carried out. Attention must be given to the inclusion and exclusion criteria that were used, on the sampling methods, on the baseline characteristics of the patients that were enrolled in the study. External validity refers to the possibility of generalizing conclusions beyond the study sample or the study population. External validity includes population validity and ecological validity. Lastly, the article covers potential threats to external validity that must be considered when analyzing a study.

El análisis crítico de la literatura es un proceso de evaluación que permite al lector formarse una idea del potencial de error en los resultados de un estudio, ya sea por sesgo -error sistemático- o confusión. El análisis crítico no entrega una sentencia definitiva sobre la condición de verdad de los resultados, pero nos aproxima

indirectamente a ella, señalando hasta qué punto pueden considerarse confiables.

Se suele considerar además dentro del análisis crítico la apreciación del potencial de error por azar, la correcta interpretación de los resultados que arroja el estudio, y la

aplicabilidad de los mismos a un determinado contexto de pacientes. No vamos a discutir si eso es pertinente o no. Todos esos componentes son necesarios al abordar un trabajo de investigación; sin embargo, es importante que la secuencia sea correcta.

Si un estudio tiene un alto potencial de error, lo que se haga con esos resultados también lo tendrá. Si el estudio sobre una intervención es de baja calidad pero su resultado es atractivo (ejemplo: disminución de la mortalidad en un 20%), podemos sentirnos tentados a aplicar la intervención haciendo caso omiso de la mala calidad del estudio. Antes de hacerlo, piense que podríamos someter a un grupo de pacientes a una intervención que en realidad es inefectiva, o cuyos efectos adversos superan los beneficios reales.

El mensaje es que, antes que nada, debemos analizar la validez interna del estudio. Si parece razonablemente confiable, podemos interpretar y utilizar sus resultados según corresponda. Si no parece confiable, debemos ser cautelosos en la aplicación de los resultados, o simplemente no tomarlos en cuenta.

Como todo en epidemiología, aquí tampoco las cosas son en blanco y negro. No hay una dicotomía entre estudios "válidos" y "no válidos". Existen estudios de gran calidad, que cumplen rigurosamente los más importantes criterios de validez interna. En otro extremo, hay artículos con limitaciones metodológicas serias que les restan todo valor científico. Y encontraremos también una amplia gama de situaciones intermedias.

¿Cómo enfrentarlas para decidir si un estudio es confiable o no? Utilizando el buen juicio, aplicando sus conocimientos epidemiológicos, y mirando al resto de la investigación existente en ese campo. Recuerde que para obtener conclusiones sobre la dirección y magnitud de las asociaciones en salud es necesario, además de evaluar los estudios en forma individual, que seamos capaces de analizar el conjunto de la evidencia disponible.

La tendencia inicial en quienes aprenden análisis crítico de la literatura es a descalificar en exceso los artículos que evalúan. Esta actitud debe ser moderada para que no conduzca al inmovilismo, en el que ningún trabajo es lo suficientemente bueno o confiable ("sesgo del epidemiólogo"). Veremos, además, que ciertos diseños poseen limitaciones intrínsecas en cuanto a su capacidad de controlar los sesgos.

El escenario global de la investigación clínica se encuentra plagado de estudios con deficiencias metodológicas, y hay que aprender a convivir con ello. Para algunos, esto ha llegado a constituir un "escándalo":

D G Altman. The scandal of poor medical research. BMJ, Jan 1994; 308: 283 - 284.

Y diez años después, se siguen planteando las mismas inquietudes, esta vez con énfasis en los estudios observacionales:

E. von Elm and M. Egger. The scandal of poor epidemiological research. BMJ, October 16, 2004; 329(7471): 868 - 869.

Como veremos más adelante, lo concreto es que es común encontrar deficiencias metodológicas en los estudios en todas las áreas de la medicina.

¿Qué condiciones básicas debe cumplir un buen análisis crítico?

Debe partir de una adecuada identificación del tipo de problema y del diseño utilizado por los autores.

Un error relativamente frecuente es analizar un estudio con criterios que no corresponden al tipo de pregunta o diseño en cuestión, por ejemplo, evaluar un estudio sobre pronóstico con los criterios de un ensayo clínico, o un estudio de pruebas diagnósticas con los criterios de un estudio sobre causa.

A veces se dan condiciones que facilitan este tipo de confusiones, por ejemplo, cuando se realiza un estudio de cohorte utilizando para ello pacientes que formaron parte de un ensayo clínico, ya sea para evaluar efectos de largo plazo o para realizar otro análisis secundario:

Zhang L, Krzentowski G, Albert A, Lefebvre PJ. Risk of developing retinopathy in Diabetes Control and Complications Trial type 1 diabetic patients with good or poor metabolic control. Diabetes Care. 2001 Jul;24(7):1275-9.

Que no sea realizado mecánicamente, sino que en forma razonada.

El análisis crítico trata en gran medida de verificar, a través de la lectura del artículo, si el estudio cumple ciertos criterios o condiciones metodológicas deseables. Existen múltiples listas de chequeo que suelen utilizarse para guiar el análisis, pero completar una lista no es sinónimo de haber analizado críticamente un estudio.

Que un estudio no cumpla determinado criterio puede tener una explicación perfectamente plausible. El revisor debe considerar hasta qué punto, dadas las características del estudio, le son aplicables cada uno de los criterios de análisis. Imagine un ensayo clínico en el que se compara una intervención quirúrgica con un tratamiento farmacológico: ¿será razonable, por ejemplo, pedirle a ese estudio que sea doble ciego? Que no lo sea, ¿le resta validez? Por cierto sería ideal que fuera ciego, pero probablemente no sea razonable exigirlo, dada la dificultad obvia para cegar la intervención. O bien si en el estudio el desenlace evaluado es la mortalidad, ¿influirá sobre ella que quien "mide" la mortalidad esté ciego a la intervención? Vea cómo se expresa esto en un ejemplo real:

Boutron I, Tubach F, Giraudeau B, Ravaud P. Methodological differences in clinical trials evaluating nonpharmacological and pharmacological treatments of hip and knee osteoarthritis. *JAMA*. 2003 Aug 27;290(8):1062-70.

Que vaya más allá de la verificación de los criterios antes señalados, sino que, además, incluya una apreciación sobre la dirección y magnitud de los potenciales sesgos del estudio, y de qué forma ellos pueden estar afectando la validez de los resultados.

Como veremos más adelante, ocasionalmente un sesgo puede inclusive reafirmar la existencia de una asociación.

Concepto de validez interna y externa

La **validez interna** de un estudio ha sido definida como la capacidad que tienen los hallazgos de la investigación, de representar la verdadera relación entre una exposición y un desenlace, bajo las circunstancias particulares en que fue realizado.

¿Qué podemos extraer de esta definición? En primer lugar, que la validez interna se refiere a la condición de verdad de los hallazgos, es decir, si los resultados del estudio efectivamente dan cuenta de lo que ocurre en la población en estudio o de la cual proviene la muestra. Recuerde que el hallazgo de una asociación en un estudio puede explicarse no sólo porque esa asociación efectivamente existe, sino también puede haber sido encontrada por azar, sesgo o confusión. De estos conceptos hablaremos más adelante.

En segundo término, que la validez interna se refiere solamente a las condiciones en las que se realizó el estudio, es decir, a esos pacientes, y a las variables de exposición y resultado tal y como fueron definidas en él, y no a cualquier población ni a otras formas de medir el desenlace o a otras modalidades de la misma intervención. Al evaluar la validez interna uno espera sacar conclusiones aplicables a cualquier sujeto que cumpla los criterios de inclusión y exclusión del estudio, o que pertenezca a la "población diana".

¿Cuál es esta última? Aquella población más amplia sobre la cual se desea, en definitiva, realizar las inferencias para los objetivos del estudio. Obviamente, el autor espera que sus hallazgos sean aplicables no sólo a pacientes idénticos a los del estudio, aquellos que cumplen estrictamente los criterios, sino a un universo más amplio.

Lamentablemente, no siempre la población del estudio es representativa de la población diana. A veces cuando se diseña un estudio se fijan criterios de inclusión que tienen como objetivo limitar sesgos o que se establecen simplemente por conveniencia práctica, y que afectan la representatividad de la muestra. Por ejemplo, se incluyen o excluyen pacientes en función de la probabilidad de que adhieran al tratamiento, o de que se pierdan durante el seguimiento. Imagine que desea evitar al máximo las pérdidas de seguimiento, en cuyo caso podría excluir pacientes que viven en zonas más alejadas o que no

poseen teléfono en sus hogares. O al revés, podría extraer los sujetos de una población de la que espera alto grado de cooperación, digamos, estudiar un grupo de médicos y enfermeras. De hecho, varios estudios han sido hechos en profesionales de salud:

Kurth T, Kase CS, Berger K, Schaeffner ES, Buring JE, Gaziano JM. Smoking and the risk of hemorrhagic stroke in men. *Stroke*. 2003 May;34(5):1151-5. Epub 2003 Mar 27.

En este caso, la muestra usó una población de médicos. ¿Significa eso que al autor le interesa el comportamiento de las variables sólo en esos profesionales? Por cierto que no, su interés es la asociación entre tabaquismo y accidente vascular hemorrágico en los hombres en general, cualquiera sea su oficio.

En otras ocasiones, los pacientes se seleccionan entre aquellos que tienen más probabilidad de presentar el evento de interés, con el propósito de disminuir el tamaño muestral necesario para el estudio (en general mientras mayor es la proporción de eventos esperados, menor el número de sujetos que es necesario reclutar). O simplemente por conveniencia práctica –acceso a la información de los pacientes–, se seleccionan a partir de un grupo que no representa todo el espectro de casos sobre el cual uno esperaría aplicar los hallazgos.

En otros casos, en cambio, los criterios son poco exigentes, precisamente para asegurar la afinidad entre la población muestral y la población diana.

Entonces, para evaluar estos aspectos al analizar un estudio, deberemos fijarnos en:

- Los criterios de inclusión y exclusión utilizados.
- Los métodos de muestreo. Recuerde que la forma de obtener una muestra realmente representativa de una determinada población es realizar un muestreo aleatorio -al azar- de la misma; sin embargo, la mayoría de las investigaciones clínicas obtienen los pacientes de manera consecutiva. A esto se suma, en particular en los ensayos clínicos, que por diversas razones, del total de pacientes elegibles siempre existe una fracción que finalmente no ingresa al estudio, y que puede hacer que la representatividad de la muestra sea -a lo menos- dudosa:

Britton A, McKee M, Black N, McPherson K, Sanderson C, Bain C. Threats to applicability of randomised trials: exclusions and selective participation. *J Health Serv Res Policy*. 1999 Apr;4(2):112-21.

Gross CP, Mallory R, Heiat A, Krumholz HM. Reporting the recruitment process in clinical trials: who are these patients and how did they get there? *Ann Intern Med*. 2002 Jul 2;137(1):10-6.

- Las características basales (la clásica Tabla 1 de cualquier artículo) de los pacientes enrolados en el estudio.

- La descripción de las variables del estudio: criterios de resultado, criterios de exposición (ejemplo, la forma en que se aplicó la intervención).

Nota importante: como ocurre frecuentemente en epidemiología, en esto también hay diferencias de opinión. Para muchos la validez interna se refiere exclusivamente a los sujetos de la muestra, y por lo tanto no debe ser extendida a la población "diana" a la que hemos hecho mención aquí, en especial cuando la muestra no es representativa.

Por otra parte, la **validez externa** se refiere entonces a las inferencias que podemos hacer más allá de la población muestral o aquella que concierne al estudio. Para determinar si los resultados del estudio son extrapolables o aplicables a una población distinta, debemos utilizar la racionalidad biológica y nuestros conocimientos teóricos sobre la enfermedad.

Algunos distinguen dentro de la validez externa:

- La validez poblacional: referida a las características de los individuos (por ejemplo, culturales, raciales, etarias, etc.).
- La validez ecológica: referida a las distintas condiciones de contexto o ambientales en las que se puede dar el fenómeno o la intervención.

En general, las amenazas a la validez externa de un estudio en las que uno debiera reparar incluyen:

- Que no exista una descripción explícita del tratamiento experimental.
- Que no exista descripción suficiente de las cointervenciones aplicadas a los pacientes.

- Efecto *Hawthorne* (la atención causa diferencias): los sujetos se comportan distinto cuando saben que están siendo estudiados u "observados". La asociación puede ser distinta en sujetos fuera del contexto de un estudio, que no se sienten o se saben investigados. Algunos utilizan este concepto un tanto indistintamente del efecto placebo.

Leung WC, Lam HS, Lam KW, To M, Lee CP. Unexpected reduction in the incidence of birth trauma and birth asphyxia related to instrumental deliveries during the study period: was this the Hawthorne effect? *BJOG*. 2003 Mar;110(3):319-22.

- Novedad: los sujetos pueden responder al tratamiento porque es novedoso, más que por el tratamiento mismo.
- Al revés, puede que no funcione precisamente porque es muy nuevo y los sujetos todavía no se adaptan a él (después de un tiempo de adaptación sí funciona).

En el caso de quien ejecuta una intervención...

- Efecto del experimentador: la intervención sólo funciona o lo hace mejor cuando la aplica el investigador. Este fenómeno se relaciona a su vez con la llamada "curva de aprendizaje": el tratamiento funciona mejor con quienes tienen experiencia en él; llevado a la práctica, con sujetos en distinto grado de entrenamiento, la efectividad disminuye.
- Interacción del período de tiempo y el efecto del tratamiento: con el tiempo, las condiciones bajo las cuales se aplicó la intervención ya no son las mismas.



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.