

Trabajo de Investigación

Medwave. Año XI, No. 9, Septiembre 2011. Open Access, Creative Commons.

Grupos de estudio pequeños: una búsqueda de evidencias

Autores: Patricio Herrera Labarca⁽¹⁾, Carlos Valenzuela Yuraidini⁽²⁾

Filiación:

⁽¹⁾Departamento de Pediatría, Facultad de Medicina, Universidad de Chile

⁽²⁾Instituto de Ciencias Biomédicas, Facultad de Medicina, Universidad de Chile

Correspondencia: pherrera@vtr.net

doi: 10.5867/medwave.2011.09.5133

Ficha del Artículo

Citación: Herrera P, Valenzuela C. Grupos de estudio pequeños: una búsqueda de evidencias. *Medwave* 2011 Sep;11(09). doi: 10.5867/medwave.2011.09.5133

Fecha de envío: 24/6/2011

Fecha de aceptación: 24/8/2011

Fecha de publicación: 1/9/2011

Origen: no solicitado, ingresado por FTS

Tipo de revisión: con revisión externa por 2 revisores a doble ciego

Correcciones

05/09/2011 En Resultados: A petición del autor, se corrige "Los resultados del análisis no pareado de la muestra original, aparecen en la Tabla I. En ella se aprecia un OR=0,247 (IC 95%=) y, en contraste, el OR=1,74 (IC 95%=0,79-1,71), "...por"Los resultados del análisis no pareado de la muestra original, aparecen en la Tabla 1. En ella se aprecia un OR = 0,79 (IC 95% = 0,47 a 1,36), cuyo nivel de significación estadística resultó p=0,35".

Resumen

Introducción: Con el fin de buscar evidencia empírica para respaldar la proposición de que los estudios sin hipótesis, sobre todo los de "pequeños grupos de pacientes" analizados sin hipótesis previa pueden dar lugar a resultados erróneos, llevamos a cabo un análisis siguiendo un modelo de fraccionamiento de una muestra mayor con el resultado de su análisis ya conocido.

Métodos: Se usó una muestra de 584 pacientes calculada según una hipótesis operacional cuyos resultados ya fueron publicados, segmentándola en seis subgrupos a los que se practicó el mismo procedimiento analítico y las mismas estadísticas.

Resultados: Los resultados mostraron que en la mitad de los grupos hubo asociación en una dirección (mayor que 1) y la otra mitad en sentido inverso (menor que 1). Sólo un grupo alcanzó nivel de significación estadística menor que 0,05.

Conclusión: Estos resultados son una ilustración basada en datos empíricos, de que estudios pequeños analizados sin hipótesis explícita previa producen resultados inconsistentes, no tienen el valor de los diseños analíticos completos -con hipótesis- y, por ello, su importancia eventual reside sólo en su calidad de descriptivos.

Abstract

Introduction: With the aim of seeking for empiric evidence as to support the hypothesis that small-size studies analyzed with no explicit operational hypothesis may lead to erroneous conclusions, we carried out an analysis of subgroups resulting from the breakdown of a previous published large study and using on them the same analytical procedures.

Methods: The 584 patients from a sample as calculated upon an operational hypothesis whose test results are already published, were grouped in six subgroups 100 individuals each, on whom the same analytical procedures were performed using the same statistical methods.

Results: The outcomes showed that half the groups had association statistics in one direction (above one) whereas the other half had it on the opposite (below one). Only one subgroup reached statistical significance ($p < 0.05$).

Comments: These results illustrate with empirical evidence that small-size study groups analyzed with no explicit hypothesis yield inconsistent results, this way supporting the hypothesis that these small-size groups yields do not have the meaning of complete analytical ones and hence, their eventual importance lies only on descriptive grounds.

Palabras clave: epidemiology, clinical research, analysis, validity, small study groups

Abreviaturas:

SHU: Síndrome Hemolítico-Urémico

OR: *Odds Ratio*. Término estadístico de traducción discutida, se ha traducido como oportunidad relativa, razón de ventaja o razón de azar.

Introducción y objetivos

Gran parte de la producción de la actividad en investigación clínica vaciada en la literatura –y, frecuentemente, presentada en congresos y otras reuniones- consisten en grupos pequeños de pacientes¹. En revistas del ámbito anglosajón en que se publica investigación clínica –referente máximo para el cuerpo médico en general- también abunda este tipo de artículos, bajo diversas denominaciones según el criterio de sus autores y aceptación por parte de los comités editoriales.

Sólo por mencionar un estudio hecho por los autores recientemente, en vías de publicación, citaremos como ejemplo los que tratan de la asociación entre infecciones por *Escherichia coli* O157:H7, uso de antibióticos y Síndrome Hemolítico-Urémico (SHU)^{2 3 4 5 6}. Ninguno de los estudios tenía una hipótesis operacional explícita sino, más bien, eran series de casos ulteriormente divididos en dos bloques dicotómicos (“con” y “sin” algún factor o efecto) según variables a elección de los autores (en este caso, ni siquiera sólo las necesarias para analizar la asociación, sino también otras), por lo que sus resultados no pueden considerarse producto de un modelo de diseño analítico.

Esta última afirmación ha sido objeto de debate académico no hecho público y la existencia de estos “modelos” de diseño de investigación no figuran en la literatura especializada. Con el fin de ilustrar la validez de la proposición de que **los grupos de estudio cuyo tamaño no es previsto en función de una hipótesis conceptual seguida de una operacional, tienen sólo calidad de descriptivos** y sus resultados deben ser sólo considerados como fuente de información para la investigación analítica, llevamos a cabo el estudio que pasamos a describir.

Métodos

Con el fin de saber si de un universo acotado por una hipótesis previa cuyos resultados conocemos⁸, la aplicación de exactamente los mismos métodos de análisis estadístico (tanto la medida de asociación, en este caso el OR) como los niveles de significación (“p”), aplicados a subconjuntos aleatorios del mismo, arrojará resultados que pudiesen diferir sustancialmente de lo que consideramos “la realidad” (los resultados obtenidos del universo mayor, que contiene los subconjuntos). Para ello se tomó una población constituida por 256 casos confirmados bacteriológicamente como “fiebre tifoidea” y

328 controles pareados por edad y sexo, reordenando los datos de modo de anular el pareo aleatorio –con lo cual los casos y los controles se reordenaron por su número correlativo, que es una nueva forma de hacer la muestra aleatoria para esta nueva experiencia. A continuación se segmentó la muestra “grande” original en 6 submuestras (“bloques”) de 50 casos y 50 controles (la última no fue así por la razón casos:controles de 1:1,3). Hecho esto utilizaron, con la misma hipótesis conceptual pero sin hipótesis operacional, la misma medida de asociación y los mismos procedimientos de significación estadística originales (excepto el análisis pareado). A los casos de fiebre tifoidea los denominamos “CASOS” y “CONTROLES” al resto de la muestra, constituida por niños sanos. Las categorías de exposición están representadas por “A” o fenotipo sanguíneo A y el resto de los fenotipos sanguíneos (B y O) están representados por “B”. La medida de asociación fue el OR y la de dispersión, los intervalos de confianza del 95% (IC 95%). La prueba de significación estadística fue la de Chi cuadrado y la de probabilidad exacta de Fisher cuando fue necesario; para el cálculo de OR de tablas que contuviesen un 0, se usó el método de Valenzuela⁹.

Resultados

Los resultados del análisis no pareado de la muestra original, aparecen en la Tabla I. En ella se aprecia un OR=0,79 (IC 95%=0,47 a 1,36), cuyo nivel de significación estadística resultó p=0,35.

Grupo	Caso	Control	OR	IC 95%	p< (*)
A	22	37			
B	234	291	0,79	0,47-1,36	0,35
Total	256	328			

(*) Cálculo según Chi cuadrado.

Tabla I. Análisis sobre la muestra no fragmentada.

Del examen de los resultados de los “bloques” (Tabla II), en tres de éstos el OR es menor que 1 lo que, en el contexto de la hipótesis conceptual del trabajo original “el fenotipo sanguíneo A II (o A+AB) está asociado a la probabilidad de sufrir de fiebre tifoidea”, resultaría ser protector para los pacientes que lo llevan en sus tejidos. En el resto de los “bloques” el OR es mayor que 1, lo que significaría que los pacientes con este fenotipo tendrían mayor probabilidad que el resto de la población en la muestra de contraer fiebre tifoidea, lo que los sindicaría como susceptibles a ella.

No. bloque	Grupo	Casos	Controles	OR	IC 95%	P (*)
1	A	4	13			
	B	46	37	0,247	0,07-0,82	0,03
2	A	5	7			
	B	45	43	0,68	0,20-2,32	0,76
3	A	3	2			
	B	47	48	1,53	0,24-9,59	1
4	A	5	3			
	B	45	47	1,74	0,39-1,71	0,71
5	A	5	3			
	B	45	46	1,7	0,38-7,55	1
6	A	0	9			
	B	6	70	0,71	0,14-3,61	0,54 (**)
Total		256	328			

(*) Cálculo según Fisher exacto. (**) OR e IC calculados según Valenzuela Y. C. "Dos soluciones para la estimación de 'odds ratios' con ceros". Rev.Med. Toshiba Chile 1993; 121:1441-1444.

Tabla II. Análisis sobre la muestra fragmentada en bloques.

Discusión y conclusiones

Este estudio ilustra cómo una determinada serie de casos sin hipótesis operacional –aunque hubiesen tenido una conceptual en los términos del trabajo original8- se aleja imprevisiblemente de la realidad. Si se hubiese elaborado la hipótesis operacional correspondiente, se habría llegado a un tamaño muestral parecido al estudio "grande". El no haberlo hecho muestra lo que puede suceder si, al margen de una hipótesis explícita operacional -que implica estipulaciones que el autor puede elegir, pero dentro de un margen con restricciones metodológicas-, se emprende la aplicación de pruebas de significación estadística que pueden o no ser apropiadas para el caso. En este ejemplo, el calcular riesgos relativos es erróneo (RR= 0,76 o 0,84, según se dispongan los datos en columnas y filas), pues el enunciado de la hipótesis conceptual y el diseño elegido (caso-control) determinan la hipótesis operacional y, de ella, la metodología para calcular el tamaño de la muestra para este objeto y no para otros. El que las cifras sean bastante parecidas es sólo un fenómeno incidental.

Cien pacientes es un grupo que puede considerarse intermedio entre los llamados grupos "chicos" y los mayores, en este caso con la particularidad de permitir examinar 50 pacientes con una característica de interés y 50 controles lo que, aparentemente, promete resultados interesantes. Sin embargo, el azar operó como para hacer los resultados de seis estudios independientes y en mucho comparables entre sí, contradictorios no sólo por no lograr siempre significación estadística sino –peor- por lograr esta apetecida meta sólo por azar, en uno de los bloques. Más aún, la asociación no apareció siempre en la misma dirección (las hubo mayores y menores que 1).

Algunas conclusiones anexas que pueden proponerse de este análisis son:

1. El trabajo original nos sugiere, con una probabilidad de 0,05 bilateral y poder de 0,80, que no existe asociación entre el fenotipo sanguíneo A II y la

susceptibilidad a la fiebre tifoidea. Ni este hecho ni el OR en cierto modo imprevisto, además de su falta de significación estadística pueden entenderse como un fracaso: el que no existiera tal asociación no figuraba en el paradigma de la fiebre tifoidea.

2. El tamaño muestral no puede determinarse sin una hipótesis conceptual además de una operacional que enfrentará al investigador con la necesidad de a) elegir un diseño, b) respetar explícitamente las exigencias metodológicas en el análisis de los resultados del diseño elegido y, c) sólo cumplidas estas etapas juzgar, en base a la observación de los cambios que significan distintos valores de la medida de asociación, nivel de error tipo I, poder y proporción de casos y controles en la muestra, cuál es la combinación que hace a su diseño factible y eficiente.
3. También es evidente que ni siquiera haciendo tal ejercicio se asegura: a) que los resultados afirmen la hipótesis, ni b) una cierta dirección de la presunta asociación. La dirección de ésta puede ser intuida por el autor, pero –en rigor y en la mayoría de los diseños- no puede tener certeza de ella.
4. Que, sin este cálculo, que sólo se justifica si existen las hipótesis mencionadas, se corre el riesgo de a) encontrar un $p \leq 0,05$ no tan sólo por simple azar, sino por ignorar condiciones de muestreo que no podrían pasar por alto al tener hipótesis y diseños de investigación y b) una direccionalidad espuria ya que, dependiendo del bloque de segmentación que elijamos, puede aparecer mayor o menor que 1, lo que lleva en cada caso a resultados abiertamente falsos.
5. Sin conocer las hipótesis conceptual y operacional, se pudo suponer que alguna de las tablas "pequeñas" (bloques) correspondía al seguimiento de dos cohortes prospectivas o retrospectivas, o a un ensayo controlado. Esto es posible si los autores, recopilado

un cierto número de pacientes y habiéndole echado una mirada a su distribución, plantean una hipótesis *a posteriori*, sin las indispensables estipulaciones.

No estará de más agregar que la recolección de datos de casos y controles se hizo después de haber elaborado el diseño para probar la hipótesis central, es decir, *a priori*. Esto no hace de este estudio (y, por consiguiente, de los datos de los "bloques") uno de tipo prospectivo sino que, por la naturaleza del problema y derivado de ello, el diseño elegido resulta ser necesariamente analítico, retrospectivo, controlado; lo que en propiedad debe llamarse "caso-control".

Más consideraciones válidas a partir de esta experiencia excederían el objetivo de su presentación.

Referencias

1. Araujo M. Estudios "pequeños": ¿cuál es la importancia del tamaño en los estudios clínicos? *Medwave* 2011;11(06). [↑ | CrossRef | PubMed |](#)
2. Bell BP, Griffin PM, Lozano P, Christie DL, Kobayashi JM, Tarr PI. Predictors of hemolytic uremic syndrome in children during a large outbreak of Escherichia coli O157:H7 infections. *Pediatrics*. 1997 Jul;100(1):E12. [↑ | PubMed |](#)
3. Ikeda K, Ida O, Kimoto K, Takatorige T, Nakanishi N, Tataru K. Effect of early fosfomicin treatment on prevention of hemolytic uremic syndrome accompanying Escherichia coli O157:H7 infection. *Clin Nephrol*. 1999 Dec;52(6):357-62. [↑ | PubMed |](#)
4. Slutsker L, Ries AA, Maloney K, Wells JG, Greene KD, Griffin PM. A nationwide case-control study of Escherichia coli O157:H7 infection in the United States. *J Infect Dis*. 1998 Apr;177(4):962-6. [↑ | PubMed |](#)
5. Proulx F, Turgeon JP, Delage G, Lafleur L, Chicoine L. Randomized, controlled trial of antibiotic therapy for Escherichia coli O157:H7 enteritis. *J Pediatr*. 1992 Aug;121(2):299-303. [↑ | CrossRef | PubMed |](#)
6. Cimolai N, Carter JE, Morrison BJ, Anderson JD. Risk factors for the progression of Escherichia coli O157:H7 enteritis to hemolytic-uremic syndrome. *J Pediatr*. 1990 Apr;116(4):589-92. [↑ | CrossRef | PubMed |](#)
7. Wong CS, Jelacic S, Habeeb RL, Watkins SL, Tarr PI. The risk of the hemolytic-uremic syndrome after antibiotic treatment of Escherichia coli O157:H7 infections. *N Engl J Med*. 2000 Jun 29;342(26):1930-6. [↑ | CrossRef | PubMed |](#)
8. Herrera P, Valenzuela CY, Arias H, Olivari F, Terán C, Ubilla C, Bravo P, Farías P, Oviedo I. Fiebre Tifoidea en el Niño: Asociaciones de los Fenotipos Sanguíneos ABO, Rh y MNSs. *Rev Med Chil*. 1992 Sep;120(9):986-93. [↑ | PubMed |](#)
9. Valenzuela C. Dos soluciones para la estimación de "odds ratios" con ceros. *Rev Med Chil*. 1993 Dec;121(12):1441-4. [↑ | PubMed |](#)



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.