

MBE en Cirugía Pediátrica. Lectura crítica de artículos. Etiología

V. Ibáñez Pradas¹, V. Modesto Alapont²

¹Servicio de Cirugía Pediátrica. Hospital General de Castellón. ²Sección de UCI Pediátrica. Servicio de Anestesiología. Hospital Infantil La Fe, Valencia.

Nota del editor

Este es el último de la serie de siete artículos de MBE publicados en nuestra revista desde Mayo de 2005 y escritos por los Dres. Ibáñez y Alapont. En ellos se han expuesto con gran calidad conceptos a veces complicados de asimilar, pero cada vez más importantes para la necesaria lectura crítica de la literatura. Queremos agradecerles el enorme esfuerzo dedicado a colaborar en la Formación Continuada de nuestra Sociedad aportando una guía de gran utilidad que convendrá tener siempre a mano.

*La culpa nunca cae al suelo,
siempre cae sobre alguien (o algo)*

INTRODUCCIÓN

Para los humanos, la forma más natural de entender el mundo que nos rodea es mediante las relaciones de causalidad. Quizá la pregunta más clásica de la investigación científica sobre los fenómenos de la naturaleza es la búsqueda de sus causas. También para los médicos, conocer la etiología de una enfermedad o los factores de riesgo de determinados procesos es una parte fundamental del conocimiento. De ello se derivan consecuencias para la práctica médica habitual, como atribuir un efecto nocivo a un agente que en realidad no lo tiene, o viceversa. En este artículo trataremos de desarrollar una guía de lectura crítica para los artículos sobre etiología, que también pueden encontrarse en los textos sobre MBE como “harm”, traducido como “daño” o más correctamente como “factores de riesgo”. Debe señalarse que con el término “factor de riesgo” se denominan aquellas variables que aumentan la probabilidad de que un sujeto vea modificada su estado de salud, bien positivamente (factor protector), bien negativamente (factor nocivo o productor de enfermedad). Como en los otros capítulos sobre lectura crítica seguiremos un esque-

ma de preguntas que evalúen la validez del estudio, la cuantificación de la magnitud de los resultados y su aplicabilidad en nuestro medio. Sobre este aspecto debe señalarse que mientras en textos clásicos de MBE puede encontrarse una plantilla “ad hoc” para la evaluación general de artículos de etiología o daño (Tabla I)⁽¹⁾, en otros medios como la red CAS-Pe sólo se ofrecen las guías para la valoración específica de estudios de cohortes o de caso-control. Nosotros utilizaremos la primera opción como guía general, aunque también entrando en las particularidades de los distintos tipos de estudio.

El mejor estudio para evaluar la relación causal es el ensayo clínico: de hecho se utiliza para saber si la causa de una curación es un cierto tratamiento. Pero para averiguar causalidad entre una exposición a determinado agente nocivo y una enfermedad, debería realizarse un ensayo clínico en el que los participantes fueran asignados de forma aleatoria a dos grupos (los expuestos y los no expuestos al factor de riesgo), y luego se siguieran para ver en cuál de los dos aparecían más casos de la enfermedad. Es evidente que este tipo de estudio es éticamente inaceptable, ya que –ni siquiera por mor de incrementar el conocimiento científico– el principio de “no maleficencia” nos impide someter a un grupo de personas a la acción de un supuesto agente nocivo. Por ello, generalmente los ensayos clínicos no constituyen una fuente de información clínica adecuada de causalidad para efectos nocivos: su tamaño muestral suele estar calibrado para la observación de la respuesta terapéutica del fármaco en cuestión, y no para el efecto adverso. De todas maneras, un efecto adverso detectado en un ensayo clínico sigue siendo evidencia de primer orden de causalidad.

Así que en el terreno de las causas, sólo nos queda apoyar nuestro conocimiento en los estudios observacionales. En deter-

Correspondencia: Dr. V. Ibáñez Pradas. Servicio de Cirugía Pediátrica. Hospital General de Castellón. Avda Benicàssim s/n, 12004 Castellón.
Email: ibanez_vic@hotmail.com

Recibido: Febrero 2006

Aceptado: Febrero 2007

Tabla I Preguntas para evaluar la validez de un artículo sobre un agente supuestamente nocivo

¿Son válidos los resultados del estudio?

1. ¿Existen grupos de pacientes claramente definidos, y son éstos similares en todo el resto de factores u otras causas excepto la exposición al factor de riesgo estudiado? (Control de la Confusión).
2. La exposición y los resultados clínicos ¿se han medido de manera idéntica en todos los grupos (p.e., la medición de los resultados fue objetiva [v.g. muerte] o “ciega” a la exposición)?
3. El seguimiento de los pacientes del estudio ¿fue completo y suficientemente prolongado como para que se produjera el resultado?
4. Los resultados del estudio ¿cumplen los supuestos de causalidad?:
 - Criterios Mayores:
 - a) Precedencia temporal (sine qua non): “La causa siempre precede a la consecuencia”
¿Queda claro que la exposición precedió a la aparición del resultado?
 - b) Plausibilidad biológica: ¿Existe un mecanismo biológico que dá sentido a la asociación?
 - c) Consistencia: ¿La asociación es coherente con los resultados encontrados en otros estudios diferentes?
 - d) Exclusión de factores de confusión (explicaciones alternativas):
¿Se ha hecho un análisis capaz de minimizar el efecto de los factores de confusión?
 - Criterios Menores
 - e) ¿Existe un gradiente dosis-respuesta?
 - f) Efecto del Cese de la exposición
¿Hay estudios que encuentran que la incidencia de la enfermedad disminuye al cesar la exposición?

¿Cuáles son los resultados?

- ¿Cuál es la magnitud de la asociación entre exposición y resultado? (riesgo relativo/odds ratio)
- ¿Cuál es la precisión de los resultados? (intervalo de Confianza)

Aplicabilidad de los resultados

- ¿Nuestro paciente es tan distinto de los sujetos incluidos en el estudio como para no poder aplicar sus resultados?
- ¿Cuál es el riesgo de nuestro paciente de beneficio-daño secundario al agente?
- ¿Cuáles son las preferencias de nuestro paciente?
- ¿Cuáles son los tratamientos alternativos?

minados casos se pueden realizar estudios de cohortes para estudiar prácticas de riesgo aceptadas por la sociedad (por ejemplo comparando fumadores y no fumadores). Pero son estudios epidemiológicos muy costosos (en recursos y en tiempo de seguimiento), y en muchas ocasiones, para no dar “palos de ciego”, los investigadores necesitan con urgencia hipótesis plausibles que guíen con eficiencia las investigaciones futuras. Para ello, en los años 70 y 80 se acabaron de desarrollar los fundamentos teóricos de los llamados estudios caso-control o caso-no caso.

Son, por definición, estudios retrospectivos, y en ellos el investigador recorre un camino que es el inverso al desarrollo temporal natural de los acontecimientos: tenemos los casos de enfermedad que ya se han producido, y viajamos al pasado para estudiar si los enfermos han estado más expuestos a determinados factores de riesgo que los no enfermos. Como luego comprenderá, la particular idiosincrasia de los estudios de caso-control, les hace especialmente indicados en los escenarios en los que la enfermedad es poco frecuente, una característica habitual en muchas de las que pertenecen al ámbito de la cirugía pediátrica. De todas formas, tanto en estudios de cohortes como de caso-control, ya que no se va a realizar aleatorización, para obtener la mejor evidencia causal de los diseños observacionales se deberá exigir que los autores realicen un adecuado análisis de los posibles factores de confusión que puedan influir en la relación causal que se está estudiando. Para ello es imprescindible que se haya realizado un análisis estadístico multivariable, generalmente mediante regresión lineal, logística, de Poisson o de Cox (según la variable de resultado que se esté utilizando).

Puesto que muchos de los artículos sobre factores de riesgo van a pertenecer a este grupo, nos detendremos un momento en el manejo de la principal medida de efecto utilizada en los estudios de casos y controles: la odds ratio (... vamos, no ponga esa cara. Ya hablamos de ella en el capítulo de pruebas diagnósticas)

ESTUDIOS CASO-CONTROL: LA ODDS RATIO

Su interpretación es la misma que la del riesgo relativo (RR): el valor OR = 1 representa el “no efecto”, el valor OR < 1 simboliza un factor protector, y OR > 1 indica un factor productor del evento que estudiamos.

En el capítulo sobre tratamiento presentamos el riesgo relativo (RR) como medida de efecto, utilizada en ensayos clínicos y estudios de cohortes. En este caso la variable de exposición (a un factor de riesgo, a un tratamiento) es la que se utiliza para clasificar los grupos de pacientes o cohortes: expuestos/No expuestos, Tratados/No tratados, etc. Pero si la incidencia de enfermedad fuera muy baja (p. ej., del 1%), si se utilizara un estudio de cohortes -aunque fueran retrospectivas-, para obtener 10 casos en el grupo de control ¡¡deberíamos reunir 1000 pacientes!!, ¡¡y otros 1000 pacientes en el otro grupo para poder establecer comparaciones fiables!! Y si la tasa de incidencia es muy baja, la velocidad a la que aparecen los casos podría hacer que hubiera de tardarse hasta 25 años en obtener datos fiables. Sólo por razones logísticas (sobre todo presupuestarias), muchas veces estos estudios serían completamente irrealizables. Para eso se han desarrollado los estudios con diseño de Casos-Controles. Se trata de un diseño retrospectivo en el que, por el contrario, es la variable de resultado la que se utiliza para montar los grupos de pacientes: Casos o Enfermos y No casos o Controles. El investigador recorre al revés la flecha del tiempo y busca en el pasado de cada uno de esos grupos la diferencia en exposición a los

Tabla II

Estudio de cohortes		Hepatitis		
		Sí	No	
Transfusión	Sí	300	5	305
	No	100	45	145
		400	50	450

distintos factores de riesgo en que estamos interesados. Pero actuar contra natura y viajar hacia atrás en el tiempo puede traer al investigador graves problemas, que ilustraremos mejor en el siguiente ejemplo ficticio.

Supongamos que, aunque nosotros no lo sepamos, el riesgo basal de hepatitis es 68,97% ($R_0 = 68,97\%$), pero que si el paciente es sometido a una transfusión este riesgo se eleva a 98,36% ($R_{exp} = 98,36\%$). Ello conduce a un Riesgo Relativo de hepatitis de 1,43 ($RR = R_{exp}/R_0 = 1,43$), por lo que la transfusión se comporta como factor de riesgo para esta enfermedad. Sea un investigador que pertenece a un grupo muy potente (recibe muchas subvenciones y dispone de una gran infraestructura y de horas remuneradas o becarios de investigación) y que está interesado en estudiar esa relación causal transfusión \rightarrow hepatitis. Para ello diseña un estudio de cohortes en el que sigue a 145 pacientes no expuestos y 305 sometidos a transfusión, y espera suficiente tiempo para que aparezca en todos la enfermedad. Una vez acabado el seguimiento del último paciente, con los resultados del estudio obtiene la siguiente tabla de 2x2 (Tabla II):

Fíjese que el investigador, con su diseño determina el tamaño muestral de las dos cohortes (expuestos y no expuestos), que están situadas en las filas de la tabla. Es el paso del tiempo el que irá rellenando las columnas y determinando el número de los enfermos o no enfermos. Con su cálculo de riesgos, usando los totales por filas –los tamaños muestrales de cada cohorte– como denominadores, al final del estudio el investigador obtendrá:

$$R_0 = 100/145 = 0,6897$$

$$R_{exp} = 300/305 = 0,9836$$

$$RR = R_{exp}/R_0 = 1,426$$

Seguir las normas del desarrollo temporal de los acontecimientos le ha permitido acceder al conocimiento real sobre la hepatitis y las transfusiones.

Pero supongamos ahora que es usted el que está interesado en la misma hipótesis causal: no dispone de ninguna subvención ni de tiempo remunerado para investigar (y mucho menos de un becario), y como mucho sólo puede convencer a un residente para que –en sus ratos libres– revise los datos retrospectivos de 90 historias del archivo del hospital. Recoge todos los datos de los enfermos de hepatitis, por cada caso elige uno o dos controles emparejados, y clasifica a los pacientes de su muestra utilizando la variable de resultado en enfermos (hepatitis Sí) y No enfermos. Seguidamente, investiga en el pasado de estos pacientes si se les transfundió o no. Con ello obtiene la siguiente tabla de 2x2 (Tabla III):

Tabla III

Estudio de caso-control		Hepatitis		
		Sí	No	
Transfusión	Sí	30	5	35
	No	10	45	55
		40	50	90

Tabla IV

		Efecto: enfermedad		
		Sí	No	
Causa	Sí	a	b	a+b
	No	c	d	c+d
		a+c	b+d	n

Como puede observarse, con su diseño de caso-control usted ha fijado el tamaño muestral de las columnas de la tabla: de los grupos determinados por la variable de resultado. Este artificio “contra natura” produce que ahora los totales por filas no correspondan a los denominadores de los riesgos: a los tamaños muestrales que deberían tener las cohortes de expuestos y no expuestos para que el curso del tiempo hiciera que aparecieran un total de 40 hepatitis y 50 sanos. Si pretende utilizar el cálculo de riesgos usando los totales por filas como denominadores, obtendría el siguiente resultado:

$$R_0 = 10/55 = 0,1818$$

$$R_{exp} = 30/35 = 0,8571$$

$$RR = R_{exp}/R_0 = 4,7143$$

Una estimación completamente errónea sobre el riesgo real de la transfusión.

Pero oiga ¡que uno es pobre, pero “honrao”!, y sobre todo listo... No tiene usted los medios del otro investigador, pero estudiando un poquito descubre con asombro que existe una alternativa con la que usted puede extraer conocimiento causal “real” a partir de estos estudios retrospectivos. Veámoslo: Si analiza una tabla 2x2 típica (Tabla IV):

Verá que lo que realmente nos interesa para establecer una relación causal es el RR de enfermar $= (a/a+b) / (c/c+d)$, un cociente entre dos probabilidades. Pero ¿y si el verdadero lenguaje con el que la naturaleza expresa la incertidumbre fuera el lenguaje de las “Odds”? En este caso, deberíamos expresar la relación causal calculando la Razón de Odds (en inglés Odds Ratio, OR) ¿No? Cuando (estudio de cohortes) la tabla tiene fijas las filas, la Odds de enfermar en los no expuestos (Causa No) es $O_0 = c/d$, y la Odds de enfermar en los expuestos (Causa Sí) es $O_{exp} = a/b$. Con ambos números podemos calcular el parámetro que nos interesa realmente: el OR de enfermar. Pero cuando (estudio de casos-contróles) la tabla tiene fijas las columnas, sólo puede calcularse la Odds de exposición en los no enfermos (Enfermedad No) $ONE = b/d$, y la Odds de exposición en los enfermos (Enfermedad Sí) OE

= a/c. A primera vista parecería que calcular la OR de exposición no tiene para nosotros ningún sentido, pero lo verdaderamente milagroso del caso es que si calculamos la OR, tanto si es la OR de enfermar como si es la OR de exposición, ¡ambos resultados coinciden! Así:

$$\text{OR exposición} = \text{O}_E/\text{O}_{NE} = (a/c)/(b/d) = a*d/b*c = (a/b)/(c/d) = \text{O}_{\text{exp}}/\text{O}_0 = \text{OR enfermar}$$

$$\text{OR exposición} = \text{OR enfermar} = \text{OR}$$

Es decir, si utilizamos un diseño caso-control, sólo seremos capaces de conocer la realidad sobre hepatitis y transfusiones si medimos la relación causal mediante la OR. Comprobémoslo: Los riesgos “reales” eran $R_0 = 68,97\%$, $R_{\text{exp}} = 98,36\%$, y $RR = 1,43$. Tanto si hubieramos hecho un estudio de cohortes ($OR = 45*300 / 5*100$), como un caso-control ($OR = 30*45 / 5*10$), el investigador hubiera obtenido el mismo resultado sobre el riesgo de hepatitis en las transfusiones: una OR de 27.

La OR puede utilizarse también para expresar relaciones causales obtenidas de estudios de cohortes, pero debemos reconocer que el lenguaje de las Odds –por lo menos a los latinos– nos resulta antipático. Así que solemos usar más el RR. El problema es que no existe manera de calcular un RR si sólo se dispone de una OR. De hecho no tienen porqué coincidir: $RR = (a/a+b) / (c/c+d)$, y $OR = a*d / b*c$. La única manera de aproximar el valor de un RR con el de una OR es asumiendo que la incidencia de la enfermedad es muy baja (el llamado “supuesto de enfermedad rara”), p. ej inferior al 5%. Bajo este supuesto, el número de enfermos es despreciable frente al tamaño muestral de ambas cohortes, y se cumple que $a+b \approx b$ y que $c+d \approx d$, con lo que: $RR = (a/a+b) / (c/c+d) \approx (a/b) / (c/d) = a*d / b*c = OR$.

El supuesto de enfermedad rara se suele cumplir cuando se estudian relaciones causales entre factores de riesgo y enfermedades, pero no cuando se estudia, p. ej., la tasa de respuesta a un tratamiento. Cuanto más frecuente es la tasa de respuesta, más inexacto es aproximar el RR con el OR. Zhang y Yu⁽²⁾ han demostrado que, conociendo el OR y el Rc (riesgo en la cohorte de control):

$$RR = \frac{OR}{(1-R_c) + (R_c * OR)}$$

Además, se puede calcular el NND a partir de una Odds Ratio (OR), sabiendo la Tasa de evento esperado para este efecto adverso (daño) de tus pacientes si NO estuvieran expuestos al factor de riesgo o, lo que es lo mismo, el riesgo de enfermar de los no expuestos (R_0). Para ello se aplica la siguiente fórmula:

$$NND = \frac{R_0(OR-1)+1}{R_0+(OR-1) \times (1-R_0)}$$

VALIDEZ DEL ESTUDIO

Supuesto práctico

Supongamos ahora que le avisan para valorar a un prematuro de 1100 g de peso y 10 días de vida que, en el contexto

Tabla V Preguntas para evaluar la validez de un estudio caso-control

¿Son válidos los resultados del estudio?

Preguntas de eliminación

- ¿Existe una pregunta clínica claramente definida?
- ¿El tipo de estudio es el adecuado para responder a la pregunta (el resultado a estudio es raro o perjudicial)?

Preguntas de detalle

- ¿Los casos se reclutaron correctamente? Evaluar el sesgo de selección:
 - Definición precisa de “caso”
 - Representan una población definida
 - Sistema fiable de selección
 - Tamaño muestral adecuado
- ¿Los controles se reclutaron correctamente? Evaluar el sesgo de selección:
 - Representativos de una población definida
 - Selección aleatorizada en base a una población
 - Número suficiente de controles
- ¿La exposición se midió de forma precisa para minimizar posibles sesgos?
 - Definición clara de la exposición. Medida precisa
 - Utilización preferente de variables objetivas
 - Métodos de medida similares en ambos grupos
 - Relación temporal correcta
- ¿Qué factores de confusión han tenido en cuenta los autores?
- ¿Se han tenido en cuenta los factores de confusión en el diseño y/o análisis?

¿Cuáles son los resultados?

- ¿Cuál ha sido la magnitud del efecto del tratamiento?
- ¿La OR se modifica sustancialmente con los ajustes por variables confundidoras?
- ¿Con qué precisión se ha estimado el efecto? Mire el valor de la p y de los intervalos de confianza
- Un efecto grande es difícil de ignorar: ¿Puede deberse al azar, sesgo o confusión?

Aplicabilidad de los resultados

- ¿Los resultados del estudio son aplicables en mi medio?
- ¿Los resultados coinciden con otra evidencia disponible? Considerar ensayos clínicos, revisiones sistemáticas, estudios de cohortes, y otros caso-control.

de una sepsis, presenta distensión abdominal y una radiografía con silencio abdominal. El paciente está conectado a ventilación mecánica, precisa de soporte con inotropos, y ha recibido tratamiento con indometacina en los 3 primeros días de vida por un ductus arterioso persistente. La exploración abdominal no le parece concluyente, pero la neonatóloga de guardia se inclina a pensar que el paciente presenta una perforación intestinal neonatal (PIN), y para justificar su sospecha le presenta un artículo publicado en el *Journal of Peri-*

Tabla VI Características de la muestra del estudio de Paquette et al.⁽³⁾

Población de origen

Pretérminos con peso al nacimiento < 1500 g nacidos entre el 1/1/97 y el 31/12/2003 en 2 centros hospitalarios de California.

Definición de caso (n=16)

Pacientes con diagnóstico de perforación intestinal neonatal (PIN) ocurrida en las 2 primeras semanas de vida, a partir de los registros de la base de datos de la UCIn.

Se define la PIN como una perforación de 5-10 mm en borde antimesentérico del ileon terminal, bien aislada, bien como perforaciones múltiples, en ausencia de isquemia intestinal secundaria con necrosis.

Criterios de exclusión:

- Enterocolitis necrosante
- Enfermedad de Hirschsprung
- Ileo meconial
- Anomalías intestinales
- Malrotación/vólvulo
- Perforación gástrica
- Perforación intestinal a partir de los 14 días de vida

Definición de control (n=32)

Dos controles por cada caso, emparejados en función del peso (± 100 g).

natology⁽³⁾ que estudia a pacientes con PIN en los primeros 14 días de vida, y que cifra el riesgo de perforación con una OR de 9,6. Vd le pide una copia del artículo y, dada la magnitud de la fuerza de asociación causal que ofrece, se toma su tiempo para realizar una lectura detenida y, si fuera preciso, una breve búsqueda a través de Medline.

Siguiendo con el esquema de preguntas presentado en la tabla V, en principio parece correcto plantear un estudio de casos y controles para estudiar la relación entre indometacina y perforación intestinal neonatal, ya que se trata de un proceso con baja incidencia (1-7% de los recién nacidos de muy bajo peso). Las características del estudio se muestran en la tabla VI. A priori parece correcto acudir a la base de datos del hospital para localizar a los pacientes con PIN. Tal vez quede la duda de los diagnósticos utilizados al principio del periodo de estudio, ya que la distinción entre PIN y enterocolitis necrosante comenzó a diferenciarse en la literatura a partir de 1997⁽⁴⁾, pero les concederemos el beneficio de la duda.

Selección de los casos

El tamaño muestral, 16 casos, parece una muestra escasa, pero muchas veces no puede reclutarse más muestra así que, en este caso, los autores han estudiado la muestra de que disponían. En cualquier estudio es recomendable hacer un cálculo aproximado del tamaño muestral necesario, que puede realizarse fácilmente a partir de programas de libre acceso como Epidat⁽⁵⁾. La paradoja es que para el cálculo del tamaño muestral siempre necesitaremos información que no tenemos, como la proporción de expuestos al factor de ries-

go (tanto en los casos como en los controles) o la OR que pretendemos detectar con el estudio. Será a partir de nuestro conocimiento sobre el tema y de la literatura previa como estimemos estos valores. Si, como los autores de este estudio, sólo reunimos 16 casos de PIN, pues no hace falta que calculemos más: eso es todo lo que tenemos.

Selección de los controles

Éste suele ser el talón de Aquiles de estos estudios, y su selección suele ocupar un lugar central en el diseño de estos estudios. Lo esencial es que los controles deberían ser una muestra representativa de la población de la que van a salir los casos. Aunque no nos detendremos aquí en estas particularidades, sólo para ilustrar lo complicado que puede llegar a ser su selección diremos que el periodo de tiempo durante el cual un sujeto puede ser elegido como control debería ser igual al periodo de tiempo durante el que podría convertirse en caso. Así, si en este estudio sólo se incluyen casos de PIN desarrollados en los primeros 14 días de vida (Tabla IV), para cada caso debería buscarse un control nacido como mucho 2 semanas antes. Y si tuviéramos 2 posibles controles, uno de 7 días de vida y otro de 14 días de vida, el segundo debería tener el doble de probabilidades que el primero de convertirse en control, ya que ha contribuido con el doble de tiempo que el primero. No se preocupe, ya que Vd no es el diseñador del estudio, y no es necesario entrar en tanto detalle para realizar una lectura crítica.

Se trata de un estudio “emparejado” (matched), lo que significa que se utilizan determinadas variables presentes en los casos para seleccionar a los controles (en este caso, el peso). La utilidad de esta maniobra es la mejora de la eficiencia estadística (mayor potencia estadística y mayor precisión de la estimación). De todas formas, en algún estudio se han seleccionado dos muestras de controles, con y sin emparejamiento, para comparar ambos resultados⁽⁶⁾. En caso de utilizar emparejamiento debe tenerse en cuenta que las variables utilizadas para realizarlo no podrán valorarse como factores de riesgo. Esto es evidente, ya que si utilizo el peso, como hicieron en el artículo que nos ocupa, la distribución de esta variable será la misma tanto en los casos como en los controles y, por lo tanto, su efecto será el mismo en ambos grupos y su OR será 1 (sin efecto). Así que, si sospechamos que una variable es la responsable de la enfermedad (en este caso, la exposición a indometacina), ésta no deberá utilizarse para la selección de los controles. Se puede emparejar por tantas variables como quiera, pero cuantas más utilice, más complicado será encontrar controles para todos los casos. Lo que sí está estudiado es la razón entre el número de casos y de controles: a coste similar y sin limitaciones de sujetos, la razón que minimiza el coste para una determinada potencia estadística es una razón 1:1. Si el número de casos es escaso (como en el estudio de Paquette⁽³⁾), se pueden incluir más casos para aumentar la potencia estadística (cosa que los autores no hicieron), pero ya que el aumento de potencia es mínimo cuando se incluyen más de 4 controles por caso, ésta suele ser la relación

Tabla VII Influencia de la razón caso:control. Ejemplo calculado para una proporción de casos expuestos del 40%, y de controles expuestos del 20%, para un nivel de confianza del 95%

A. Ilustra cómo aumentar el número de controles aumenta la potencia estadística para un número de casos dado. Para una muestra de 16 casos.

Razón caso:control	Potencia estadística (%)
1:1	22,9
1:2	32,2
1:3	37
1:4	39,3

B. Ilustra cómo, para una potencia estadística dada, seleccionar más controles reduce el número de casos necesarios. Para una potencia del estudio del 80%.

Razón caso:control	Casos necesarios
1:1	82
1:2	60
1:3	52
1:4	48

máxima: 4 controles por cada caso. En la tabla VII se muestra un ejemplo de la influencia de la razón entre casos y controles sobre la potencia estadística y el tamaño muestral.

Los autores no utilizaron selección aleatoria de los controles, una opción que no les hubiera costado demasiado, ya que los seleccionaban de una base de datos.

Medición de la exposición

Por lo que respecta a la exposición al factor de riesgo, será preferible la obtención de datos objetivos, ya que la subjetividad puede influir más en un grupo que en otro. Piense en un estudio que intente relacionar la toma de medicación durante el embarazo con la presencia de malformaciones congénitas: si la información se recoge mediante entrevista, la madre de un niño con malformación posiblemente recuerde con más detalle la medicación que tomó. En el artículo analizado los autores consultaron en las historias clínicas en qué momento se administró la droga y cuál fue la dosis, así que, en principio, la información en las historias debiera ser la misma tanto en un grupo como en otro, pero también sería deseable que la persona que recoge la información no conociera la hipótesis de estudio.

Sin embargo, en este estudio no parece que haya una definición clara de la exposición. No se trata de indometacina sí o no, sino que los autores han elaborado toda una serie de nuevas variables para medir la exposición (Tabla VIII). Como ya vimos en capítulo sobre pronóstico, cada investigador decide cómo codificar sus variables, y esto puede suponer un problema en muchas ocasiones. En este estudio, además, está el problema del limitado tamaño muestral, como veremos más adelante.

Tabla VIII Variables de control en el estudio de Paquette et al. (y sólo con 16 casos!!). Las variables marcadas con x fueron utilizadas en el análisis multivariante

Variable	
1. Edad gestacional	x
2. Peso al nacimiento	
3. Adecuación para la edad gestacional	
4. Administración de esteroides neonatales	
5. Condiciones maternas: hipertensión inducida, pre-eclampsia, polihidramnios, oligohidramnios, placenta previa, abrupcio placentae.	
6. Abuso de sustancias	
7. Gestación múltiple	
8. Sexo	
9. Apgar al minuto/ 5 minutos	x
10. Tipo de soporte ventilatorio	
11. Presencia de fuga aérea pulmonar	
12. Posición del catéter arterial (alto T6-T9; bajo L3-L4)	x
13. Deposiciones antes de la administración de indometacina o esteroides y antes de la perforación	x
14. Uso de contraste oral, enemas (salino o contraste) y/o supositorio de glicerina	
15. Dosis máxima de vasopresor/inotropo, y edad a la administración	x
16. Administración de inotropos?15 mcg/kg/m durante > 72 horas	
17. Momento y dosis administrada de indometacina	x
18. Momento y dosis administrada de dexametasona	x
19. Alimentación antes de la perforación (si/no) y tipo	
20. Antibioterapia antes de la perforación	
21. Hiper magnesemia (niveles?2,2 mg/dl)	
22. Cirugía (si/no) y momento de realización	
23. Drenaje peritoneal y momento de colocación	
24. Ductus arterioso persistente (DAP) en ecocardiografía	
25. Ligadura quirúrgica del DAP	
26. Positividad en cultivo de sangre/orina/tubo endotraqueal	
27. Niveles de cortisol antes de la administración de corticosteroides	
28. Medicación gastrointestinal (metoclopramida, eritromicina, antiH ₂)	

Control por factores de confusión

El último paso para valorar la validez es comprobar si se ha controlado por factores de confusión. Un factor de confusión se define como aquel factor relacionado tanto con la exposición (p.ej. indometacina) como con el resultado (p. ej. PIN) pero que no constituye un elemento intermedio en la relación causal (Fig. 1). No se preocupe, no es necesario andar elucubrando si determinado factor es un eslabón intermedio o no. Es más sencillo ver qué factores han controlado los investigadores y valorar, desde su conocimiento del tema, si considera que falta alguno relevante.

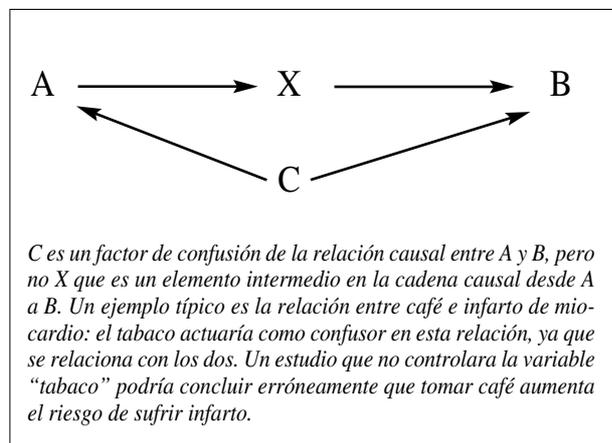


Figura 1. Factor de confusión. Esquema.

En la tabla IX se ofrecen todas las variables que recogieron los autores del estudio. A pesar del apabullante número de variables, observe que sólo 7 se incluyeron en el análisis multivariante. Aquí debemos hacer un comentario sobre la relación entre el tamaño muestral y el número de variables analizadas: realizar un análisis multivariante no significa carta blanca para analizar todas las posibles combinaciones que se nos ocurran, y si nuestra muestra es pequeña, el número de variables a introducir en el multivariante necesariamente debe ser limitado. No está bien establecido el número de casos necesarios para estimar una regresión logística (el tipo de análisis multivariante aplicable a un estudio caso-control), pero una regla muy extendida obtenida a partir de estudios de simulación⁽⁷⁾ es la de incluir 10 casos del evento en estudio (en este caso, la perforación intestinal) por cada variable a controlar. Eso significa que con 16 casos de PIN podríamos incluir... ¡¡1,5 variables!!.

Ya ve, no importa que los autores hayan controlado el efecto de más de 20 variables. Ellos saben (y nosotros también) que el efecto de todas esas posibles variables ha de ser valorado conjuntamente, no por separado; y que con 16 casos no pueden hacer grandes cosas. Eso no significa que su experiencia e información aportada no sean útiles, pero controlar por 7 variables con sólo 16 casos puede llevar a conclusiones no demasiado fiables. Un problema más: cuantas más variables incluya en relación a una muestra pequeña, mayor es la probabilidad de encontrar relaciones simplemente por azar.

¿CUÁLES SON LOS RESULTADOS?

Lamentablemente, el estudio de Paquette⁽³⁾ no ofrece las cifras que necesitamos para calcular directamente la OR. Nos ofrece directamente el resultado del análisis multivariante, en el que "curiosamente" sólo aparece una variable como estadísticamente significativa: la administración conjunta de ≥ 3 dosis de indometacina y dexametasona, con una OR de 9,6

Tabla IX Codificación de la exposición a indometacina y dexametasona

1. Primer día de administración de dexametasona
2. Primer día de administración de indometacina
3. N° de dosis de indometacina antes de la dexametasona
4. N° de dosis de dexametasona antes de la indometacina
5. Dosis totales de indometacina > 3
6. 0-3 dosis de dexametasona
7. 4-6 dosis de dexametasona
8. Horas entre dosis (entre indometacina y dexametasona y viceversa)
9. Dosis total de indometacina
10. Dosis total de dexametasona
11. Exposición a ≥ 3 dosis de tanto de indometacina como de dexametasona

(IC 95% 1,22-75,71). Una de las cosas en las que debe fijarse es si el ajuste por variables de confusión modifica sustancialmente el resultado, y en este caso parece que así ocurre, porque el resultado del análisis univariante ofrecido por los autores muestra una OR de 2,2 (IC 95% 1,2-4). ¿Se debe este cambio a una variable de interacción que aumenta el efecto de la exposición a la indometacina y la dexametasona, o este cambio es el resultado de una relación espuria provocada al incluir en el análisis multivariante variables confundidoras?

No podemos saber si ese cambio tan enorme (pasamos de 2 perforaciones en los expuestos a multiplicar esta cifra por 4) se debe o no al efecto de la confusión, pero lo que sí sabemos es que este resultado es muuuuuuy impreciso: mire el intervalo de confianza. Puede que la verdadera OR sea 1,2 (un cambio muy discreto) o puede que sea 75,7 (un cambio espectacular), y compruebe también que los autores han omitido el valor de la p. Un dato más con respecto a los intervalos de confianza: que el resultado haya sido 9,6 no hace más probable este resultado que cualquiera de los otros valores del intervalo. Es decir, tan probable es que la OR sea 1,2, como 9,6, como 75. Seguro que en otros ámbitos de la vida no aceptaría un resultado así (imagine que se tratara del presupuesto de una reforma en su casa).

Como regla general, recuerde que un efecto grande es difícil de ignorar, pero en el caso que nos ocupa, el pequeño tamaño muestral, la codificación de las variables y la escasez de datos proporcionados sobre el análisis multivariante plantean dudas acerca de los resultados ofrecidos.

APLICABILIDAD DE LOS RESULTADOS

Ahora debemos valorar si nuestro paciente es tan diferente de los del estudio analizado hasta el punto de no poder aplicar sus resultados. El tipo de paciente es similar, pero existe una diferencia fundamental: nuestro paciente no ha recibido tratamiento con dexametasona, y en el estudio analizado, el

tratamiento aislado con indometacina no se relacionó con un aumento de riesgo para PIN, así que, según el artículo elegido, no habría más riesgo de PIN en un paciente tratado con indometacina.

El siguiente paso, si el estudio analizado le plantea dudas, es buscar información en otras fuentes, y eso es exactamente lo que Vd hizo: Una breve búsqueda por Medline le hace descubrir que en el mismo número de la revista *Journal of Perinatology*⁽⁸⁾ hay otro estudio caso-control, con resultados menos espectaculares, pero más precisos, puesto que la muestra de estudio estuvo formada por 633 casos y 581 controles. Siguiendo una metodología similar y tras realizar un análisis multivariante, este nuevo estudio cuantifica la relación entre indometacina y PIN con una OR de 1,86 (IC 95% 1,4-2,5).

Y llegando a este punto, Vd, que ya le ha cogido el gusto a esto de obtener información por su cuenta, se da cuenta que a partir de una OR y del riesgo en el grupo de no expuestos (R_0) puede obtener una estimación del NND (recuerde la fórmula presentada en la introducción). ¿De dónde obtener el riesgo de PIN en el grupo no expuesto? Pues de una cohorte que se haya seguido prospectivamente. En la introducción comentábamos que un efecto adverso detectado en un ensayo clínico es una buena fuente de evidencia, y ahí es exactamente donde acudimos: Dos ensayos clínicos^(9,10) realizados sobre pacientes con peso inferior a 1500 g mostraron un riesgo de PIN en la cohorte no expuesta del 5%. Así que con la OR de 1,86 y un R_0 de 0,05 es fácil obtener el NND: por cada 28 pacientes < 1500 g tratados con indometacina habría 1 caso de PIN. Visto de esta forma, la indometacina no parece un factor tan claramente asociado a la perforación intestinal, ¿no le parece?

BIBLIOGRAFÍA

1. Straus SE, Richardson WS, Glasziou P, Haynes RB. *Medicina basada en la evidencia. Cómo practicar y enseñar la MBE*. 3ª Ed. Churchill-Livingstone, Madrid 2006.
2. Zhang J, Yu KF. What's de Relative Risk? A method of correcting the Odds Ratio in Cohort Studies of Common Outcomes. *JAMA* 1998;**280**:1690-1.
3. Paquette L, Friedlich P, Ramanathan R, Seri I. Concurrent use of indomethacin and dexamethasone increases the risk of spontaneous intestinal perforation in very low birth weight neonates. *J Perinatol* 2006;**26**:486-92.
4. Moss L, Dimmitt RA, Henry MCW, Geraghty N, Efron B. A meta-analysis of peritoneal drainage versus laparotomy for perforated necrotizing enterocolitis. *J Pediatr Surg* 2001;**36**:1210-3.
5. <http://dxsp.sergas.es/default.asp>.
6. Glaser N, Barnett P, McCaslin I, Nelson D, Trainor J, Louie J, et al. Risk factors for cerebral edema in children with diabetic ketoacidosis. *N Engl J Med* 2001;**344**:264-69.
7. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol* 1996;**49**:1373-9.
8. Attridge J, Clark R, Walter MW, Gordon PV. New insights into spontaneous intestinal perforation using a national data set: (1) SIP is associated with early indomethacin exposure. *J Perinatol* 2006;**26**:93-9.
9. Schmidt B, Davis P, Moddemann D, Ohlsson A, Roberts RS, Saigal S, et al. Long-term effects of indomethacin prophylaxis in extremely-low-birth-weight infants. *N Engl J Med* 2001;**344**:1966-72.
10. Watterberg KL, Gerdes JS, Cole CH, Aucott SW, Thilo EH, Mammel MC, et al. Prophylaxis of early adrenal insufficiency to prevent bronchopulmonary dysplasia: A multicenter trial. *Pediatrics* 2004;**114**:1649-57.