

EVIDENCIAS EN PEDIATRÍA

Toma de decisiones clínicas basadas en las mejores pruebas científicas
www.evidenciasenpediatria.es

Fundamentos de medicina basada en la evidencia

Métodos de ajuste de sesgos. Análisis estratificado

Ochoa Sangrador C¹, Molina Arias M², Ortega Páez E³

¹Servicio de Pediatría. Hospital Virgen de la Concha. Complejo Asistencial de Zamora. Zamora. España.

²Servicio de Gastroenterología. Hospital Universitario La Paz. Madrid. España.

³UGC Góngora. Distrito Granada-Metropolitano. Granada. España.

Correspondencia: Manuel Molina Arias, mma1961@gmail.com

Palabras clave en español: análisis estratificado; estadística; inferencia estadística; métodos de ajuste.

Palabras clave en inglés: stratified analysis; statistics; statistical inference; adjustment methods.

Fecha de recepción: 5 de septiembre de 2022 • **Fecha de aceptación:** 21 de septiembre de 2022

Fecha de publicación del artículo: 28 de septiembre de 2022

Evid Pediatr. 2022;18:31.

CÓMO CITAR ESTE ARTÍCULO

Ochoa Sangrador C, Molina Arias M, Ortega Páez E. Métodos de ajuste de sesgos. Análisis estratificado. Evid Pediatr. 2022;18:31.

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=2022;18:31>.

©2005-22 • ISSN: 1885-7388

Métodos de ajuste de sesgos. Análisis estratificado

Ochoa Sangrador C¹, Molina Arias M², Ortega Páez E³

¹Servicio de Pediatría. Hospital Virgen de la Concha. Complejo Asistencial de Zamora. Zamora. España.

²Servicio de Gastroenterología. Hospital Universitario La Paz. Madrid. España.

³UGC Góngora. Distrito Granada-Metropolitano. Granada. España.

Correspondencia: Manuel Molina Arias, mma1961@gmail.com

En documentos previos de Fundamentos se abordaron los errores epidemiológicos. Entre ellos se mencionaron los errores sistemáticos o sesgos: sesgos de selección, de información y de análisis. Se adelantaba en su momento que los dos primeros solo podían ser controlados en la fase de diseño del estudio, mientras que los sesgos de análisis, confusión e interacción podían ser controlados con diferentes procedimientos en la fase de análisis.

Comenzaremos recordando brevemente lo que denominamos confusión e interacción.

Se produce **confusión** cuando la asociación entre la exposición en estudio y el efecto puede ser explicada por una tercera variable, la variable de confusión. La presencia de tal factor de confusión altera los resultados y puede producir un aumento o disminución del efecto, o incluso cambiar la dirección del mismo.

Para que un factor sea de confusión necesita cumplir unos requisitos:

- I) Debe ser factor de riesgo independiente de la enfermedad, tanto para los expuestos como para los no expuestos en la población.
- II) Debe estar asociado al factor de exposición que estudiamos en la población de donde provienen los casos.
- III) No debe ser una consecuencia de dicha exposición; es decir, no debe ser un paso intermedio en la cadena o secuencia de causalidad entre la exposición y la enfermedad.

Veamos un ejemplo. En un estudio en el que queremos estimar el beneficio de la lactancia materna para evitar enfermedades infecciosas, tendremos que controlar la clase social de la familia, ya que podría comportarse como factor de confusión. La clase social cumpliría los requisitos de un factor de confusión, ya que se asocia a la enfermedad (riesgo de infección), al influir en los cuidados higiénicos; asimismo, se asocia a la exposición (tipo de lactancia) y no forma parte de la cadena causal entre exposición y enfermedad (el tipo de lactancia no produce cambios de clase social que puedan afectar a la incidencia de infección).

Existe **interacción** o modificación de efecto cuando la asociación entre exposición y efecto varía según las categorías de una

tercera variable, la variable modificadora de efecto. Si no analizamos la modificación del efecto para los diferentes valores de esa tercera variable, la estimación de la asociación será incompleta.

Veamos un ejemplo. En un estudio queremos estimar el riesgo de la toma durante la gestación de inhibidores de la dihidrofolato reductasa (ej.: cotrimoxazol) sobre el riesgo de malformaciones congénitas en el feto. En este estudio tendremos que controlar si las gestantes tomaban o no vitaminas y, en concreto, ácido fólico durante la gestación, ya que el efecto podría cambiar en función de esta tercera variable. De hecho, la toma de los inhibidores de la dihidrofolato reductasa solo aumenta el riesgo en los que no toman ácido fólico. La toma de ácido fólico se comporta como un factor de interacción o modificación del efecto. Cuando hay interacción, el análisis no puede limitarse a estimar el riesgo en el conjunto de gestantes, sino que debe hacerse por separado, en las que lo tomaban y en las que no.

Los principales **métodos de ajuste** empleados para controlar sesgos son el análisis estratificado y el análisis multivariante. En este capítulo nos centraremos en el análisis estratificado, que nos permitirá describir de una forma sistemática los fenómenos de confusión e interacción y nos proporcionará procedimientos para su ajuste.

ANÁLISIS ESTRATIFICADO

La estratificación es una técnica estadística que permite detectar y/o controlar sesgos de análisis, tanto la confusión como la interacción. Cuando exploramos la asociación entre dos variables, estimando, por ejemplo, medidas de riesgo, como el riesgo relativo o la *odds ratio*, si separamos la muestra en subgrupos de una tercera variable, obtendremos estimaciones de asociación dentro de cada subgrupo. Si las estimaciones por subgrupos son diferentes de las calculadas para toda la muestra, puede existir algún sesgo de análisis a corregir. Habitualmente se consideran diferentes cambios de un 10% en las estimaciones. Cuando esto ocurre la estimación agrupada no será válida.

Veamos cómo podemos usar el análisis estratificado para controlar factores de confusión e interacción.

Confusión

Si sospechamos que existe un factor de confusión en nuestro análisis, la forma más sencilla de comprobarlo es hacer un análisis estratificado, haciendo un análisis por separado de los sujetos con o sin dicho factor de confusión. Si la asociación entre exposición y enfermedad disminuye (en ocasiones puede aumentar) en cada estrato (por ejemplo, riesgos relativos u *odds ratios* que se acercan a 1) con respecto a la existente en el análisis global, existe un factor de confusión. Si hacemos una ponderación de las estimaciones de asociación de cada estrato obtendremos la estimación verdadera de la asociación entre exposición y enfermedad. Aunque no es habitual, la confusión también podría producirse en el sentido contrario, con un aumento del efecto en cada estrato y, por lo tanto, en la estimación ajustada global.

Veamos un ejemplo. En un estudio de cohortes se siguieron a 482 niños asmáticos, midiendo la incidencia acumulada de crisis asmática en un periodo en función de la exposición o no a tabaco en el hogar. Presentaron crisis un 62,5% (60/96) de los expuestos a tabaco en el hogar, frente a un 36,2% (140/386) de los no expuestos. Ese mayor riesgo correspondía a un riesgo relativo (RR) de 1,72 (intervalo de confianza del 95% [IC 95]: 1,40 a 2,11), que supone un aumento estadísticamente significativo del 72%. Para poder estar seguros de que esa asociación no es la consecuencia de la existencia de algún factor de confusión se valoraron otras posibles variables. En la **Tabla 1** se presentan las frecuencias de crisis en función de la exposición a tabaco, estratificando por exposición alta o baja a contaminación ambiental.

Lo primero que llama la atención es que la diferencia de riesgos entre los expuestos a tabaco y los no expuestos ha descendido considerablemente en ambos estratos, aunque sigue siendo mayor en los expuestos. Como consecuencia, en ninguno de los dos estratos la diferencia es estadísticamente significativa (en ambos el valor nulo "1" está incluido en el intervalo): en el estrato de alta contaminación el RR del tabaco es 1,14 (IC 95%: 0,91 a 1,43), en el estrato de baja contaminación el RR es 1,30 (IC 95: 0,78 a 2,19). Por ello, si quere-

mos estimar el riesgo real tendremos que estimar el riesgo ajustado, haciendo el cálculo ponderado a partir de las estimaciones de ambos estratos (1,14 y 1,30). En la **Figura 1** se presentan las fórmulas para el cálculo manual de los riesgos ajustados, tanto para RR como para *odds ratios* (OR); más adelante explicaremos algunas herramientas informáticas que facilitan su cálculo, con las salidas de resultados para este y otros supuestos. El RR ajustado es 1,18 (IC 95: 0,95 a 1,45). En este ejemplo, la contaminación ambiental se comporta como factor de confusión en la relación entre tabaco y asma. Cuando el efecto observado en el análisis crudo (sin ajustar) se modifica en ambos estratos en la misma dirección, habitualmente atenuándose, respecto al valor nulo (en este caso el 1), tenemos un factor de confusión.

Interacción o modificación del efecto

La existencia de interacción en la relación entre dos variables implica que la asociación entre estas variables es diferente cuando se hace un análisis por subgrupos de una tercera variable. Esto es, existe heterogeneidad entre subgrupos en la asociación o efecto observados.

La interacción puede observarse en modelos estadísticos multiplicativos (riesgos relativos, *odds ratios*, etc.) y/o en modelos estadísticos aditivos (diferencias de medias o proporciones, riesgos atribuibles, etc.). Lo habitual es que se explore la modificación del efecto en modelos multiplicativos.

Veamos un ejemplo. En un ensayo clínico han asignado a una muestra de 944 niños de alto riesgo de enfermedad celiaca a introducir el gluten entre los 4 y 6 meses o más tarde, para saber si influía en el riesgo de desarrollar la enfermedad. El 9,2% de los que recibieron precozmente (44/475) fueron diagnosticados de enfermedad celiaca frente a un 7,6% (36/469) de los que lo introdujeron a partir de los 6 meses. La introducción precoz de gluten se asoció a un riesgo de celiacía un 21% mayor (RR 1,21; IC 95 0,79 a 1,84), aunque la magnitud no era suficientemente precisa para poder asumir diferencias, ni a favor ni en contra (el intervalo de confianza

TABLA 1. FRECUENCIAS DE CRISIS ASMÁTICAS EN FUNCIÓN DE EXPOSICIÓN A TABACO EN EL HOGAR ESTRATIFICANDO POR EXPOSICIÓN A CONTAMINACIÓN AMBIENTAL. DATOS FIGURADOS

Contaminación ambiental	Exposición a tabaco en hogar	Crisis asmática		
		Sí	No	
Alta	Sí	50 (71,4%)	20	70
	No	50 (62,5%)	30	80
	Subtotal	100 (66,6%)	50	150
Baja	Sí	10 (38,4%)	16	26
	No	90 (29,4%)	216	306
	Subtotal	100 (30,1%)	232	332
Todos	Sí	60 (62,5%)	36	96
	No	140 (36,2%)	246	386
	Total	200 (41,4)	282	482

Figura 1. Fórmulas de Mantel-Haenszel para estimar RR y OR ajustados

Estudio de cohortes			
Estrato i	Efecto		
Exposición	Sí	No	
Sí	a _i	b _i	n _{1i}
No	c _i	d _i	n _{0i}
	m _{1i}	m _{0i}	T _i

$$RR_{MH} = \frac{\sum_{i=1}^s \frac{a_i n_{0i}}{T_i}}{\sum_{i=1}^s \frac{c_i n_{1i}}{T_i}}$$

Estudio de casos y controles			
Estrato i	Exposición		
Enfermedad	Sí	No	
Sí	a _i	b _i	n _{1i}
No	c _i	d _i	n _{0i}
	m _{1i}	m _{0i}	T _i

$$OR_{MH} = \frac{\sum_{i=1}^s \frac{a_i d_{0i}}{T_i}}{\sum_{i=1}^s \frac{b_i c_{1i}}{T_i}}$$

incluía el I). Sin embargo, queríamos saber si algún otro factor ha podido influir en los resultados, en concreto el sexo de los pacientes.

En la **Tabla 2** se presentan las frecuencias de enfermedad celiaca en niños de alto riesgo en función de la introducción precoz o tardía de gluten, estratificando por sexos.

Lo primero que observamos es que el riesgo se invierte de un estrato a otro. En las niñas, la introducción precoz parece aumentar el riesgo, mientras que en los niños parece disminuir. Esto se traduce en estimaciones de riesgo también invertidas, con importantes diferencias entre estratos: el RR de la introducción precoz de gluten para niñas es 1,92, de manera que se comporta como factor de riesgo en niñas y parece de protección en niños, aunque en los niños la estimación no sea significativa (el intervalo de confianza incluye el 1). Haciendo el cálculo ponderado de ambos estratos, el RR ajustado es 1,21 (IC 95: 0,79 a 1,84), similar al estimado crudo global anteriormente. En este ejemplo, el sexo se comporta como factor de interacción o modificador del efecto y ni la estimación cruda ni la ponderada son válidas, solo las estimaciones de cada estrato.

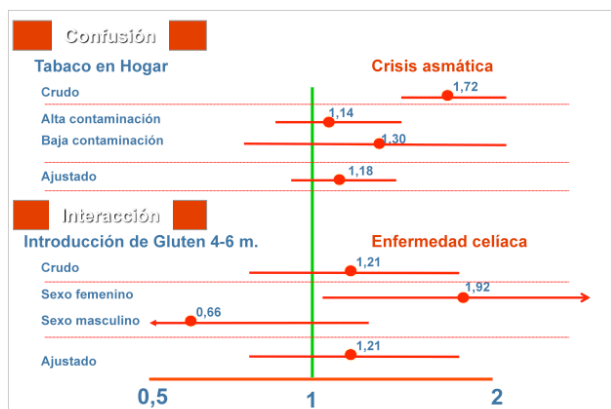
En la **Figura 2** se presentan las estimaciones crudas, por estratos y ajustadas para los ejemplos de confusión e interacción de las **Tablas 1 y 2**. Si comparamos las estimaciones crudas y ajustadas, con los factores de confusión son diferentes, mientras que con los de interacción son similares. Además, podemos ver que, cuando hay confusión, las estimaciones de cada estrato se desplazan en el mismo sentido, habitualmente hacia el valor nulo, como ocurre en nuestro caso, aunque podrían también alejarse. Sin embargo, en la interacción las estimaciones por estratos se alejan de la estimación cruda, en sentidos opuestos; en nuestro ejemplo, alejándose del valor nulo, aunque esto dependerá del tipo de interacción.

Es muy importante tener en cuenta que confusión e interacción se interpretan de forma distinta. Siempre que encontremos una variable de confusión, nuestro objetivo será controlar su efecto y tratar de estimar una medida de asociación ajustada. Sin embargo, las variables modificadoras de efecto reflejan una característica de la relación entre exposición y efecto cuya intensidad depende de la variable modificadora. En estos casos, aunque podemos calcular una medida ajustada, como hacemos con los factores de confusión, esta describe peor la asociación entre exposición y efecto que las esti-

TABLA 2. FRECUENCIAS DE ENFERMEDAD CELIACA EN NIÑOS DE ALTO RIESGO EN FUNCIÓN DE LA INTRODUCCIÓN PRECOZ O NO DE GLUTEN GLOBAL Y POR ESTRATOS DE SEXO (ESTUDIO PREVENT-CD)

Sexo	Introducción de gluten	Enfermedad celiaca		
		Sí	No	
Femenino	4-6 meses	31 (13,5%)	197	228
	≥6 meses	16 (7,0%)	210	226
	Subtotal	47 (10,3%)	407	454
Masculino	4-6 meses	13 (5,2%)	234	247
	≥6 meses	20 (8,2%)	223	243
	Subtotal	33 (6,7%)	457	490
Todos	4-6 meses	44 (9,2%)	431	475
	≥6 meses	36 (7,6%)	433	469
	Total	80 (8,4%)	864	944

Figura 2. Ejemplos de estimaciones crudas, por estratos y ajustadas de un factor de confusión y otro de interacción



maciones por estratos. Por lo tanto, lo que debemos hacer cuando descubrimos una modificación de efecto es describirla y tratar de interpretarla.

En ocasiones, una tercera variable es, a la vez, factor de confusión y variable modificadora de efecto. En estos casos, la interacción siempre es más importante que la confusión, y es preciso ofrecer los resultados por estratos o subgrupos.

Aunque la detección de interacciones puede tener importantes implicaciones prácticas, su búsqueda exagerada, creando subgrupos numerosos o de escasa importancia clínica, puede dar lugar a hallazgos espurios o falsos. En general, siguiendo el principio de parsimonia, los modelos explicativos más simples son preferidos a los modelos complejos.

PASOS DEL ANÁLISIS ESTRATIFICADO

Veamos los pasos a seguir en el análisis estratificado:

1. Calculamos el RR o el OR en crudo (es decir, sin estratificar).
2. Estratificamos por la tercera variable a controlar y calculamos el RR u OR específico de cada estrato.
3. Evaluamos la homogeneidad de las estimaciones del efecto entre los estratos. Aunque la homogeneidad puede ser valorada comparando las magnitudes de las estimaciones, existen test estadísticos para contrastarla (ej.: test de homogeneidad de Breslow-Day, test de homogeneidad de Mantel-Haenszel), que proporcionan las calculadoras y paquetes estadísticos, aunque cuando los estratos tienen pocos recuentos estos test pueden perder potencia estadística.
4. Si hay homogeneidad en las estimaciones del efecto entre los estratos, calculamos el RR u OR general ajustado mediante la fórmula de Mantel-Haenszel. En la **Figura 1** se presentan las fórmulas de Mantel-Haenszel para estimar RR y OR ajustados.

5. Si hay homogeneidad entre estratos y existen diferencias entre las estimaciones cruda y ajustada, existe sesgo de confusión. Calculando la diferencia entre ambas estimaciones podemos estimar la magnitud de la confusión; habitualmente, una diferencia superior al 10% se considera suficiente, e importante si es superior al 20%.
6. Si hay heterogeneidad entre estratos, probablemente existe interacción o modificación del efecto. En ese caso debemos informar las estimaciones de cada estrato por separado.

Para realizar análisis estratificados recomendamos emplear calculadoras epidemiológicas, como la versión 3.1 de Epidat (solo disponible para Windows: [www.sergas.es/Saude-publica/Epidat-3-1-descargar-Epidat-3-1-\(espanol\)](http://www.sergas.es/Saude-publica/Epidat-3-1-descargar-Epidat-3-1-(espanol))). Otra calculadora disponible es la versión en línea de EpiInfo (desarrollado por los CDC), OpenEpi (www.openepi.com/TwoByTwo/TwoByTwo.htm) y las versiones App instalables en móviles Android e IOs. Para cálculos con solo dos estratos se puede usar Statulator (<http://statulator.com/stat/StratifiedAnalysis.html>). El lenguaje de programación R cuenta también con paquetes especializados, como EpiR o EpiTools.

En la **Figura 3** se presenta la salida de resultados de los análisis estratificados, realizados con Epidat 3.1, de los datos de las **Tablas 1 y 2**. Muestra el cálculo de los RR para cada estrato, los globales crudos y los combinados o ajustados por estratos de Mantel-Haenszel. Asimismo, se presentan los test de homogeneidad, en este caso, los de Mantel-Haenszel, que permiten identificar si los estratos son o no homogéneos.

En la **Figura 3A** podemos ver que para el estudio del asma el RR crudo global es 1,72 (paso 1) y que los RR se alejan de ese valor en el mismo sentido manteniéndose cercanos entre sí (1,14 y 1,30 con intervalos solapados) (paso 2). También vemos que la prueba de homogeneidad indica que no hay heterogeneidad entre estratos ($p = 0,63$) (paso 3), por lo que procedemos a buscar la estimación combinada o ajustada (paso 4), en este caso un RR de 1,18, que es claramente diferente de la estimación cruda global. Podemos, por tanto, asumir que la contaminación ambiental es un factor de confusión para la asociación entre tabaquismo familiar y asma, por lo que solo la estimación ajustada es válida (paso 5). Además, vemos que la diferencia entre las estimaciones cruda y ajustada es importante, superior al 20%.

En la **Figura 3B** vemos los RR crudo global y por estratos del estudio de la enfermedad celiaca (pasos 1 y 2); los RR de cada estrato se alejan de la estimación cruda (1,20) y entre sí (1,92 para niñas y 0,63 para niños). Comprobamos que existe heterogeneidad significativa entre estratos ($p = 0,0151$) (paso 3), por lo que no podemos combinar sus resultados y solo son válidas las estimaciones por estratos (paso 6). En este análisis vemos que las estimaciones crudas y combinada son iguales, pero ninguna de las dos informa del riesgo, solo las estimaciones por estratos.

Figura 3. Se presenta la salida de resultados de los análisis estratificados, realizados con Epidat 3.1 de los datos de las Tablas 1 (A) y 2 (B)

A.- Riesgo de Asma asociado a Tabaquismo familiar Sibilantes por estratos de contaminación ambiental alta (1) o baja (2)			
Estrato	RR	IC (95, 0%)	
1	1,142857	0,912312	1,431662 (Katz)
2	1,307692	0,780361	2,191368 (Katz)
Crudo	1,723214	1,405630	2,112553 (Katz)
Combinado (M-H)	1,181097	0,955669	1,459700
Ponderado	1,167759	0,949889	1,435601
Prueba de homogeneidad			
	Ji-cuadrado	gl	Valor p
Combinado (M-H)	0,2314	1	0,6305
Ponderado	0,2198	1	0,6392
B.- Riesgo de enfermedad celíaca asociado a la introducción precoz de gluten por estratos de sexo femenino (1) o masculino (2)			
Estrato	RR	IC (95, 0%)	
1	1,920504	1,081111	3,411616 (Katz)
2	0,639474	0,325451	1,256491 (Katz)
Crudo	1,206784	0,791656	1,839594 (Katz)
Combinado (M-H)	1,207640	0,791882	1,841679
Ponderado	1,210306	0,781311	1,874849
Prueba de homogeneidad			
	Ji-cuadrado	gl	Valor p
Combinado (M-H)	5,9077	1	0,0151
Ponderado	5,9077	1	0,0151

El análisis estratificado tiene dos limitaciones importantes a la hora de controlar factores de confusión: (1) si hay más de un factor de confusión, la aplicación de esta fórmula es laboriosa debido al mayor número de estratos y además obliga a tener un tamaño de muestra relativamente grande; y (2) este método requiere que los factores de confusión continuos se recodifiquen en un número limitado de categorías, lo que podría generar confusión residual (un fenómeno particularmente relevante cuando la variable se clasifica en unos pocos estratos). Algunas de estas limitaciones se pueden superar usando técnicas de análisis multivariante.

BIBLIOGRAFÍA

- Arezina R, Duolao W. Source and control of bias. En: Duolao W, Bakhai A (eds.). Clinical trials. A practical guide to design, analysis and reporting. Londres: Remédica; 2006. p. 55-64.
- Argimón Pallás JM, Jimenez Villa J. Métodos de investigación clínica y epidemiológica. Barcelona: Elsevier; 2006.
- Bases metodológicas de la investigación clínica y epidemiológica. En: Argimón Pallás JM, Jiménez Villa J (eds.). Métodos de investigación clínica y epidemiológica. Madrid: Elsevier; 2004. p. 8-15.

- Confusión y modificación del efecto. En: Argimón Pallás JM, Jiménez Villa J (eds.). *Métodos de investigación clínica y epidemiológica*. Madrid: Elsevier; 2004. p. 278-88.
- Molina Arias M, Ochoa Sangrador C. Errores en epidemiología. Errores sistemáticos. Factores de confusión y modificación del efecto. *Evid Pediatr*. 2016;12:16.
- Murray KW, Duggan A. Understanding confounding in research. *Pediatr Rev*. 2010;31:124-6.
- Rosner B. *Fundamentals of Biostatistics*, 7th Edition. Boston: Brooks/Cole, Cengage Learning 2011.
- Rothman K.J. *Epidemiología Moderna*. Madrid: Díaz de Santos, 1987.
- Sordillo JE, Scirica CV, Rifas-Shiman SL, Gillman MW, Bunyavanich S, Camargo CA Jr, *et al.* Prenatal and infant exposure to acetaminophen and ibuprofen and the risk for wheeze and asthma in children. *J Allergy Clin Immunol*. 2015;135:441-8.
- Woodward M. *Epidemiology study design and data analysis*. London; Chapman & Hall/CRC, 1999.
- Vriezinga SL, Auricchio R, Bravi E, Castillejo G, Chmielewska A, Crespo Escobar P, *et al.* Randomized feeding intervention in infants at high risk for celiac disease. *N Engl J Med* 2014;371:1304-15.