

ANÁLISIS ESTRATIFICADO EN ESTUDIOS  
EPIDEMIOLÓGICOS DE OBSERVACIÓN  
AJUSTE DEL RIESGO RELATIVO

Por: Kahl-Martin Colimon S,  
Profesor de Epidemiología FNSP - U. de  
Antioquia  
Medellín - COLOMBIA

Junio de 1986

INTRODUCCION

En ocasiones, cuando se plantea el análisis de un estudio observacional, se ve la necesidad de controlar una variable porque se sospecha su posible participación como variable de confusión. Una de las formas de realizar este control en el diseño de estudio, y la única en la etapa de ejecución del análisis, es por medio de la estratificación.

Estratificación:

La estratificación consiste en organizar grupos homogéneos internamente y heterogéneos entre sí, con respecto a la variable que supuestamente produce confusión.

Se obtiene entonces una estimación de la medida de asociación en cada estrato, y finalmente se ponderan las diversas estimaciones en una sola medida (global y ajustada).

Por ejemplo, si en un estudio de observación se desea controlar la variable edad por estratificación (después de que la revisión de la literatura indique que pueda ser una variable de confusión en la relación de un factor de riesgo  $F_r$  y de un efecto  $E$ ) distribuida en cuatro categorías o estratos, de 15-24

años, de 25-34, de 35-44 y de 45-64 años, se obtendrán las 4 tablas de contingencia (de 2 por 2 en este ejemplo) numeradas de 1 a 4 según el procedimiento siguiente:

Para el análisis estratificado, será necesario realizar varios pasos, a saber:

- 1) Estratificar según categorías de importancia de dicha variable.
- 2) Calcular la estimación del riesgo relativo para cada estrato o categoría.
- 3) Calcular el Chi-cuadrado para cada estrato.
- 4) Obtener el factor de ponderación en cada estrato.
- 5) Obtener el riesgo relativo global ajustado.
- 6) Interpretar los resultados.

Los pasos citados se ilustrarán utilizando un ejemplo para un estudio de casos y controles.

\* Se agradece la colaboración de los siguientes profesores de Epidemiología de la FNSP - U. de Antioquia Juan Luis Londoño F. y Gabriela Ospina de Arboleda.

EDAD 15 – 24 AÑOS

	E +	E -	
Fr+	a1	b1	m1 1
Fr-	c1	d1	m2 1
	g1 1	g2 1	n 1

EDAD 25 – 34 AÑOS

	E +	E -	
Fr+	a2	b2	m1 2
Fr-	c2	d2	m2 2
	g1 2	g2 2	n 2

EDAD 35 – 44 AÑOS

	E +	E -	
Fr+	a3	b3	m1 3
Fr-	c3	d3	m2 3
	g1 3	g2 3	n 3

EDAD 45 – 64 AÑOS

	E +	E -	
Fr+	a4	b4	m1 4
Fr-	c4	d4	m2 4
	g1 4	g2 4	n 4

ANALISIS ESTRATIFICADO EN ESTUDIO DE "CASOS Y CONTROLES"

AJUSTE DE RIESGO RELATIVO

Los aspectos mencionados anteriormente se ampliarán en este modelo para el estudio de casos y controles.

Paso No. 1 : El esquema general de las tablas de contingencia de 2 por 2 para estudios de casos y controles es el siguiente para cada estrato:

en donde:

- Fr+ : Presencia del factor de riesgo.
- Fr- : Ausencia del factor de riesgo.
- E+ : Presencia del efecto.
- E- : Ausencia del efecto.
- a1 : Casos expuestos al factor de riesgo.
- b1 : Controles expuestos al factor de riesgo.
- c1 : Casos no expuestos al factor de riesgo.

	E +	E -	
Fr+	ai	bi	m1 i
Fr-	ci	di	m2 i
	g1 i casos	g2 i controles	ni

- d1 : Controles no expuestos al factor de riesgo.
- ni : o sea (ai + bi + ci + di).
- g1i : Total de casos en el estrato, o sea (ai + ci).
- g2i : Total de controles en el estrato (bi + di).
- m1i : Total de sujetos en el estrato con el Fr+ , es sea (ai + bi).
- m2i : Total de sujetos en el estrato con el Fr- , es decir (ci + di).

Paso No. 2: La estimación del riesgo relativo en cada estrato se obtendrá así:

$$R_{ri} = \frac{a_i * d_i}{b_i * c_i}$$

Paso No. 3: El chi-cuadrado en cada estrato, con 1 grado de libertad (fórmula corta) será:

$$X^2_{1i} = \frac{(a_i d_i - b_i c_i)^2 * n_i}{g_{1i} * g_{2i} * m_{1i} * m_{2i}}$$

Paso No. 4: El cálculo del factor de ponderación  $w_i$  en cada estrato, en estudios de casos y controles se traduce por:

$$w_i = \frac{b_i * c_i}{n_i}$$

Paso No. 5: La estimación del riesgo relativo global ajustado ( $R_{r-g-a}$ ) se obtiene así:

$$R_{r-g-a} = \frac{\sum (R_{ri} * W_i)}{\sum (w_i)} = \frac{\sum \left( \frac{a_i d_i}{b_i c_i} * \frac{b_i c_i}{n_i} \right)}{\sum \left( \frac{b_i * c_i}{n_i} \right)}$$

$$= \frac{\sum \left( \frac{a_i * d_i}{n_i} \right)}{\sum \left( \frac{b_i * c_i}{n_i} \right)}$$

en donde  $\sum$  significa Sumatoria.

Paso No. 6: Es conveniente la obtención del Chi-cuadrado global correspondiente al riesgo relativo global ajustado. Para el efecto, se aplicará la fórmula siguiente:

$$X^2_{g-a} = \frac{\left( \sum \frac{(a_i d_i - b_i c_i)^2}{n_i} \right)^2}{\sum \left( \frac{g_{1i} * g_{2i} * m_{1i} * m_{2i}}{(n_i - 1) (n_i)^2} \right)}$$

Si la variable por la cual se controla (mediante la estratificación), la edad en este ejemplo, no actúa como factor de confusión, no hará diferencia apreciable entre las medidas globales de asociación obtenidas con la ponderación o sin ella.

Ejemplo numérico:

Los siguientes datos extraídos de un estudio de casos y controles correlacionan un factor de riesgo ocupacional con el cáncer de vejiga. Como se presume que la edad pueda ser un factor de confusión en esta relación, se estratifican los resultados en 3 grupos de edad: de 20-59 años, de 60-74 años, y de 75 años y más.

Paso No. 1.: Los datos se presentan en las siguientes tablas, en donde las 3 primeras representan la  $i = 3$  estratos de edad, y la última la combinación de todas las edades.

EDAD 20 – 59 AÑOS

	E+	E-	
Fr+	41	29	
Fr-	55	77	
	96	106	202

EDAD 60 – 74 AÑOS

	E+	E-	
Fr+	65	41	
Fr-	99	134	
	164	175	339

EDAD 75 AÑOS Y MAS

	E+	E-	
Fr+	30	25	
Fr-	66	68	
	96	93	189

TODAS LAS EDADES

	E+	E-	
Fr+	136	95	
Fr-	220	279	
	356	374	730

en donde

F = Exposición a factor de riesgo ocupacional.  
E = Cáncer de Vejiga.

Para facilitar el trabajo mecánico del cálculo del factor de ponderación, del riesgo relativo en cada estrato, del riesgo relativo global ajustado y del chi-cuadrado de asociación para cada estrato, se presentarán en un solo cuadro de trabajo los resultados obtenidos en los 3 estratos y la combinación de las 3 tablas.

CUADRO DE TRABAJO

Edad	ai	bi	ci	di	ni	g1 i	g2 i	m1 i	m2 i	wi	Rri	X2
20-59	41	29	55	77	202	96	106	70	132	7.896	1.98	5.24
60-74	65	41	99	134	339	164	175	106	233	11.973	2.15	10.3
75 y +	30	25	66	68	189	96	93	55	134	8.730	1.24	0.43
Todas edad	136	95	220	279	730	356	374	231	499	---		

Paso No. 2: Estimación del Rri en cada i estrato:

En el estrato o categoría de 20 – 59 años se tendrá:

$$Rr_{20-59} = \frac{41 * 77}{29 * 55} = 1.98$$

Paso No. 3: Obtención del Chi-cuadrado en cada i estrato:

En el estrato de 20 – 59 años, sería:

$$\chi^2_{20-59 \text{ años}} = \frac{(41*77 - 29*55)^2}{1} \times \frac{202}{96 * 106 * 70 * 132} = 5.24$$

Paso No. 4: Cálculo del factor de ponderación  $w_i$  en cada estrato.

Para el estrato de 20 — 59 años, se tendrá:

$$w_i = \frac{29 * 55}{202} = 7.89$$

Paso No. 5: Estimación del riesgo relativo global ajustado

La fórmula del riesgo relativo global ajustado  $Rr_{g-a}$  es la siguiente:

$$Rr_{g-a} = \frac{\sum (w_i * R_{ri})}{\sum (w_i)}$$

$$= \frac{(7.89 * 1.98) + (11.97 * 2.15) + (8.73 * 1.24)}{7.89 + 11.97 + 8.73} = 1.82$$

Esta fórmula que sigue para el riesgo relativo global ajustado  $Rr_{g-a}$  es igual a la anterior:

$$Rr_{g-a} = \frac{\sum \left( \frac{a_i * d_i}{n_i} \right)}{\sum \left( \frac{b_i * c_i}{n_i} \right)}$$

$$= \frac{\frac{41 * 77}{202} + \frac{65 * 134}{339} + \frac{30 * 68}{189}}{\frac{29 * 55}{202} + \frac{41 * 99}{339} + \frac{25 * 66}{189}} = 1.82$$

Paso No. 6: Interpretación

Adicionalmente se puede calcular el chi-cuadrado global ajustado de asociación de la manera siguiente:

$$\left( \frac{41 * 77 - 29 * 55}{202} + \frac{65 * 134 - 41 * 99}{339} + \frac{30 * 68 - 25 * 66}{189} \right)^2 =$$

$$\frac{96 * 106 * 70 * 132}{201 * 202^2} + \frac{164 * 175 * 106 * 233}{338 * 339^2} + \frac{96 * 93 * 55 * 134}{188 * 189^2}$$

$$\chi^2_{g-a} = 13.99, \text{ correspondiente a } p < 0.001.$$

Sería benéfico para el estudio de casos y controles calcular los límites de confianza para el riesgo relativo. De los dos límites, superior e inferior, el límite inferior de confianza es el más importante en un estudio con hipótesis de causalidad en donde el límite inferior debe ser mayor de 1.

La fórmula general para el intervalo de confianza (IC) del riesgo relativo es la siguiente:

$$IC_{(1-\alpha)\%} = Rr \left( 1 \pm Z_{(1-\alpha/2)} \sqrt{\frac{1}{Rr^2}} \right)$$

en donde:

$$Rr_{g-a} = 1.82$$

$$\alpha = 0.05$$

$$Z_{(1-\alpha/2)} = 1.96$$

$$\chi^2_{g-a} = 13.99$$

Dicha fórmula se traduce entonces por:

IC ( 95%) para el riesgo relativo  $g-a$ ;

$$= 1.82 \left( 1 \pm 1.96 / \sqrt{13.99} \right)$$

De allí, el límite superior de confianza (LSC) para el riesgo relativo  $g-a$  es:

$$LSC = 1.82 \left( 1 + 1.96 / \sqrt{13.99} \right) = 1.82 (1 + 0.524)$$

$$= 1.82 (1.524) = 2.49$$

El límite inferior de confianza para el riesgo relativo  $g-a$  (LIC):

$$LIC = 1.82 \left( 1 - 0.524 \right) = 1.82 (0.476) = 1.33$$

De acuerdo con dichos resultados, se podría afirmar con una confianza muy alta (del 95%) que el valor del riesgo relativo se encuentra entre 1.33 y 2.49 (después de controlar el efecto de la

edad). Si se trata de demostrar que existe asociación positiva entre el Fr y el E, se advertirá que el valor menor del riesgo relativo es de 1.33, y que por lo tanto existe, al menos una débil asociación.

#### BIBLIOGRAFIA

1. ARMITAGE, P. Statistical Methods in Medical Research New York, Willey, 1971.
2. COLIMON, K.M. Fundamentos de Epidemiología – Medellín, Colimón, 1978.
3. KLEINBAUM, D., KUPPER, L., MORGENSTERN, H. Epidemiologic Research, Principles and Quantitative Methods –London, Lifetime Learning Publications, 1982.
4. LILIENFELD, A., LILIENFELD, D. – Fundamentos de Epidemiología – Bogotá, Fondo Educativo Interamericano, 1983.
5. LONDOÑO, J.L. Análisis Fundamental en Estudios Epidemiológicos de Tipo Analítico. Medellín, Publicación Interna Facultad Nacional de Salud Pública U. de Antioquia, 1981.
6. LONDOÑO, J.L. Obtención del riesgo relativo ajustado por el método de Mantel Haenszel, Medellín. Publicación Interna Facultad Nacional de Salud Pública U de Antioquia, 1981.
7. MACMAHON, B., PUGH, T., Epidemiology: Principles and Methods – Boston, Little Brown, 1970.
8. MANTEL, N., HAENSZEL, W., Statistical Aspects of the Analysis of Data from Retrospective Studies of Disease. J. N. Cancer Inst. 22, 719-748, 1959.
9. MIETTINEN, O.S., Estimability and Estimation in Case-referent studies. Am J of Epidem 103, 498-502.