

## Métodos de integración de odds ratio basados en meta-análisis usando modelos de efectos fijos y aleatorios útiles en salud pública

### Integration Methods of Odds Ratio Based on Meta-Analysis Using Fixed and Random Effect Models Useful in Public Health

MÓNICA CATALÁN<sup>1,a</sup>, M. PURIFICACIÓN GALINDO<sup>2,b</sup>, JAVIER MARTÍN<sup>2,c</sup>,  
VÍCTOR LEIVA<sup>1,d</sup>

<sup>1</sup>DEPARTAMENTO DE ESTADÍSTICA, UNIVERSIDAD DE VALPARAÍSO, VALPARAÍSO, CHILE

<sup>2</sup>DEPARTAMENTO DE ESTADÍSTICA, UNIVERSIDAD DE SALAMANCA, SALAMANCA, ESPAÑA

---

#### Resumen

Un meta-análisis integra información proveniente de varios estudios con el propósito de generar un resultado común para un problema determinado. En la literatura nos encontramos con varios métodos de integración de resultados, siendo el más básico el método de integración de niveles de probabilidad y, con una complejidad mayor, el método de integración del tamaño del efecto. Este último hace uso de modelos de efectos fijos y aleatorios. En este estudio, comparamos los resultados de dos métodos de estimación del tamaño del efecto basados en un meta-análisis usando modelos de efectos fijos y aleatorios. La medida del tamaño del efecto considerada en este estudio es el odds ratio, debido a que esta medida es usada frecuentemente en revisiones sistemáticas de varios temas de interés en salud pública, tales como cáncer cérvico uterino, colecistectomía laparoscópica, enfermedades cardiovasculares, enfermedad de Parkinson y tabaquismo. Las conclusiones de este trabajo indican las condiciones de aplicabilidad de los estimadores analizados del odds ratio en función de la magnitud del efecto poblacional, de la variabilidad entre estudios, del tamaño del meta-análisis y de los tamaños muestrales de tales estudios.

**Palabras clave:** bioestadística, ensayos clínicos, medicina, tamaño del efecto.

---

<sup>a</sup>Profesora auxiliar. E-mail: monica.catalan@uv.cl

<sup>b</sup>Profesora titular. E-mail: pgalindo@usal.es

<sup>c</sup>Profesor titular. E-mail: jmv@usal.es

<sup>d</sup>Profesor titular. E-mail: victor.leiva@uv.cl

### Abstract

Meta-analysis integrates information from different studies to generate a common response to a determined problem. In the literature, we find several integration methods of results, with the integration method of levels of probability being the more basic and, with a greater complexity, the integration method of the effect size, which uses fixed and random effect models. In this study, we compare the results of two estimation methods of the effect size based on meta-analysis using fixed and random effect models. The measure of the effect size considered here is the odds ratio, due to this measure is frequently used in systematic reviews of several topics of interest in public health, such as heart diseases, laparoscopic colectomy, Parkinson disease, tobacco addiction and uterine cervical cancer. Conclusions of this work indicate the applicability conditions of the analyzed estimators of the odds ratio in function of the size of the population effect, of the variability among studies, of the size of the meta-analysis and of the sample sizes of such studies.

**Key words:** Biostatistics, Clinical trials, Effect size, Medicine.

## 1. Introducción

Un meta-análisis integra resultados de varios estudios con el fin de generar una respuesta común frente a un problema de investigación determinado (Glass 1976, Martín, Donaldson, Villarroel, Parmar, Ernst & Higginson 2002, Catalán & Galindo 2003, Burguillo, Martín, Barrera & Bardsley 2010). En general, una revisión sistemática es actualmente reconocida como la búsqueda organizada de literatura de un tema específico, mientras que un meta-análisis estudia de manera estadística esa información que ha sido organizada previamente (Glass 1976, Rosenthal 1984, Vamvakas 2011). La integración de niveles de probabilidad fue uno de los primeros métodos estadísticos usados para sintetizar cuantitativamente los resultados de un conjunto de estudios. Para ese fin, se han desarrollado diversos métodos con la limitación que éstos sólo permiten determinar si se rechaza o no una hipótesis nula, sin indicar cuál es el tamaño del efecto o el grado de influencia de cada estudio en el resultado que se genera (Rosenthal 1984). Los métodos de integración, cuando el tamaño del efecto es el odds ratio (OR), el riesgo relativo o la diferencia de riesgo, presentan un gran desarrollo en la literatura científica, dado que proporcionan mayor información sobre la magnitud del efecto, permitiendo inferir un resultado desde los obtenidos de un conjunto de estudios (Rosenthal 1984, Hedges & Olkin 1985).

El OR es una de las medidas del tamaño del efecto más comúnmente utilizada en ensayos clínicos aleatorizados, donde la variable de respuesta dicotómica se registra para dos conjuntos de sujetos, usualmente llamados grupo tratado y grupo control. Además, el OR se utiliza en otros estudios de interés clínico como las asociaciones con factores de riesgo o en las pruebas de diagnóstico. Sin embargo, para la aplicación de los modelos principales para la integración del tamaño del efecto en meta-análisis, es más apropiado trabajar con el logaritmo del OR estimado, dado que éste cumple con mayor facilidad el supuesto de normalidad (Turner, Omar, Yang, Goldstein & Thompson 2000, Leyland &

Goldstein 2001, Catalán & Galindo 2003). Antecedentes actualizados siguen evidenciando que el OR es una medida del tamaño del efecto presente en revisiones sistemáticas por meta-análisis en varios temas de interés en salud pública, tales como cáncer cérvico uterino (Rydzewska, Tierney, Vale & Symonds 2010), colecistectomía laparoscópica (Claros, Manterota, Vial & Sanhueza 2007, Zhou, Zhang, Wang & Hu 2009), enfermedades cardiovasculares (Moores, Jackson, Shorr & Jackson 2004, Cornelissen 2007, Dentali, Douketis, Lim & Crowther 2007), enfermedad de Parkinson (Allam, Del Castillo & Navajas 2003, Stowe, Ives, Clarke & van Hilten 2008) y tabaquismo (Jiménez-Ruiz, Riesco, Ramos & Barrueco 2008).

En un meta-análisis, los métodos de integración del tamaño del efecto se han analizado desde dos perspectivas, y éstas son:

- (i) Considerando un modelo de efectos fijos (M1): en este caso, la hipótesis de partida es la existencia de un único tamaño del efecto poblacional y sólo se considera la variabilidad debido al muestreo, o
- (ii) Considerando un modelo de efectos aleatorios (M2): en este otro caso, se parte de una megapoblación de tamaños del efecto y, por tanto, se contempla una nueva variabilidad debido a la diferencia entre estudios. Cada estudio estima un tamaño del efecto de esa población.

No obstante, de acuerdo a lo que se describe en la literatura, la elección entre estos dos modelos (M1 y M2) es un tema de extensa discusión para los investigadores meta-analíticos (Hedges & Vevea 1998, Berlin, Laird, Sacks & Chalmers 1989). Por una parte, el modelo de efectos fijos asume homogeneidad de los parámetros correspondientes a los efectos de los estudios, de modo que el tamaño del efecto es una constante fija desconocida que debe ser estimada. Por otra parte, el modelo de efectos aleatorios supone heterogeneidad de los parámetros correspondientes a los efectos de los estudios, y así cada estudio representa una población. Por consiguiente, este último tipo de modelos (M2) permite descomponer la varianza de los resultados de los estudios en una parte que corresponde a la variación muestral y otra que refleja las diferencias reales entre estos estudios.

Existen varios métodos que se pueden usar para estimar los parámetros en cada tipo de modelo (M1 y M2), haciendo que la decisión de utilizar uno u otro método en el desarrollo de un meta-análisis resulte más compleja. Dos de los estimadores más utilizados en los meta-análisis para la integración del OR bajo el modelo M1 en el campo clínico son el estimador clásico de media ponderada (conocido como DerSimonian-Laird y que llamaremos “clásico”) y el de Peto (que llamaremos “peto”) (Petitti 1994). Bajo el modelo M2, el estimador más utilizado es el de media ponderada, que incluye la estimación de la variabilidad entre estudios propuesta por DerSimonian & Laird (1986). Otro estimador muy utilizado en estudios clínicos es el de Mantel-Haenzel, pero éste tiene problemas en la estimación de su variabilidad. Existen también otros estimadores propuestos en la bibliografía sobre el tema que el lector interesado puede revisar en (Greenland & Salvan 1990). En consecuencia, para elegir el tipo de modelo y el método de estimación que se debe usar en un estudio de meta-análisis, es necesario considerar las características de los estudios que intervienen en el meta-análisis y el problema para el que se pretende obtener un resultado común. Esto quiere decir que, por

una parte, si bien los estudios que se integran tratan un problema similar, éstos pueden presentar varias características relacionadas, por ejemplo, al número y al tipo de pacientes en cada uno de ellos, a las diferencias en su diseño y al lugar donde estos estudios se realizan. Por otra parte, los resultados que proporciona un método de integración basados en meta-análisis podrían depender del tamaño del efecto que se pretende estimar, es decir, de un efecto del tratamiento mayor o menor, de la varianza entre los estudios, del número de estudios involucrados en el meta-análisis y del número de individuos considerados.

Frente a las alternativas de elección entre modelos de efectos fijos y de efectos aleatorios para la estimación del tamaño del efecto mediante meta-análisis (Hedges & Vevea 1998), el presente estudio responde a la pregunta de investigación acerca de qué diferencias existen entre los métodos de estimación clásico y peto en los modelos M1 y M2 cuando el tamaño del efecto es el OR. Esto permite valorar el impacto de las distintas condiciones en las diferencias obtenidas por un modelo o por otro. En consecuencia, la hipótesis de investigación es que existen diferencias entre los resultados que proporcionan los métodos de integración del tamaño del efecto en meta-análisis.

El objetivo principal de este artículo es comparar los resultados que proporcionan dos métodos de estimación del tamaño del efecto (clásico y peto) bajo los dos modelos considerados habitualmente en meta-análisis (M1 y M2). Específicamente, se pretende conocer el comportamiento de estos estimadores en función de la magnitud del efecto poblacional, de la variabilidad entre estudios, del tamaño del meta-análisis y de los tamaños muestrales de los estudios. En el caso del modelo de efectos fijos M1, se consideran el estimador clásico y el estimador peto, métodos que llamaremos ef-clásico y ef-peto, respectivamente. En el caso del modelo de efectos fijos M2, se utilizan los estimadores clásico y peto incluyendo además el estimador de la variabilidad entre estudios propuesto por Dersimonian-Laird, métodos que llamaremos ea-clásico y ea-peto, respectivamente.

El resto de este artículo está organizado como sigue. En la sección 2 describimos los materiales y métodos de este estudio. En la sección 3 presentamos los resultados del estudio. En la sección 4 discutimos los resultados obtenidos en la sección 3. En la sección 5 bosquejamos las conclusiones de este trabajo.

## 2. Materiales y métodos

En esta sección, proporcionamos los materiales y métodos de este estudio que incluyen la definición de las unidades en estudio y las variables a considerar, la generación del meta-análisis y los métodos de integración del tamaño del efecto.

En este trabajo se diseñó un estudio de simulación donde se generaron los datos necesarios para un conjunto de 81 meta-análisis a los que se les aplicaron los dos métodos para estimar los parámetros de los modelos M1 y M2 para integración del tamaño del efecto. Específicamente, en este artículo (i) describimos los resultados generados por los dos métodos de estimación para el conjunto de meta-análisis y (ii) determinamos si existen diferencias entre los métodos de estimación y entre los

modelos, en relación al valor estimado del tamaño del efecto, permitiendo valorar el impacto de las distintas condiciones en las diferencias obtenidas por un modelo o por otro. Recalamos que la hipótesis de investigación es que existen diferencias entre los resultados que proporcionan los métodos de integración del tamaño del efecto en meta-análisis.

## 2.1. Unidades de análisis y variables

Las unidades en estudio son los meta-análisis. En cada uno de ellos se consideran los modelos M1 y M2 basados en el supuesto distribucional de normalidad para la integración del tamaño del efecto (OR estimado) mediante la aplicación de los métodos ef-clásico, ef-peto, ea-clásico y ea-peto. Las variables en estudio son los resultados para el OR estimado, su logaritmo ( $\log(\widehat{OR})$ ) y la varianza de  $\log(\widehat{OR})$  estimada.

## 2.2. Meta-análisis

Debido a que el objetivo principal de este artículo es comparar los resultados que proporcionan los métodos de estimación clásico y peto del tamaño del efecto bajo los modelos M1 y M2 usando meta-análisis, entonces necesitamos un número importante de meta-análisis. Este número nos permite obtener los datos a nivel de cada estudio considerado en cada meta-análisis y los datos específicos de los individuos. Sin embargo, obtener una cantidad grande de meta-análisis basados en estudios reales para lograr el objetivo de este estudio está fuera de nuestro alcance, por lo que se optó por la alternativa de generar los datos individuales para los estudios que intervienen en cada meta-análisis a través de un proceso de simulación. Para esto, se estableció que la variable respuesta dentro de cada estudio corresponde a la presencia o a la ausencia de una enfermedad bajo un factor de exposición o de riesgo como lo es un tratamiento o un control. La medida del efecto considerada aquí en cada estudio es el OR correspondiente al odds ratio del grupo tratado en relación al grupo control y dado por

$$OR = \frac{p_t/(1-p_t)}{p_c/(1-p_c)}$$

donde  $p_t$  y  $p_c$  son las probabilidades de presencia de la enfermedad en los grupos tratado y control, respectivamente. De esta manera,  $\log(OR) = \text{logit}(p_t) - \text{logit}(p_c)$ , donde  $\text{logit}(p_t) = \log(p_t/(1-p_t))$  y  $\text{logit}(p_c) = \log(p_c/(1-p_c))$  son las funciones logito correspondientes.

Para hacer inferencias estadísticas para el parámetro OR, es necesario disponer de la distribución del estimador del OR. Sin embargo, ya que el OR está acotado inferiormente en cero, puesto que por definición éste no puede tomar valores negativos, y el OR no tiene una cota superior, su estimador suele seguir una distribución asimétrica que impide asumir una distribución normal. Entonces, para evitar este problema, se suele trabajar con el logaritmo del OR y así suponer una

distribución normal para el logaritmo natural del estimador del OR,  $\widehat{\text{OR}}$ , esto es,

$$\log(\widehat{\text{OR}}) \sim N(E[\log(\widehat{\text{OR}})], \text{Var}[\log(\widehat{\text{OR}})])$$

donde  $\log(\widehat{\text{OR}}) = \text{logit}(\widehat{p}_t) - \text{logit}(\widehat{p}_c)$ , con  $\text{logit}(\widehat{p}_t) \sim N(\theta_t, \sigma_t^2)$  y  $\text{logit}(\widehat{p}_c) \sim N(\theta_c, \sigma_c^2)$ . Aquí,

- (i)  $E[\log(\widehat{\text{OR}})] = \theta$  es el valor esperado del estimador del logaritmo natural del OR o tamaño del efecto y
- (ii)  $\text{Var}[\log(\widehat{\text{OR}})] = \tau^2 = \sigma_t^2 + \sigma_c^2$  es la varianza verdadera entre estudios, donde  $\sigma_t^2$  y  $\sigma_c^2$  son las varianzas de los grupos tratado y control, respectivamente.

Entonces, para llevar a cabo el proceso de simulación de los datos, se consideran los parámetros siguientes:

- (i) Tamaño del efecto ( $\theta$ );
- (ii) Varianza poblacional entre estudios ( $\tau^2$ );
- (iii) Número de estudios del meta-análisis ( $J$ );
- (iv) Número de individuos dentro de cada estudio ( $n$ );
- (v) Número de individuos dentro de cada estudio del grupo tratado ( $n_t$ ) y
- (vi) Número de individuos dentro de cada estudio del grupo control ( $n_c$ ).

Para asignar el número de individuos en cada estudio, se definió un indicador de la proporción de estudios en un meta-análisis ( $p$ ) con un número de individuos determinado ( $n$ ). Para cada uno de los parámetros de simulación ( $\theta, \tau^2, J, n$ ) y el indicador ( $p$ ), se seleccionaron tres escenarios distintos con valores considerados como “bajo”, “moderado” y “alto”, basándonos en los valores que utilizan algunos meta-análisis descritos en la literatura (Turner et al. 2000, Coomarasamy, Papaioannou, Gee & Khan 2001). Usando también como referencia estos estudios previos y de acuerdo al número de parámetros establecidos y a los valores de cada uno de ellos, se generaron datos para un total de 81 meta-análisis, que reúnen 1.755 estudios y 739.080 individuos. Estos 739.080 individuos corresponden a la suma de todos los individuos de todos los estudios en todos los meta-análisis. Estos datos se generaron basados en (i) tres valores de tamaño del efecto poblacional (indicado como  $\log(\text{OR})$ :  $-0,106$ ,  $-0,714$ ,  $-1,599$ ), (ii) tres valores de la varianza poblacional ( $0,015$ ;  $0,15$  y  $0,8$ ), (iii) tres cantidades de estudios dentro de cada meta-análisis ( $10$ ,  $20$  y  $35$ ), (iv) tres cantidades de individuos dentro de cada estudio ( $20$  para el grupo tratado y  $20$  para el grupo control,  $150$  para el grupo tratado y  $150$  para el grupo control y  $500$  para el grupo tratado y  $500$  para el grupo control) y (v) la proporción de estudios con un tamaño específico de individuos dentro de cada meta-análisis. Esto se explica porque generalmente los estudios que forman parte de un meta-análisis tienen un número distinto de individuos. En este trabajo se establecieron los porcentajes siguientes de estudios dentro de un meta-análisis con

un número distinto de individuos: (a) 30 % de estudios con 40 individuos en total, 60 % de estudios con 300 individuos en total y 10 % de estudios con 1.000 individuos en total; (b) 10 % de estudios con 40 individuos en total; 70 % de estudios con 300 individuos en total y 20 % de estudios con 1.000 individuos en total; y (c) 10 % de estudios con 40 individuos en total; 50 % de estudios con 300 individuos en total y 40 % de estudios con 1.000 individuos en total. Los 1.755 estudios corresponden a la suma de todos los estudios generados según lo establecido anteriormente. Los 81 meta-análisis resultan a partir de multiplicar 3 tamaños del efecto, 3 valores para las varianzas, 3 cantidades de estudios y 3 proporciones de estudios dentro de cada meta-análisis ( $3 \times 3 \times 3 \times 3 = 81$ ).

De esta manera, sobre la base del supuesto de normalidad y dados los valores de los parámetros de simulación  $(\theta, \tau^2, J, n)$  y el indicador  $(p)$  dados en la tabla 1, se obtienen los valores de la media y la varianza del logit( $\hat{p}_t$ ) para el grupo tratado y del logit( $\hat{p}_c$ ) para el grupo control. La generación de datos para cada estudio del meta-análisis se realiza usando el algoritmo siguiente de cuatro pasos:

**Paso 1.** Generar  $J$  observaciones de logit( $\hat{p}_t$ ) y logit( $\hat{p}_c$ ) desde una distribución normal con media  $\theta$  y varianza  $\tau^2$  establecidas.

**Paso 2.** Calcular las probabilidades de tener la enfermedad en los grupos tratado y control,  $p_t$  y  $p_c$ , respectivamente, para cada una de las  $J$  observaciones generadas en el Paso 1.

**Paso 3.** Obtener los datos individuales en los grupos tratado y control para cada uno de los  $J$  estudios dentro del meta-análisis desde distribuciones binomiales con parámetros  $n$  y  $p_t$ , y  $n$  y  $p_c$ , donde, como se mencionó,  $n$  es el número de individuos y  $p_t$  y  $p_c$  son las probabilidades de presentar la enfermedad en los grupos tratado y control, respectivamente. Para cada uno de los niveles de los parámetros establecidos (bajo, moderado, alto), como se mencionó, se utilizaron 3 opciones para la proporción de estudios en un meta-análisis con un número determinado de individuos. Estas opciones son (ver últimas tres filas de la tabla 1):

- (i) 30 % de los estudios con  $n = 40$ , 60 % con  $n = 300$  y 10 % con  $n = 1000$ ;
- (ii) 10 % de los estudios con  $n = 40$ , 70 % con  $n = 300$  y 20 % con  $n = 1000$ ;
- (iii) 10 % de los estudios con  $n = 40$ ; 50 % con  $n = 300$  y 40 % con  $n = 1000$ .

**Paso 4.** Resumir las respuestas individuales de los grupos tratado y control en cada uno de los  $J$  estudios dentro de un meta-análisis en una tabla de contingencia  $2 \times 2$ , cuyas variables dicotómicas son la enfermedad (presencia/ausencia) y el factor de exposición (tratamiento/control).

El proceso de generación de datos basado en el algoritmo anterior se debe ejecutar para cada uno de los 81 meta-análisis en estudio usando los valores dados en la tabla 1. Una vez generados los datos individuales que corresponden a la

TABLA 1: Escenario del estudio de simulación.

| Parámetro | Valores establecidos |          |        |
|-----------|----------------------|----------|--------|
|           | Bajo                 | Moderado | Alto   |
| $\theta$  | -0,106               | -0,714   | -1,599 |
| $\tau^2$  | 0,015                | 0,15     | 0,8    |
| $J$       | 10                   | 20       | 35     |
| $n$       | 40                   | 300      | 1000   |
| $n_t$     | 20                   | 150      | 500    |
| $n_c$     | 20                   | 150      | 500    |
| $p_t$     | 0,10                 | 0,16     | 0,06   |
| $p_c$     | 0,11                 | 0,28     | 0,24   |
| $p$       | 0,3                  | 0,6      | 0,1    |
|           | 0,1                  | 0,7      | 0,2    |
|           | 0,1                  | 0,5      | 0,4    |

respuesta de los individuos (739.080 en total) en cada estudio (1.755 en total), se deben estimar el OR, su logaritmo ( $\log(\text{OR})$ ) y la varianza del estimador del logaritmo del OR para cada estudio dentro de los 81 meta-análisis.

### 2.3. Métodos de integración del tamaño del efecto

Considere el modelo de efectos fijos (M1)

$$Y_j = \theta + \varepsilon_j, \quad \varepsilon_j \sim N(0, \sigma_\varepsilon^2)$$

y el modelo de efectos aleatorios (M2)

$$Y_j = \mu_j + \varepsilon_j, \quad \mu_j = \theta + u_j, \quad u_j \sim N(0, \tau^2), \quad \varepsilon_j \sim N(0, \sigma_\varepsilon^2)$$

donde

- (i)  $Y_j$  es la variable respuesta en el estudio  $j$ -ésimo;
- (ii)  $\theta$  es el tamaño del efecto;
- (iii)  $\varepsilon_j$  es el error aleatorio;
- (iv)  $\sigma_\varepsilon^2$  es la varianza del error aleatorio;
- (v)  $\mu_j$  es el tamaño del efecto en el estudio  $j$ -ésimo;
- (vi)  $u_j$  es el error en el estudio  $j$ -ésimo y
- (vii)  $\tau^2$  es la varianza entre estudios.

Para el modelo M1, el estimador de  $\theta$  y su error estándar están dados por

$$\hat{\theta} = \frac{\sum_{j=1}^J w_j Y_j}{\sum_{j=1}^J w_j} \quad \text{y} \quad \sigma_{\hat{\theta}} = \frac{1}{\sqrt{\sum_{j=1}^J w_j}}$$



donde  $w_j = 1/\text{Var}[Y_j]$ , con  $\text{Var}[Y_j] = \sigma_\varepsilon^2$  conocida. En este modelo, se consideran los métodos de estimación ef-clásico (Petitti 1994) y ef-peto (Yusuf, Peto, Lewis, Collins & Sleight 1985) para la integración del OR. Para el modelo M2, el estimador de  $\theta$  y su error estándar están dados por

$$\hat{\theta} = \frac{\sum_{j=1}^J w_j^* Y_j}{\sum_{j=1}^J w_j^*} \quad \text{y} \quad \sigma_{\hat{\theta}} = \frac{1}{\sqrt{\sum_{j=1}^J w_j^*}}$$

donde  $w_j^* = 1/(\sigma_\varepsilon^2 + \hat{\tau}^2)$ . En este modelo, se consideran los métodos de estimación ea-clásico (DerSimonian & Laird 1986) y ea-peto (Martín 1995) para la integración del OR. Así, en general, para M1 y M2, un intervalo de confianza (IC) del  $100 \times (1 - \alpha)\%$  para  $\theta$  está dado por

$$\text{IC}(\theta)_{100 \times (1 - \alpha)\%} = \left[ \hat{\theta} \pm z_{1 - \alpha/2} \sigma_{\hat{\theta}} \right]$$

donde  $z_{1 - \alpha/2}$  es el percentil  $1 - \alpha/2$  de la distribución normal estándar. Los métodos de estimación son aplicados a cada meta-análisis en estudio a través de un programa computacional disponible en la literatura. Específicamente, dados el número de integraciones por realizar y la información específica requerida para este estudio, se utilizó un programa computacional desarrollado en Excel por Martín (1995). Previamente, los resultados de este programa fueron contrastados con otros programas comerciales tales como Metawin (Rosenberg, Adams & Gurevitch 2000) y uno de libre acceso como Mix v1.56 (Bax, Yu, Ikeda, Tsuruta & Moons 2006).

Las variables en estudio para los 81 meta-análisis son los resultados alcanzados de la aplicación de los métodos de estimación considerados. Específicamente, el OR estimado, su logaritmo  $-\log(\widehat{\text{OR}})$  y la varianza de este logaritmo obtenidos con los métodos de estimación ef-clásico, ea-clásico, ef-peto y ea-peto. Estos 81 meta-análisis fueron divididos en tres conjuntos de 27 meta-análisis cada uno, de acuerdo al tamaño del efecto poblacional definido para el estudio. En el primer grupo se consideró una eficacia baja del tratamiento ( $\text{OR} = 0,90$ ), es decir, que el riesgo de presentar la enfermedad en el grupo control varía muy poco con respecto al grupo que recibe el tratamiento; específicamente, la variación es de 1,1. En el segundo grupo se consideró un efecto moderado del tratamiento ( $\text{OR} = 0,49$ ), donde el riesgo de presentar la enfermedad en el grupo control es el doble que en el grupo tratado. En los 27 meta-análisis restantes correspondientes al tercer grupo se consideró un efecto alto del tratamiento ( $\text{OR} = 0,20$ ), lo que significa que el riesgo de presentar la enfermedad en el grupo control es cinco veces mayor que en el grupo tratado. Para comparar el  $\log(\widehat{\text{OR}})$  mediante los métodos ef-clásico, ea-clásico, ef-peto y ea-peto en cada uno de los tres grupos de meta-análisis ( $\text{OR} = 0,90$ ;  $\text{OR} = 0,49$  y  $\text{OR} = 0,20$ ), se utilizó la prueba  $t$ -Student para diferencia de medias. Habitualmente, la representación gráfica de los resultados del tamaño del efecto de los estudios involucrados en un meta-análisis se hace mediante el forest plot o gráfico de OR o riesgo relativo (Abrams & Jones 1995, Rodríguez 2002, Moores et al. 2004). En este estudio, utilizamos el forest plot con fines prácticos para observar los intervalos de confianza y los OR estimados con los distintos métodos de estimación, representando además el tamaño del efecto poblacional establecido a priori.

### 3. Resultados

En esta sección proporcionamos los resultados más relevantes de este estudio. Sin embargo, si el lector interesado requiere resultados más específicos, éstos pueden solicitarse a los autores (Catalán 2003). Específicamente, con respecto al OR promedio, los dos métodos de estimación aplicados a los 81 meta-análisis arrojaron los siguientes resultados:

- (i) Para los 27 meta-análisis donde el OR poblacional presenta un efecto bajo de tratamiento ( $OR = 0,90$ ), el OR estimado promedio es igual para todos los métodos, tomando un valor de 0,95. Sin embargo, el método ef-peto para el modelo M1 presenta una variación mayor.
- (ii) Para los 27 meta-análisis donde el OR poblacional presenta un efecto moderado del tratamiento ( $OR = 0,49$ ), el OR estimado promedio es igual tanto para el método ef-clásico como para ea-peto, tomando éste un valor de 0,52. Lo mismo sucede con los métodos ef-peto y ea-clásico, donde el OR estimado promedio toma un valor de 0,51. La variación mayor se observa con el método ef-clásico, y la variación menor con el método ea-peto.
- (iii) Para el conjunto de meta-análisis donde el OR poblacional presenta un efecto alto del tratamiento ( $OR = 0,20$ ), el OR estimado promedio con el método ef-clásico es 0,24, mientras que con el método ea-clásico este valor es 0,20. Los métodos ef-peto y ea-peto proporcionan el mismo resultado promedio, que toma un valor de 0,25. La variación mayor se observa con el método ef-clásico, y la variación menor con el método ea-peto.

En cuanto a la significación estadística del OR estimado, se tiene que:

- (i) Cuando el efecto del tratamiento es bajo ( $OR = 0,90$ ), el modelo M1 presenta el porcentaje mayor de valores significativos (ef-peto = 41 %, ef-clásico = 33 %), a diferencia de los métodos de estimación para el modelo M2, donde el porcentaje de OR estimado significativo es de un 7 % en ambos métodos (ea-clásico y ea-peto).
- (ii) Cuando el efecto del tratamiento es moderado ( $OR = 0,45$ ), los métodos de estimación bajo el modelo M1 entregan el 100 % de valores significativos, y un 93 % con los métodos de estimación para el modelo M2.
- (iii) Cuando el efecto del tratamiento es alto ( $OR = 0,20$ ), todos los métodos de estimación entregan un 100 % de resultados significativos.

Los resultados de las pruebas *t*-Student muestran que:

- (i) Para el conjunto de meta-análisis donde el efecto del tratamiento es bajo ( $OR = 0,90$ ) y moderado ( $OR = 0,49$ ), no se observan diferencias significativas para el OR estimado promedio entre los dos métodos de estimación empleados en el análisis (valores- $p > 0,10$ ).

- (ii) Cuando el efecto del tratamiento es alto ( $OR = 0,20$ ), el método ea-clásico en el modelo M2 difiere significativamente ( $\text{valor-}p < 0,02$ ) de los métodos ef-clásico, ef-peto y ea-peto, los que presentan valores mayores para el OR estimado promedio.

En general, el método de estimación de Peto usado para estimar los parámetros del modelo M1 y su adaptación para el modelo M2 proporcionan resultados similares.

Sobre la base de los resultados anteriores, se generó una representación gráfica basada en el forest plot de los IC para el OR sobre el conjunto de meta-análisis, donde se reflejan las diferencias entre los métodos de estimación. En las figuras 1-3 se observa en términos generales que la amplitud de los IC para el OR es diferente a medida que la varianza entre estudios aumenta. Específicamente,

- (i) Desde las figuras 1 y 2 se observa que los métodos ef-clásico y ea-clásico dan resultados similares y que lo mismo sucede con los métodos ef-peto y ea-peto. Sin embargo, se observa que los métodos ef-peto y ea-peto sobrestiman el OR, y generalmente el IC no contiene su valor verdadero.
- (ii) Desde la figura 3 se observa que los IC para el tamaño del efecto son más amplios, lo que refleja una variabilidad mayor entre los estudios ( $\tau^2 = 0,8$ ). Es en esta figura donde se observa una diferencia mayor entre los métodos de estimación. Específicamente, entre el método clásico para el modelo M2 y el resto, siendo el método ea-clásico el que proporciona resultados más aproximados al tamaño verdadero del efecto.

## 4. Discusión

En meta-análisis, una de las primeras decisiones que un investigador debe tomar es la elección del modelo bajo el cual va a integrar sus resultados. Esta decisión puede tomarla el investigador a priori en función del conocimiento metodológico del tema revisado. Sin embargo, en la bibliografía hay mucha controversia acerca de la elección del modelo bajo el cual se integran los resultados. Schmidt, Oh & Hayes (2009) consideraron que la aplicación del modelo de efectos fijos está limitada a estudios muy similares entre ellos. Por el contrario, Peto (1987) y Thompson & Pocock (1991) propusieron un modelo de efectos fijos argumentando que la integración debe limitarse a los estudios revisados con hipótesis que hagan referencia a los mismos y rechazar hipótesis inferenciales más ambiciosas que son difíciles de contrastar. Una práctica generalizada es la de partir del modelo de efectos fijos, y ante la presencia de heterogeneidad, el investigador debe utilizar factores elegidos a priori que puedan explicar esta heterogeneidad. En el caso de que el investigador no sea capaz de explicar dicha heterogeneidad, él debería considerar un modelo de efectos aleatorios, aunque Pocock & Hughes (1990) y Greenland & Salvan (1990) concluyeron que ninguna aproximación es buena en presencia de una heterogeneidad fuerte. Una de las conclusiones que se extrae de este trabajo es la diferencia entre los resultados del tamaño del efecto estimado que proporcionan los

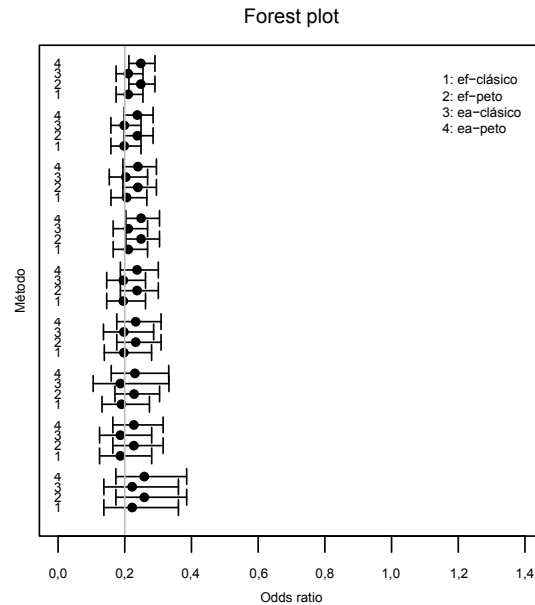


FIGURA 1: IC para el OR con 4 métodos de estimación para 9 meta-análisis, donde el tamaño del efecto verdadero es  $OR = 0,20$  y la varianza entre estudios es pequeña ( $\tau^2 = 0,015$ ).

modelos M1 y M2. Estas diferencias van a depender del tamaño del efecto que se pretende estimar y de la varianza entre estudios. Cuando el efecto de tratamiento en la población es bajo ( $OR = 0,90$ ) o moderado ( $OR = 0,49$ ), no se observan diferencias significativas entre los métodos de estimación para modelos de efectos fijos y aleatorios. Las diferencias mayores se producen con la mayor variabilidad considerada ( $\tau^2 = 0,8$ ). Sin embargo, cuando el efecto de tratamiento es alto ( $OR = 0,20$ ), se detectan diferencias significativas entre los métodos. Esta diferencia refleja tanto el tamaño del efecto en la población como la variabilidad presente entre los estudios. Cuando la varianza entre estudios es grande, el modelo más adecuado es M2, como cabría esperar, lo que se observa al comparar los resultados del método clásico de la media ponderada bajo los modelos M1 y M2. Nuestros resultados coinciden con los obtenidos por Berlin et al. (1989), donde se encuentra que el método ef-peto es similar al método ea-clásico cuando la heterogeneidad no es muy grande. No obstante, entre una gama de procedimientos de estimación descritos en la literatura (DerSimonian & Laird 1986), existen algunos métodos que no reflejan las características de la población en estudio. Uno de éstos es el método de Peto cuando se considera un modelo de efectos aleatorios, ya que este método proporciona resultados similares al método de Peto cuando se considera un modelo de efectos fijos. Esta similitud puede deberse a que el estimador de Peto para el modelo M2 se ha construido de manera artificial a partir del estimador de Peto considerando un modelo de efectos fijos. Por tanto, deducimos que el estimador de Peto para modelos de efectos aleatorios no es una buena aproximación cuando sólo existen efectos fijos. En cuanto a la significación del valor estimado del OR,

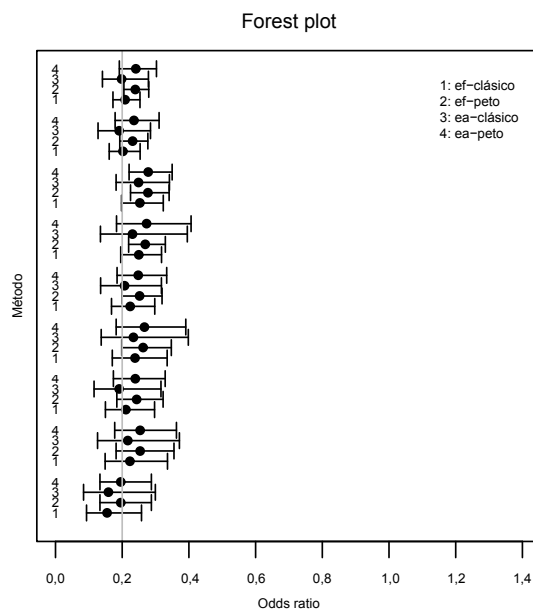


FIGURA 2: IC para el OR con 4 métodos de estimación para 9 meta-análisis, donde el tamaño del efecto verdadero es  $OR = 0,20$  y la varianza entre estudios es moderada ( $\tau^2 = 0,15$ ).

se tiene que los modelos de efectos fijos son más sensibles para detectar un efecto de tratamiento pequeño. Sin embargo, cuando el efecto de tratamiento en estudio es alto, los dos métodos de estimación empleados entregan resultados significativos, pero se diferencian en la amplitud de los intervalos de confianza y en cómo el OR estimado se aproxima a su valor verdadero. Engels, Schmid, Terrin, Olkin & Lau (2000) analizaron varios meta-análisis que utilizan el OR y riesgos absolutos concluyendo que, al comparar los modelos M1 y M2, existe un incremento lógico de los errores estándar en el modelo M2, lo que produce menos significatividad. El método clásico de la media ponderada es el que proporciona las mejores estimaciones del OR y, según se observa, los valores estimados con el método de Peto usando los modelos M1 y M2 son mayores que el valor verdadero, es decir, hay una sobreestimación del efecto de tratamiento. El estimador de Peto bajo el supuesto de efectos fijos es insesgado bajo la hipótesis de independencia, aunque el sesgo aumenta cuando el OR poblacional se aleja de un valor igual a uno. Además, este sesgo es mayor cuando los estudios individuales son balanceados, situación poco común en los estudios no experimentales (Greenland & Salvan 1990).

Tal como se ha planteado ampliamente en la literatura, para la realización de un meta-análisis se hace necesario conocer en profundidad el problema en estudio, de tal forma que se tengan referencias acerca de la variabilidad de los resultados en cada estudio involucrado y del efecto de tratamiento que se pretende estimar (Egger, Smith & Phillips 1997). El crecimiento en el número de meta-análisis publicados en los últimos años ha llevado a elaborar normas de publicación para

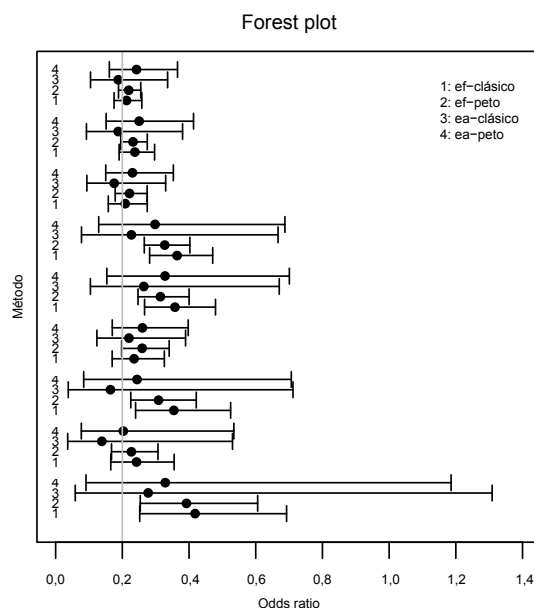


FIGURA 3: IC para el OR con 4 métodos de estimación para 9 meta-análisis, donde el tamaño del efecto verdadero es  $OR = 0,20$  y la varianza entre estudios es alta ( $\tau^2 = 0,8$ ).

facilitar su interpretación y utilización. Sobre esto, en la propuesta QUOROM (quality of reporting of meta-analysis) se identifican varios ámbitos por abordar y, específicamente, el punto 14 trata sobre la síntesis cuantitativa de los datos y sobre los métodos de combinación de los resultados (Moher, Cook, Eastwood, Olkin & Rennie 1994, Urrutia, Tort & Bonfill 2005). Moher, Liberati, Tetzlaff & Altman (2009) revisaron y actualizaron las líneas maestras para la realización de meta-análisis que se denomina PRISMA (preferred reporting items for systematic reviews and meta-analysis).

Como en cualquier investigación, en estudios de meta-análisis la elección del método de análisis de datos se fundamenta en las características del problema y en los objetivos e hipótesis que se plantean. Por esta razón, se hace necesario conocer las ventajas y desventajas de los procedimientos estadísticos que se podrían utilizar. En este sentido, el resultado de la estimación del tamaño del efecto en un meta-análisis no depende solamente del diseño de los estudios individuales, sino que también puede depender del tipo de modelo y del procedimiento de estimación que se emplea. Para más antecedentes, el lector interesado puede revisar las guías de Cochrane, las que son ampliamente utilizadas y validadas para meta-análisis en salud (<http://www.cochrane-handbook.org>).

## 5. Conclusiones

En este estudio hemos comparado los resultados de dos métodos de estimación del tamaño del efecto basados en meta-análisis usando modelos de efectos fijos y aleatorios. La medida del efecto considerada en este estudio fue el odds ratio. Esto se debe a que tal medida del tamaño del efecto está presente en revisiones sistemáticas de temas de interés en el campo clínico procedentes de diseños experimentales, observacionales o pruebas de diagnóstico. Hemos observado diferencias entre los dos modelos analizados en cuanto al porcentaje de estimaciones de odds ratio significativos. Con respecto al tamaño del efecto estimado, cuando el efecto del tratamiento es alto, el resultado del método de DerSimonian-Laird difiere significativamente de los otros métodos, mientras que el método de Peto y su versión adaptada presentan resultados similares. En presencia de heterogeneidad entre los estudios, el método de DerSimonian-Laird es el que más se aproxima a los resultados verdaderos, mientras que el método de Peto no es una buena aproximación.

## Agradecimientos

Los autores desean agradecer a las editoras invitadas de este número especial en bioestadística de la Revista Colombiana de Estadística, Dra. Liliana López Kleine y Dra. B. Piedad Urdinola, y a dos árbitros anónimos por sus comentarios constructivos sobre la versión original de este artículo que condujeron a esta versión final mejorada. La investigación de Víctor Leiva fue parcialmente financiada por el proyecto FONDECYT 1120879 del gobierno de Chile.

[Recibido: septiembre de 2011 — Aceptado: febrero de 2012]

## Referencias

- Abrams, K. & Jones, D. (1995), 'Meta-analysis and the synthesis of evidence', *IMA Journal of Mathematics Applied in Medicine and Biology* **12**, 297–313.
- Allam, M. F., Del Castillo, A. S. R. & Navajas, F. (2003), 'Enfermedad de Parkinson temprana y tabaco: metanálisis', *Revista de Neurología* **12**, 1101–1103.
- Bax, L., Yu, L. M., Ikeda, N., Tsuruta, H. & Moons, K. G. M. (2006), 'Enfermedad de Parkinson temprana y tabaco: metanálisis', *BMC Medical Research Methodology* **6**.
- Berlin, J. A., Laird, N. M., Sacks, H. S. & Chalmers, T. (1989), 'A comparison of statistical methods for combining event rates from clinical trials', *Statistics in Medicine* **8**, 141–151.
- Burguillo, F. J., Martín, F. J., Barrera, I. & Bardsley, W. G. (2010), 'Meta-analysis of microarray data: The case of imatinib resistance in chronic myelogenous leukemia', *Computational Biology & Chemistry* **34**, 184–192.

- Catalán, M. (2003), Los modelos multinivel como herramienta de análisis de datos biomédicos con estructura jerárquica, Tesis doctoral, Universidad de Salamanca, Departamento de Estadística.
- Catalán, M. & Galindo, M. P. (2003), 'Utilización de los modelos multinivel en investigación sanitaria', *Gaceta Sanitaria* **17**(3), 35–52.
- Claros, N., Manterota, C., Vial, M. & Sanhueza, A. (2007), 'Efectividad de la profilaxis antibiótica en el curso de la colecistectomía laparoscópica electiva. Revisión sistemática de la literatura', *Revista Chilena de Cirugía* **59**, 353–359.
- Coomarasamy, A., Papaioannou, S., Gee, H. & Khan, K. S. (2001), 'Aspirin for the prevention of preeclampsia in women with abnormal uterine artery doppler: A meta-analysis', *Obstetrics and Gynecology* **98**, 861–866.
- Cornelissen, V. A. (2007), 'Incidence of cardiovascular events in white-coat, masked and sustained hypertension versus true normotension: a metaanalysis', *Journal of Hypertension* **25**, 2193–2198.
- Dentali, F., Douketis, J. D., Lim, W. & Crowther, M. (2007), 'Combined aspirin-oral anticoagulant therapy compared with oral anticoagulant therapy alone among patients at risk for cardiovascular disease: A meta-analysis of randomized trials', *Archives of Internal Medicine* **167**, 117–124.
- DerSimonian, R. & Laird, N. (1986), 'Meta-analysis in clinical trials', *Controlled Clinical Trials* **7**, 177–188.
- Egger, M., Smith, G. D. & Phillips, A. N. (1997), 'Meta-analysis: Principles and procedures', *British Medical Journal* **315**, 1533–1537.
- Engels, E. A., Schmid, C. H., Terrin, N., Olkin, I. & Lau, J. (2000), 'Heterogeneity and statistical significance in meta-analysis: An empirical study of 125 meta-analyses', *Statistics in Medicine* **19**, 1707–1728.
- Glass, G. V. (1976), 'Primary, secondary and meta-analysis of research', *Educational Researcher* **6**, 3–8.
- Greenland, S. & Salvan, A. (1990), 'Bias in the one-step method for pooling study results', *Statistics in Medicine* **9**, 247–252.
- Hedges, L. & Olkin, I. (1985), *Statistical Methods for Meta-Analysis*, Academic Press, New York.
- Hedges, L. V. & Vevea, J. L. (1998), 'Fixed and random effects in meta-analysis', *Psychological Methods* **3**, 486–504.
- Jiménez-Ruiz, C., Riesco, J. A., Ramos, A. & Barrueco, M. (2008), 'Recomendaciones para el tratamiento farmacológico del tabaquismo. Propuestas de financiación', *Archivos de Bronconeumología* **44**, 213–219.
- Leyland, A. H. & Goldstein, H. (2001), *Multilevel Modelling of Health Statistics*, Wiley, New York.



- Martín, J. (1995), Métodos estadísticos en meta-análisis, Ph.d. thesis, Universidad de Salamanca, España.
- Martín, J., Donaldson, A. N. A., Villarroel, R., Parmar, M. K. B., Ernst, E. & Higginson, I. J. (2002), 'Efficacy of acupuncture in asthma: Systematic review and meta-analysis of published data from 11 randomised controlled trials', *European Respiratory Journal* **20**, 846–852.
- Moher, D., Cook, D. J., Eastwood, S., Olkin, I. & Rennie, D. (1994), 'Improving the quality of reporting of meta-analysis of randomized controlled trials: The quorum statement', *Lancet* **354**, 1896–1900.
- Moher, D., Liberati, A., Tetzlaff, J. & Altman, D. J. (2009), 'The prisma group', *PLoS Medicine* **6**, 1–6.
- Moore, L., Jackson, W., Shorr, A. & Jackson, J. (2004), 'Meta-analysis: Outcomes in patients with suspected pulmonary embolism managed with computed tomographic pulmonary angiography', *Annals of Internal Medicine* **141**, 866–875.
- Pettiti, D. B. (1994), *Meta-Analysis, Decision Analysis, and Cost-effectiveness Analysis*, University Press, Oxford.
- Peto, R. (1987), 'Why do we need systematic overviews of randomized trials?', *Statistics in Medicine* **6**, 233–240.
- Pocock, S. T. & Hughes, M. D. (1990), 'Estimation issues in clinical trials and overviews', *Statistics in Medicine* **9**, 657–671.
- Rodríguez, G. (2002), 'Entendiendo los diagramas de odds-ratio de las revisiones sistemáticas', *CES Medicina* **16**, 66–72.
- Rosenberg, M., Adams, D. C. & Gurevitch, J. (2000), *MetaWin 2.0: Statistical software for meta-analysis*, Sinauer Associates, Sunderland.
- Rosenthal, R. (1984), *Meta-Analytic Procedures for Social Research*, Sage, Beverly Hills.
- Rydzewska, L., Tierney, J., Vale, C. L. & Symonds, P. R. (2010), 'Neoadjuvant chemotherapy plus surgery versus surgery for cervical cancer', *Cochrane Database of Systematic Reviews* **1**. CD007406.
- Schmidt, F. L., Oh, I. & Hayes, T. L. (2009), 'Fixed- versus random-effects models in meta-analysis: Model properties and an empirical comparison of differences in results', *British Journal of Mathematical & Statistical Psychology* **62**, 97–128.
- Stowe, R. L., Ives, N. J., Clarke, C. & van Hilten, J. (2008), *Tratamiento con agonistas dopaminérgicos para la enfermedad de Parkinson en sus etapas iniciales*, Wiley, Oxford. (Revisión Cochrane traducida).

- Thompson, S. G. & Pocock, S. J. (1991), 'Can meta-analysis be trusted?', *The Lancet* **338**, 1127–1130.
- Turner, R. M., Omar, R. Z., Yang, M., Goldstein, H. & Thompson, S. G. (2000), 'A multilevel model framework for meta-analysis of clinical trials with binary outcomes', *Statistics in Medicine* **19**, 3417–3432.
- Urrutia, G., Tort, S. & Bonfill, X. (2005), 'Metaanálisis (quorum)', *Medicina Clínica* **125**(1), 32–37.
- Vamvakas, E. C. (2011), Meta-analysis: A statistical method to integrate information provided by different studies, in A. M. Marchevsky & M. Wick, eds, 'Evidence Based Pathology and Laboratory Medicine', Springer, New York, pp. 149–171.
- Yusuf, S., Peto, R., Lewis, J., Collins, R. & Sleight, P. (1985), 'Beta blockade during and after myocardial infarction: An overview of randomized trials', *Progress in Cardiovascular Diseases* **27**, 335–371.
- Zhou, H., Zhang, J., Wang, Q. & Hu, Z. (2009), 'Meta-analysis: Antibiotic prophylaxis in elective laparoscopic cholecystectomy', *Alimentary Pharmacology and Therapeutics* **29**, 1086–1095.