

ODELO DE EFECTOS FIJOS Y MODELO DE **EFECTOS ALEATORIOS**

Julio Sánchez-Meca Fulgencio Marín Martínez Tania B. Huedo Medina

1. Introducción

La característica que mejor identifica al meta-análisis es el uso de los métodos estadísticos para integrar cuantitativamente los resultados de un conjunto de estudios empíricos sobre un mismo problema de investigación. Para alcanzar este objetivo, el investigador tiene que definir una medida de la magnitud del efecto que pueda obtenerse de forma homogénea en todos los estudios. El tamaño del efecto es una familia de índices estadísticos que tienen como objetivo cuantificar el grado o magnitud en que se manifiesta el fenómeno en cuestión en cada estudio empírico⁽¹⁾.

En el capítulo anterior hemos visto que dependiendo de las características del diseño de los estudios individuales y de la naturaleza de la variable dependiente, se pueden utilizar diferentes índices del tamaño del efecto. Es también habitual registrar en cada estudio individual algunas de sus características con objeto de ponerlas en relación con el tamaño del efecto y comprobar de esta forma si algunas de ellas pueden estar moderando o confundiendo el resultado del meta-análisis. Estas características pueden tener que ver con la calidad metodológica de los estudios (por ejemplo, asignación aleatoria versus no aleatoria de los pacientes a los grupos, tamaño muestral, utilización de enmascaramiento simple, doble, o ausencia de enmascaramiento, etc.), con el tratamiento aplicado (por ejemplo, principio activo del fármaco, dosis, periodicidad, tipo de intervención psicosocial, su duración, su intensidad, etc.), o bien con las características de las muestras de pacientes utilizadas (edad media de la muestra, género, gravedad de la enfermedad, estatus socioeconómico, etc.).

Una vez que se dispone de una estimación del tamaño del efecto de cada estudio y de las variables moderadoras o confusoras registradas en cada uno de ellos, el análisis estadístico de estos datos implica tomar el tamaño del efecto como la variable dependiente y las variables moderadoras o confusoras como potenciales factores explicativos de aquél. Las principales preguntas a las que puede responder un meta-análisis son:

(a) ¿Cuál es la magnitud del efecto medio?

- (b) ¿Es estadísticamente significativo el tamaño del efecto medio?
- (c) ¿Es el tamaño del efecto medio representativo de todos los estudios individuales? O lo que es lo mismo, ¿existe homogeneidad en tomo al tamaño del efecto medio?
- (d) Si no existe homogeneidad, ¿qué características de los estudios pueden estar moderando el tamaño del efecto? ¿Puede proponerse un modelo explicativo?

El objeto de este capítulo es ofrecer una introducción a los dos modelos estadísticos que actualmente están reconocidos como los más adecuados para abordar el análisis de los tamaños del efecto en un meta-análisis: los modelos de efectos fijos y de efectos aleatorios. Centraremos nuestra atención en cómo aplicar ambos modelos para abordar los apartados (a), (b) y (c) arriba propuestos, ya que el apartado (d) es objeto del siguiente capítulo. Por tanto, presentamos cómo estimar la magnitud del efecto medio de un meta-análisis, cómo evaluar la homogeneidad de los tamaños del efecto, y cómo contrastar su significación estadística y cómo construir un intervalo de confianza asumiendo ambos modelos estadísticos, proporcionando una interpretación conceptual de las diferencias entre ambos modelos, de modo que los lectores de estudios meta-analíticos sepan distinguir las consecuencias de aplicar uno u otro modelo.

2. Modelos estadísticos en meta-análisis

Existe en la actualidad un claro consenso hacia el reconocimiento de aquellos modelos meta-analíticos que en sus análisis estadísticos ponderan cada tamaño del efecto en función de su precisión. La precisión en la estimación de un tamaño del efecto está directamente relacionada con el tamaño muestral del estudio, de tal forma que cuanto mayor es el tamaño muestral mayor es su precisión. Ponderar en función de la precisión significa que el peso específico que cada tamaño del efecto individual ejerce en los cómputos meta-analíticos depende, básicamente, del tamaño muestral del estudio: los estudios con tamaños muestrales mayores ejercen un mayor peso en los cálculos meta-analíticos (por ejemplo, a la hora de calcular el tamaño del efecto medio del meta-análisis). Aun así, cabe distinguir entre dos modelos estadísticos específicamente diseñados para el meta-análisis: Los modelos de efectos fijos y de efectos aleatorios (2-5).

Los modelos de efectos fijos y aleatorios difieren en la concepción de la población de estudios de partida. En el *modelo de efectos fijos* se asume que los estudios incluidos en el meta-análisis están estimando a un mismo, y único, tamaño del efecto paramétrico o poblacional, θ , por lo que la

única variabilidad asumida en los estudios individuales es la debida a error de muestreo aleatorio, que se cuantifica estimando la varianza intra-estudio, V(T). Así pues, en el modelo de efectos fijos se asume que las diferencias entre los tamaños del efecto de los estudios individuales se deben tan sólo al hecho de que los estudios utilizan muestras de sujetos diferentes. Siendo T_i el i-ésimo tamaño del efecto de un conjunto de k tamaños del efecto independientes que están estimando a un mismo efecto poblacional, θ , el modelo matemático de efectos fijos se formula como:

$$T_i = \Theta + \varepsilon_i \tag{1}$$

donde ε_i representa el error de muestreo aleatorio a que están sometidas las estimaciones del tamaño del efecto, T_i .

En el *modelo de efectos aleatorios* se parte de la base de que los estudios estiman a una distribución de tamaños del efecto paramétricos en la población, una distribución que suele asumirse Normal, $\theta_i \sim N(\mu_{\theta}; \tau^2)$. Es decir, cada tamaño del efecto individual, T_{i} , estima a un tamaño del efecto poblacional diferente, θ_i . Esto implica que además de la variabilidad debida al error de muestreo o intra-estudio, $V(T_i)$, hay que contemplar también la variabilidad entre-estudios, τ^2 . En consecuencia, el modelo matemático incorpora dos términos de error, uno debido a la variabilidad de cada estudio, o variabilidad intra-estudio, ϵ_i , que coincide con la del modelo de efectos fijos, y otro que refleja variabilidad entre-estudios, u_i , como la desviación de cada estudio respecto del tamaño del efecto medio:

$$T_i = \mu_{\theta} + U_i + \varepsilon_i \tag{2}$$

Al haber dos términos de error, que se suponen independientes entre si, la varianza de los tamaños del efecto individuales es la suma de los dos tipos de variabilidad: la intra-estudio, $V(T_i)$, y la entre-estudios, τ^2 .

Ambos modelos implican ponderar cada estimación del tamaño del efecto por la inversa de la varianza. Por tanto, siendo w_i el factor de ponderación del i-ésimo estudio, mientras que en el modelo de efectos fijos el factor de ponderación de cada índice T_i viene determinado exclusivamente por la varianza intra-estudio, $w_i=1/V(T_i)$, en el de efectos aleatorios el factor de ponderación lo está en función de la varianza intra-estudio y la varianza entre-estudios, $w_i^*=1/[V(T_i)+\tau^2]$. En ambos casos, tanto las varianzas intra-estudio como la varianza entre-estudios tienen que ser estimadas a partir de los datos de los estudios individuales del meta-análisis. Las consecuencias de asumir

uno u otro modelo afectan al grado de generalización de los resultados del meta-análisis. En el modelo de efectos fijos, la generalización se limita a la población de estudios de características similares a los incluidos en el meta-análisis. En el modelo de efectos aleatorios, por el contrario, los resultados pueden generalizarse a una población mayor de posibles estudios.

2.1 ¿Cuál es la magnitud del efecto medio?

Tal y como se muestra en el capítulo anterior, a partir de cada estudio individual, se obtiene una estimación del tamaño del efecto, T_{μ} así como una estimación de la varianza intra-estudio, $V(T_{\mu})$, mientras que la varianza entre-estudios, τ^2 , tiene que estimarse a partir de las propias estimaciones del tamaño del efecto incluidas en el meta-análisis. Por tanto, la varianza entre-estudios es un valor constante para todos los estudios incluidos en ese meta-análisis. El procedimiento habitualmente utilizado para estimarla es el propuesto por DerSimonian y Laird⁽⁶⁾, basado en el método de los momentos:

$$\hat{\tau}^2 = \begin{cases} \frac{Q - (k-1)}{c} & (paraQ \ge k-1) \\ 0 & (paraQ < k-1) \end{cases}$$
 (3)

donde c se obtiene mediante:

$$c = \sum_{i=1}^{k} w_i - \left(\sum_{i=1}^{k} w_i^2 / \sum_{i=1}^{k} w_i\right)$$
 (4)

siendo $w_i=1/V(T_i)$ el factor de ponderación en función únicamente de la varianza intra-estudio, y Q es el estadístico de heterogeneidad, que se obtiene mediante:

$$Q = \sum_{i=1}^{k} w_i \left(T_i - \overline{T} \right)^2 \tag{5}$$

donde T es el tamaño del efecto medio ponderado desde el modelo de efectos fijos por la inversa de la varianza intra-estudio:0

(a)
$$i \mathcal{N} \sum_{i=1}^{\lambda} / i \mathcal{T}_i \mathcal{N} \sum_{i=1}^{\lambda} = \overline{\mathcal{T}}$$

Pero en el modelo de efectos aleatorios el factor de ponderación incluye no sólo varianza intra-estudio, sino también una estimación de la varianza entre-estudios:

$$\frac{z^{2} + (^{1}L)\Lambda}{I} = {^{1}M}$$

Por tanto, el tamaño del efecto medio en el modelo de efectos aleatorios, \overline{T} , se obtiene mediante:

(8)
$${}_{i}^{*}\mathcal{M}\sum_{I=i}^{\lambda}/{}_{i}T_{i}^{*}\mathcal{M}\sum_{I=i}^{\lambda}=\overline{T}$$

2.2 ¿Es estadísticamente significativo el tamaño del efecto medio? Una vez calculado el tamaño del efecto medio del meta-análisis, ya sea por el modelo de efectos fijos o aleatorios, puede obtenerse un intervalo de confanza en torno al mismo, asumiendo una distribución normal.

Para el modelo de efectos fijos el intervalo de confianza del tamaño del efecto medio se obtiene como:

$$(6) \qquad (\underline{T}) \sqrt{|T_{\alpha/2}| \pm T}$$

* Es preciso aclarar que el método de análisis estadístico aquí presentado para los modelos de efectos fijos y de efectos aleatorios se basan en ponderar por la inversa de la varianza de cada estimación del tamaño del efecto (método de Woolf). En el capítulo antenor se ha visto que existen otros métodos más específicos, como el método de Peto (Odda Ratios) y el de Mantel-Haenszel (Odda Ratios, Riesgos Relativos y Diferencias de Riesgos), asumiendo un modelo de efectos fijos.

donde $z_{\alpha/2}$ es el valor de la distribución normal estándar correspondiente a una probabilidad acumulada igual a $\alpha/2$, y siendo $(1-\alpha)$ el nivel de confianza del intervalo, y $V(\overline{T})$ es la varianza del tamaño del efecto medio, que se obtiene mediante:

$$V(\overline{T}) = 1/\sum_{i=1}^{k} w_i \tag{10}$$

Análogamente, el intervalo de confianza del tamaño del efecto medio desde el modelo de efectos aleatorios se obtiene como:

$$\overline{T}^* \pm /z_{\alpha/2} / \sqrt{V(\overline{T}^*)} \tag{11}$$

donde $V(\overline{T}^*)$ es la varianza del tamaño del efecto medio, que se obtiene mediante:

$$V(\overline{T}^*) = 1/\sum_{i=1}^k w_i^*$$
 (12)

La información que nos ofrecen estos intervalos de confianza es muy valiosa, ya que nos permiten comprobar si el tamaño del efecto medio es estadísticamente significativo, es decir, si por ejemplo el grupo experimental presenta resultados estadísticamente diferentes (y supuestamente más beneficiosos) de los del grupo de control o placebo. Para ello, sólo tenemos que comprobar si el efecto nulo se encuentra dentro del intervalo de confianza. Si dicho valor está dentro del intervalo en torno al tamaño del efecto medio, tendremos que concluir que el meta-análisis no ha sido capaz de demostrar que el tratamiento experimental presenta un resultado estadísticamente diferente del control. Si por el contrario, el efecto nulo no se encuentra dentro del intervalo y el tamaño del efecto medio obtenido en el meta-análisis es favorable al tratamiento experimental, entonces podremos afirmar que dicho tratamiento ofrece mejores resultados, desde un punto de vista estadístico, que el grupo de control.

Por otra parte, los límites superior e inferior del intervalo de confianza, así como su amplitud, nos informan del grado de precisión de la estimación del efecto medio en la población. Cuanto más estrecho sea el intervalo de

confianza obtenido, tanto mayor será la precisión de la estimación del efecto en la población. Por regla general, el intervalo de confianza que se obtiene desde el modelo de efectos aleatorios es más ancho que el que se obtiene con el modelo de efectos fijos para unos mismos datos. Ello se debe a que el modelo de efectos aleatorios incorpora en el factor de ponderación más variabilidad (varianza intra y entre-estudios) que el modelo de efectos fijos.

Sea cual sea la estimación del tamaño del efecto medio, su significación estadística y el intervalo de confianza, la interpretación del alcance de estos resultados es diferente según el modelo asumido. Recuérdese que si hemos partido de un modelo de efectos fijos, sólo podremos generalizar los resultados del meta-análisis a una población de estudios de características idénticas a las de los estudios incluidos. Por el contrario, si el modelo asumido ha sido el de efectos aleatorios, los resultados pueden generalizarse a una población más amplia de estudios con características similares, aunque no necesariamente idénticas. El problema del modelo de efectos aleatorios es que, si bien garantiza mayor representatividad de sus resultados, no está claro hasta qué punto pueden ser diferentes entre si los estudios a los que se puede generalizar. Ésta es una cuestión conceptual, no analítica, que debe ser abordada desde la teoría o la praxis sobre la que se pretenden aplicar los resultados del meta-análisis.

2.3 ¿Existe homogeneidad en torno al tamaño del efecto medio?

Otra pregunta crucial en un meta-análisis consiste en valorar si los tamaños del efecto integrados en el meta-análisis son homogéneos entre sí, es decir, difieren sólo por error de muestreo aleatorio, o si, por el contrario, existe más heterogeneidad entre ellos de la que puede explicarse por el puro azar. Esta pregunta debe responderse tanto si se ha asumido un modelo de efectos fijos como de efectos aleatorios.

En el modelo de efectos fijos, la hipótesis de homogeneidad entre los tamaños del efecto se contrasta planteando como hipótesis nula que los tamaños del efecto paramétricos que estiman los efectos de los estudios son idénticos, es decir, es el mismo:

$$H_0$$
: $\theta_1 = \theta_2 = \dots = \theta_k = \theta$ (13)

En el *modelo de efectos aleatorios*, esta misma hipótesis se formula de forma diferente, aunque realmente representa la misma idea de que no existe varianza entre-estudios:

$$H_0: \tau^2 = 0 \tag{14}$$

En ambos casos para contrastar la hipótesis nula de homogeneidad de los tamaños del efecto se utiliza la prueba de homogeneidad, empleando el estadístico Q que se formuló en la ecuación (5). Bajo esta hipótesis, el estadístico Q se distribuye según una distribución χ^2 de Pearson con k-1 grados de libertad, por lo que es posible tomar una decisión estadística acerca de la hipótesis de homogeneidad asumiendo un determinado nivel de significación.

Si la hipótesis de homogeneidad no se rechaza, entonces puede concluirse que los tamaños del efecto individuales no son más heterogéneos entre sí de lo que cabría esperar por el azar o, lo que es lo mismo, por el error de muestreo. En ese caso, el tamaño del efecto medio calculado en el meta-análisis representa adecuadamente al conjunto de los estudios meta-analizados y, por tanto, puede utilizarse como un resumen numérico apropiado de dicho meta-análisis. Si, por el contrario, el resultado del estadístico Q nos lleva al rechazo de la hipótesis de homogeneidad, entonces tendremos que concluir que los tamaños del efecto de los estudios no son homogéneos entre sí, por lo que deben existir características diferenciales de los estudios que deben estar provocando tal heterogeneidad. Estas características diferenciales de los estudios suelen denominarse variables moderadoras, de confusión, o covariables.

Los procedimientos analíticos que pueden utilizarse para este propósito (por ejemplo, la meta-regresión) son objeto del siguiente capítulo, por lo que no serán tratados aquí. Pero es importante hacer notar que tanto si se ha aplicado un modelo de efectos fijos como de efectos aleatorios, un resultado significativo del estadístico Q de heterogeneidad justificaría el análisis de variables moderadoras que puedan explicar tal heterogeneidad.

Por último, Higgins y Thompson^(7,8) han propuesto un nuevo índice de heterogeneidad denominado l^2 , que permite complementar la información aportada por el estadístico Q. Mientras que el estadístico Q sólo nos informa de si existe o no heterogeneidad, el índice l^2 permite estimar el grado de heterogeneidad de los tamaños del efecto en un meta-análisis. El índice l^2 se obtiene mediante:

$$I^{2} = \begin{cases} \frac{Q - (k - 1)}{Q} \times 100 & (paraQ \ge k - 1) \\ 0 & (paraQ < k - 1) \end{cases}$$
 (15)

El índice l^2 puede interpretarse fácilmente en términos porcentuales, en concreto, como la parte de la variabilidad total de los tamaños del efecto que se debe a la variabilidad entre-estudios, no al error de muestreo aleatorio. Además, no se deja afectar por el número de estudios, k, mientras que el estadístico Q sí que se ve afectado⁹. Higgins y Thompson^(7,8) proponen interpretar el valor de l^2 siguiendo una clasificación normativa en términos porcentuales de 25%, 50% y 75% como heterogeneidad baja, media y alta, respectivamente.

3. Ejemplo

Para ilustrar los cálculos y la interpretación de los procedimientos analíticos desarrollados, se presentan los resultados de un meta-análisis sobre la eficacia de las intervenciones psicológicas, basadas en los modelos terapéuticos conductual y cognitivo-conductual, para combatir el tabaquismo⁽¹⁰⁾. En este meta-análisis se incluyeron 13 estudios realizados en España entre los años 1980 y 1996 que comparaban un grupo de control (no tratamiento del tabaquismo) con un grupo experimental consistente en la aplicación de alguna intervención conductual o cognitivo-conductual.

En todos los estudios el resultado de eficacia se midió de forma dicotómica mediante el registro de cuántos pacientes habían dejado de fumar v cuántos continuaban fumando, en cada grupo, tras un determinado período de seguimiento. A partir de cada estudio fue posible recomponer una tabla de contingencia 2×2 en la que se registraron estas frecuencias para los dos grupos implicados y, sobre los datos de cada tabla, se calculó la razón de odds, OR, su transformación logarítmica, log(OR), y la varianza de dicha transformación V(log(OR)). La Tabla 1 recoge todos estos datos. Los OR se calcularon de forma que valores por debajo de 1 indican un mejor resultado en el grupo experimental que en el de control, mientras que valores de OR mayores que 1 indicaron un mejor resultado para el grupo de control. Como consecuencia de este cálculo, la transformación logarítmica de los OR dio lugar a valores negativos cuando el grupo experimental logró mejores resultados, quedando los valores positivos para cuando el grupo de control alcanzó un mejor resultado. Como puede observarse en la Tabla 1, todos los estudios obtuvieron valores de OR favorables al tratamiento.

Tabla 1. Datos extraídos para el meta-análisis sobre la eficacia de las intervenciones psicológicas para combatir el tabaquismo(10).

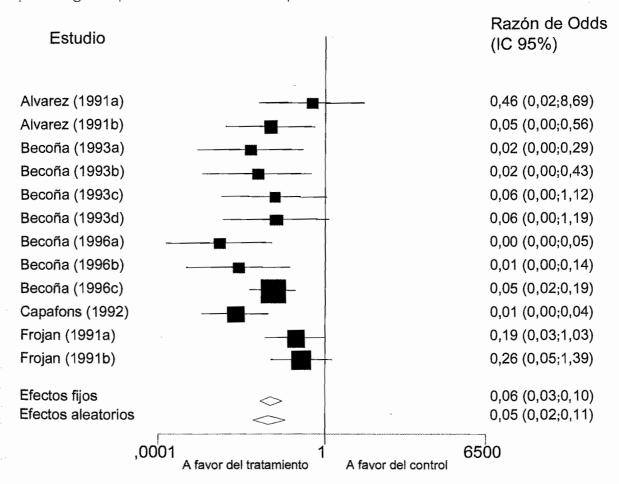
	Resultados (tablas 2×2)				Tamaño del efecto		
Estudio (año)	а	b	C	d ·	OR	log(OR)	V(log(OR))
Alvarez (1991a)	6	1	13	1	0.46	-0.7732	1.4979
Alvarez (1991b)	3	5	13	1	0,05	-3.0758	1.2690
Becoña (1993a)	8	16	16	0	0.02	-4.1598	1.4963
Becoña (1993b)	8	11	16	0	0.02	-3.7988	1.5051
Becoña (1993c)	14	7	16	0	0.06	-2.8373	1.5043
Becoña (1993d)	15	7	16	0	0.06	-2.7706	1.5028
Becoña (1996a)	7	33	40	0	0.00	-5.8911	1.4791
Becoña (1996b)	14	22	40	0	0.01	-4.8338	1.4622
Becoña (1996c)	30	51	33	3	0.05	-2.9285	0.6454
Capa- fons (1992)	2	36	25	3	0.01	-5.0106	0.9493
Frojan (1991a)	5	8	10	3	0.19	-1.6740	0.8708
Frojan (1991b)	6	7	10	3	0.26	-1.3581	0.8619

a: Número de eventos en el grupo experimental. b: Número de no eventos en el grupo experimental. c: Número de eventos en el grupo control.

d: Número de no eventos en el grupo control.

Un primer análisis que aporta una perspectiva general de los tamaños del efecto del meta-análisis consiste en el forest plot como el que se muestra en la Figura 1. En dicho gráfico el tamaño del efecto individual de cada estudio se presenta de forma numérica y gráfica junto a su intervalo de confianza. Por ejemplo, el estudio de Álvarez (1991a) obtuvo un OR=0,462, favorable al tratamiento, pero sin alcanzar la significación estadística, porque su intervalo de confianza (IC95%) incluye el efecto nulo, OR=1, (IC95% = [0,025; 8,693]). Por el contrario, el estudio de Álvarez (1991b) muestra un OR=0,05 favorable al tratamiento y estadísticamente significativo (IC95% = [0,004; 0,555]).

Figura 1. Meta-análisis para evaluar la eficacia de las intervenciones psicológicas para combatir el tabaquismo⁽¹⁰⁾.



En la parte inferior del forest plot se presenta el OR medio obtenido en el meta-análisis asumiendo un modelo de efectos fijos (OR=0,056) y un modelo de efectos aleatorios (OR=0,050), aplicando las ecuaciones (6) y (8), respectivamente, adaptadas al cálculo de estos índices del tamaño del efecto.

En torno al tamaño del efecto medio se construye un intervalo de confianza mediante las ecuaciones (9) y (10), para ambos modelos, respectivamente Los intervalos de confianza del 95% en torno al *OR* medio de ambos modelos resultan estadísticamente significativos al no incluir el efecto nulo *OR=1* (*IC95%=[0,031; 0,101]* desde el modelo de efectos fijos y *IC95%=[0,023; 0,111]* desde el modelo de efectos aleatorios). Los resultados con ambos modelos son coincidentes en este meta-análisis hacia la evidencia de un efecto consistente a favor de la eficacia de las técnicas conductuales y cognitivo-conductuales para dejar de fumar.

Resulta de interés comparar la amplitud de los intervalos de confianza para los dos modelos. Por regla general, el intervalo de confianza para el modelo de efectos aleatorios es más ancho que para el modelo de efectos fijos. En este ejemplo, se observa que la amplitud del intervalo para el OR medio del modelo de efectos aleatorios (0,088) es mayor que para el del modelo de efectos fijos (0,070). Ello se debe a que el modelo de efectos aleatorios tiene en cuenta dos fuentes de variabilidad, la intra y entre-estudios, mientras que el de efectos fijos únicamente toma en consideración la variabilidad intra-estudio, como se hace patente al examinar los factores de ponderación de uno y otro modelo, según se vio en la ecuación (7). No obstante, en el ejemplo que nos ocupa no se observan grandes discrepancias en la amplitud de los dos intervalos de confianza.

Para comprobar si el tamaño del efecto medio representa adecuadamente al conjunto de los estudios incluidos en el meta-análisis, es preciso determinar si existe más heterogeneidad entre los tamaños del efecto que pueda explicar el error de muestreo. Para ello, aplicamos la prueba Q de heterogeneidad mediante la ecuación (5). En nuestro ejemplo el resultado de la prueba queda al límite de la significación para un nivel de significación habitual (α =0,05): Q=20,209; grados de libertad=12; p=0,063. En sentido estricto, este resultado debería llevarnos a la conclusión de que no hay evidencia suficiente como para afirmar que existe más heterogeneidad de la explicada por el azar y, por tanto, que los estudios presentan tamaños del efecto homogéneos en torno al efecto medio. Sin embargo, la prueba Q no tiene suficiente potencia estadística para detectar heterogeneidad cuando el número de estudios, k, es pequeño (11-13), como ocurre en este ejemplo. En consecuencia, la prueba Q no es muy fiable bajo estas condiciones de aplicación.

Puede ayudar a examinar la presencia de heterogeneidad el cálculo de una estimación de la varianza entre-estudios, τ^2 , mediante la ecuación (3). En el ejemplo, obtenemos una varianza entre-estudios $\hat{\tau}^2$, un valor que es superior

a 0, lo que indica que existe cierto grado de heterogeneidad entre-estudios, aunque la prueba Q no haya sido capaz de detectarla. Asimismo, podemos cuantificar el grado de heterogeneidad de forma porcentual calculando el índice l^2 mediante la ecuación (15), obteniendo en el ejemplo un valor l^2 =40,62% que, según la clasificación de *Higgins y Thompson* puede calificarse de magnitud media. Estos resultados serían suficientes para justificar la búsqueda de posibles variables moderadoras de los tamaños del efecto obtenidos en los estudios, variables que podrían estar relacionadas con las características de las técnicas terapéuticas utilizadas, de las muestras de pacientes, etc. Para ello, se aplicarían técnicas de análisis de varianza y de meta-regresión tal y como se expone en el siguiente capítulo.

4. ¿Qué modelo estadístico elegir?

El problema de elegir el modelo estadístico a aplicar en un meta-análisis radica en que siempre desconocemos los parámetros poblaciones que pretendemos estimar, por lo que no podemos saber a ciencia cierta qué modelo se ajusta mejor a los datos meta-análiticos. La elección del modelo estadístico a aplicar en un meta-análisis debería ser una cuestión conceptual, no analítica. Lo ideal sería que el investigador decidiera *a priori* qué modelo desea aplicar, en función del grado de generalización que pretende dar a sus resultados, del grado de heterogeneidad que espera encontrar en los tamaños del efecto individuales y del contexto práctico al que se va a extrapolar los resultados del meta-análisis. No obstante, en ocasiones puede ser difícil tomar a priori tal decisión.

Una estrategia de compromiso que a veces se utiliza consiste en aplicar primero la prueba Q de heterogeneidad y decantarse por el modelo de efectos fijos si no resulta significativa, y por el modelo de efectos aleatorios si lo es. No obstante, esta estrategia no es aconsejable, por dos motivos principalmente. En primer lugar, porque se ha demostrado, mediante estudios de simulación, que la prueba Q posee escasa potencia estadística, en especial cuando el número de estudios del meta-análisis es pequeño (11-13). Por tanto, un resultado no significativo de la prueba Q puede no ser garantía suficiente para decantarse por el modelo de efectos fijos. En segundo lugar, el hecho de que la prueba Q sea significativa no tiene por qué conducirnos necesariamente a la adopción de un modelo de efectos aleatorios, ya que se podría mantener un modelo de efectos fijos y poner a prueba el posible influjo de variables moderadoras para tratar de explicar la heterogeneidad.

Una solución más razonable consistiría en elegir el modelo estadístico en

función del valor de la varianza entre-estudios, $\hat{ au}^2$. Dado que el estimado de dicha varianza se trunca en el valor O para evitar estimaciones negativas se puede asumir un modelo de efectos fijos cuando se dé tal situación, y adoptar un modelo de efectos aleatorios cuando la estimación de la varianza entre-estudios sea mayor que 0. De hecho, si la estimación de la varianza entre-estudios es 0 da igual asumir analíticamente uno u otro modelo, va que el factor de ponderación de los tamaños del efecto individuales será el mismo para ambos modelos. Dicho de otra forma, puede considerarse el modelo de efectos fijos como un caso particular del modelo, más general, de efectos aleatorios, cuando no existe varianza entre-estudios. Esto. unido al hecho de que la capacidad de generalización de los resultados es mayor en efectos aleatorios que en efectos fijos, hace que actualmente se considere más plausible el modelo de efectos aleatorios, ya que es más difícil imaginar que en la realidad toda una población de estudios realizados sobre un mismo problema esté estimando exactamente a un mismo efecto poblacional. Pero ante el meta-análisis de una revisión sistemática concreta nunca puede saberse con certeza si el modelo más apropiado es uno u otro, va que desconocemos los parámetros poblacionales.

Por último, cuando no es posible anticipar qué modelo estadístico es el que mejor se adecua a los datos meta-analíticos, puede llevarse a cabo un análisis de sensibilidad, consistente en aplicar los dos modelos y comprobar si los resultados son coincidentes, en términos de tamaño del efecto medio, significación estadística e intervalo de confianza. En tal caso, los resultados del meta-análisis serán más fiables; en caso contrario, sería preciso realizar análisis de sensibilidad más detallados, examinando qué tamaños del efecto son los que pueden estar provocando las divergencias en los resultados y cuáles pueden ser las causas de tales divergencias.

BIBLIOGRAFÍA

- 1. Glass GV, McGaw B, Smith ML. Meta-analysis in Social Research. Sage: Newbury Park, 1981.
- 2. Erez A, Bloom MC, Wells MT. Using random rather than fixed effects models in meta-analysis: Implications for situational specificity and validity generalization. Personnel Psychology 1996;49:275-306.
- 3. Hedges LV, Vevea JL. Fixed- and random-effects models in meta-analysis. Psychological Methods 1998;3:486-504.
- 4. Hunter JE, Schmidt FL. Fixed effects vs random effects meta-analysis models: Implications for cumulative research knowledge. International Journal of Selection & Assessment 2000;8:275-92.
- 5. Overton RC. A comparison of fixed-effects and mixed (random-effects) models for meta-analysis tests of moderator variable effects. Psychological Methods 1998;3:354-79.
- 6. DerSimonian R, Laird N. Meta-analysis in clinical trials. Contr Clin Trials 1986;7:177-88.
- 7. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002;21:1539-58.
- 8. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. Br Med J 2003;327:557-60.
- 9. Huedo-Medina T, Sánchez-Meca J, Marín-Martínez F, Botella J. Assessing heterogeneity in meta-analysis: Q statistic or I² index? Psychological Methods (en prensa).
- 10. Sánchez Meca J, Marín Martínez F, Olivares J, Rosa Al. Variables influyentes en el tratamiento de la adicción al tabaco: Un estudio de las tasas de abstinencia en España. Psicología Conductual 1999;7:301-21.
- 11. Cornwell JM. Monte Carlo comparison of three tests for homogeneity of independent correlations. Educational & Psychological Measurement 1993;53:605-18.

- 12. Harwell M. An empirical Study of Hedge's homogeneity test. Psychological Methods 1997;2:219-31.
- 13. Sánchez-Meca J, Marín-Martínez F. Homogeneity tests in meta-analysis: A Monte Carlo comparison of statistical power and Type I error. Quality & Quantity 1997;31:385-99.



EVISIONES SISTEMÁTICAS

en las Ciencias

A Contract of Salud a través de la sintesis de la Evidencia Científica de la vida

Coordinadores:

José Luis R. Martin · Aurelio Tobías Garcés · Teresa Seoane Pillado



■ Créditos

© Revisiones Sistemáticas en las Ciencias de la Vida. El concepto de Salud a través de la síntesis de la Evidencia Científica.

Edita: FISCAM - Fundación para la Investigación Sanitaria en Castilla-LaMancha

ISBN: 84-7788-420-X

Depósito Legal: TO-840-2006

Diseño e impresión: Rail Comunicación

Ninguna parte de esta obra, incluido el diseño de la cubierta puede ser reproducido, transmitido en ninguna forma o medio alguno, electrónico o mecánico, incluyendo fotocopias, grabaciones o cualquier sistema de recuperación de almacenaje de información sin obtener el permiso por escrito del editor y del coordinador.