

# Capítulo VI. Revisiones sistemáticas: introducción al meta-análisis

Dolores Frías Navarro  
Héctor Monterde i Bort  
Universidad de Valencia

---

## Citar este capítulo de libro como:

Frías-Navarro, D. y Monterde-i-Bort, H. (2014). Revisiones sistemáticas: introducción al meta-análisis. En Frías-Navarro, D., Pascual-Soler, M., Badenes-Ribera, L. y Monterde-i-Bort, H., *Reforma estadística en Psicología* (Capítulo VI). Valencia: Palmero Ediciones.  
(Edita Palmero Ediciones: trabajo@copias y revelados.com)

---

**E**l movimiento de la Medicina Basada en la Evidencia (MBE) y, en general, el de la Práctica Basada en la Evidencia (PBE) exige que el personal investigador o el profesional valoren los resultados de forma crítica (Frías-Navarro y Pascual-Llobel, 2003; Pascual-Llobel, Frías-Navarro y Monterde-i-Bort, 2004). Esta valoración crítica supone cambiar las opiniones de los expertos por el uso del método científico y por la valoración de la calidad de las evidencias halladas (Frías-Navarro, 2011). En consecuencia, la opinión del personal experto basada en su propia experiencia ha dejado de ser una fuente de información de calidad.

Sin embargo, como señalan Sanchez-Meca y Botella (2010), la acumulación de información sobre un mismo problema puede dificultar la puesta en práctica del enfoque de la PBE, al hacer inviable la selección de los estudios relevantes y su lectura crítica en un tiempo lo suficientemente corto como para que los profesionales puedan afrontar esta tarea.

Por ello, las revisiones sistemáticas se consideran fuentes valiosas de información dentro del modelo de la PBE (Frías-Navarro y Pascual-Llobell, 2003; Pascual-Llobell, et al., 2004;

Sánchez-Meca, 2010; Sánchez-Meca, Boruch, Petrosino, y Rosa-Alcázar, 2002). En este contexto los estudios de revisiones sistemáticas tipo meta-análisis han cobrado una alta relevancia y presencia en las revistas más prestigiosas (Borenstein, Hedges, Higgins & Rothstein, 2009). Además, la comunidad científica ha aceptado a los Meta-análisis como la metodología que permite ofrecer las mejores evidencias sobre un problema, cuando los estudios empíricos acumulados son estudios experimentales (Sánchez-Meca y Botella, 2010).

## **Qué es una revisión sistemática**

Una revisión sistemática de la literatura implica identificar, seleccionar, evaluar e interpretar toda la información publicada y no publicada sobre una determinada cuestión de investigación. Su principal objetivo es recopilar y sintetizar de forma transparente todas las publicaciones sobre un determinado fenómeno o temática concreta.

En consecuencia, las revisiones sistemáticas son estudios secundarios que resumen la información científica disponible a través de métodos rigurosos y explícitos. Pueden ser de tipo cualitativo (resumen descriptivo) o cuantitativo (cuando se utilizan métodos estadísticos para integrar cuantitativamente los resultados de los estudios empíricos primarios).

El desarrollo de una revisión sistemática requiere el cumplimiento de las normas propias del método científico, esto es, objetividad, sistematización y replicabilidad de sus hallazgos en cada una de sus etapas. (Marín-Martínez, et al., 2009; Perestelo-Pérez, 2013; Sánchez-Meca, 2008, 2010; Sánchez-Meca y Botella, 2010; Sánchez-Meca et al., 2011; Stroup et al., 2008).

## **Revisión sistemática versus revisión narrativa**

A diferencia de una revisión narrativa (también llamada revisión tradicional, cualitativa o subjetiva), una revisión sistemática implica un proceso de investigación científico basado en el rigor y la transparencia de cada una de las decisiones adoptadas durante la elaboración de dicha revisión, tanto si es una revisión sistemática cualitativa como si se trata de una revisión sistemática cuantitativa tipo meta-análisis. Es decir, una revisión sistemática aplica el método científico en todo el proceso de revisión.

Por lo tanto, las revisiones sistemáticas difieren de las revisiones narrativas en una serie de características:

**1) Protocolo previo de revisión.** En una revisión sistemática existe un protocolo de revisión previo que facilita la transparencia de todo el proceso de revisión.

**2) Selección objetiva de los estudios.** Los artículos o trabajos de investigación que formarán la unidad de análisis de las revisiones sistemáticas son seleccionados con criterios de búsqueda objetiva, sistemática y replicable de acuerdo a unas directrices previamente especificadas en el protocolo de revisión en función de la calidad o validez de las pruebas o evidencia aportada por los estudios científicos. Generalmente los trabajos de investigación se obtienen a través de bases de datos especializadas en la temática que son detalladas en la revisión junto con las palabras clave que se utilizan en las búsquedas y los años de la revisión. Además, la búsqueda de los trabajos se completa con la revisión de las referencias bibliográficas utilizadas en los documentos seleccionados y cualquier otro procedimiento que permita obtener información no registrada en las bases de datos o no publicada (tesis doctorales, trabajos presentados a congresos...).

**3) Criterios de inclusión y exclusión de los estudios primarios.** La información de los artículos y trabajos de investigación seleccionados es extraída de forma sistemática de acuerdo a unos criterios previos de inclusión y exclusión detallados en el protocolo de revisión y son interpretados en el contexto de todos los estudios revisados.

**4) Identificación de los estudios primarios.** Los trabajos seleccionados finalmente para formar parte de la revisión sistemática (estudios primarios) son identificados de forma clara en la revisión sistemática, facilitando la replicación de los estudios. Generalmente dichos trabajos se añaden a las referencias bibliográficas del trabajo de revisión con un asterisco.

En definitiva, a diferencia de las revisiones narrativas, las revisiones sistemáticas no se basan en la selección parcial de la literatura ni utilizan criterios subjetivos de selección. La calidad de las revisiones sistemáticas está relacionada con la captación de todos los trabajos de investigación (publicados y no publicados) sobre una determinada temática. Dichos trabajos son valorados previamente desde el punto de vista de la validez de sus hallazgos para posteriormente formar parte o no de la revisión (criterios de inclusión y exclusión) donde se integran todos los hallazgos o evidencias ofreciendo un resumen cualitativo sobre la situación de la temática analizada (revisión sistemática) o un resumen cuantitativo a través del cómputo del tamaño del efecto medio (meta-análisis). Por todo ello, las revisiones sistemáticas se consideran fuentes valiosas de información dentro del modelo de la Práctica Basada en la

Evidencia (Frías-Navarro y Pascual-Llobell, 2003; Pascual-Llobell, Frías-Navarro, y Monterde-i-Bort, 2004).

## **Proceso de revisión sistemática**

La Cochrane Collaboration<sup>1</sup>, la Campbell Collaboration<sup>2</sup>, y la PRISMA Statement<sup>3</sup> señalan unas consideraciones metodológicas a tener en cuenta y los pasos a seguir a la hora de elaborar revisiones sistemáticas.

Estos organismos recomiendan la elaboración de un Protocolo de revisión previo al proceso de investigación que guíe el mismo de manera explícita y replicable. (En el mismo sentido Wright et al., 2007).

### **Protocolo de revisión**

La identificación de la evidencia científica (estudios primarios) requiere un proceso sistemático en el que se definan a priori en el protocolo cuáles serán los criterios de inclusión y exclusión de los estudios primarios, se diseñe la estrategia de localización de los estudios y se establezcan los criterios de calidad metodológica de los estudios primarios.

El principal objetivo de un protocolo es prevenir el sesgo que podría ser introducido con cambios importantes en la cuestión de investigación original como por ejemplo el sesgo en el proceso de selección de los estudios primarios o en el análisis de los datos. Gracias a la elaboración del protocolo si se detectan problemas se podrán corregir antes de iniciar el estudio propiamente dicho.

Además, con el protocolo de revisión se facilita la transparencia del proceso de investigación, se evitan las decisiones a posteriori adoptadas en función de los resultados y también se facilita la valoración externa del proceso de revisión permitiendo detectar errores u omisiones antes de finalizar el trabajo de revisión sistemática. Además permite controlar la validez de los resultados y garantizar el rigor ético de la investigación.

La elaboración del protocolo de revisión debe responder a las siguientes cuestiones:

- Por qué es necesaria una revisión sistemática y cuáles son sus objetivos

---

<sup>1</sup> Cochrane Collaboration ([www.cochrane.org](http://www.cochrane.org); [www.cochrane.es](http://www.cochrane.es))

<sup>2</sup> Campbell Collaboration ([www.campbellcollaboration.org](http://www.campbellcollaboration.org)).

<sup>3</sup> PRISMA Statement ([www.prisma-statement.org](http://www.prisma-statement.org)).

- Qué tipo de cuestión o pregunta de investigación será analizada
- Cómo se desarrollará la revisión sistemática

Así pues, en el protocolo de revisión se diseña el plan de actuación que se seguirá posteriormente en la revisión sistemática, incluyendo los argumentos que apoyen el desarrollo la revisión, sus objetivos y los métodos que se utilizarán para localizar, seleccionar y valorar críticamente los estudios.

Si se trata de una revisión sistemática cuantitativa tipo meta-análisis, el protocolo de revisión debe incluir información detallada sobre el método de análisis estadístico como por ejemplo el tipo de modelo utilizado para resumir la evidencia cuantitativa (efectos fijos/efectos aleatorios), el método de combinación de los datos o el método de estimación del tamaño del efecto y su varianza cuando las publicaciones primarias no lo proporcionan de forma directa.

### **Etapas de una revisión sistemática**

Las principales etapas de una revisión sistemática son las siguientes (Cumming, 2012; Kline, 2013; Marín-Martínez, et al., 2009; Perestelo-Pérez, 2013; Sánchez-Meca, 2008, 2010; Sánchez-Meca y Botella, 2010; Sánchez-Meca et al., 2011; Stroup et al., 2008; Wright et al., 2007):

1. Formulación de la cuestión de investigación e identificar la necesidad de una revisión:
2. Búsqueda de la literatura.
3. Selección de los estudios que formarán parte de la revisión sistemática.
4. Valoración la calidad del estudio o publicación primario que ha sido incluido en la revisión sistemática.
5. Extracción de los datos de las publicaciones.
6. Sintetizar la evidencia: resumir los datos de las publicaciones seleccionadas. Si fuese necesario, elaborar un resumen cuantitativo de los datos (trabajo de meta-análisis).
7. Elaborar el informe, evaluando e interpretando los hallazgos.
8. Publicación de la revisión sistemática.

### **A. Formulación de la pregunta de investigación**

La formulación de la pregunta de investigación supone plantear el problema objeto de estudio de forma clara y precisa.

Una pregunta bien formulada responde al acrónimo PICOS, es decir, Participantes de estudio, la Intervención/tratamiento/exposición de los participantes; las Comparaciones, las medidas de Resultados de la intervención y el tipo de Estudio (el lector interesado puede encontrar más información en Perestelo-Pérez, 2013).

Hay que tener en cuenta que cuanto más restringida sea una pregunta de investigación, más específica y focalizada será la revisión sistemática, con lo que se podría tener dificultad en encontrar estudios empíricos primarios que respondan a la cuestión de investigación (Perestelo-Pérez, 2013; Wright et al., 2007).

### **B. Búsqueda de la literatura**

El objetivo es desarrollar una revisión lo más comprehensiva y exhaustiva posible. La búsqueda de la literatura se basa en los criterios de inclusión y exclusión de los estudios empíricos primarios establecidos en el protocolo de revisión.

Se recomienda utilizar distintas estrategias de búsqueda. Algunos autores distinguen entre procedimientos de búsqueda formales e informales (Marín-Martínez, et al., 2009; Sánchez-Meca, 2008, 2010; Sánchez-Meca y Botella, 2010; Sánchez-Meca et al., 2011).

Dentro de los **procedimientos formales** de búsqueda destacan:

*-búsquedas electrónicas en al menos dos bases de datos bibliográficas (Medline, PsycInfo, EMBASE, ERIC, WOS, etc.).*

*-búsqueda manual en revistas científicas especializadas, sobre todo los 6 meses anteriores a llevar a cabo la revisión sistemática.*

*-búsqueda manual en el listado de referencias de previas revisiones, estudios relevantes y de los estudios empíricos incluidos en la revisión.*

En los **procedimientos informales** de búsqueda destacan:

*-contacto con expertos e investigadores en el ámbito de investigación para la identificación de estudios que no han sido publicados.*

*-búsqueda manual en los libros de actas de congresos, Tesis Doctorales.*

No hay que olvidar que, aun cuando se utilicen distintas estrategias de búsqueda de la literatura es imposible recuperar todos los estudios empiricos primarios existentes en un determinado campo.

### **C. Selección de los estudios**

La selección de los estudios se debe llevar a cabo al menos por dos revisores de manera independiente. Los desacuerdos entre los revisores se pueden resolver por consenso o a través de un tercer revisor.

La selección de los estudios se desarrolla en dos fases: pre-selección y selección.

En la *fase de pre-selección* se identifican los estudios primarios que cumplen los criterios de inclusión a partir de la lectura de los *abstracts* y el título del estudio. Si hay dudas sobre la inclusión o no de un estudio, se procede a su lectura a texto completo.

En la *fase de selección*, los estudios pre-seleccionados se analizan a texto completo y se decide si se incluyen o no en la revisión sistemática. Se debe justificar la exclusión de los estudios.

Se debería comprobar la fiabilidad del proceso pres-selección y selección de los estudios primarios verificando el grado de acuerdo entre los revisores (Marín-Martínez, *et al.*, 2009; Sánchez-Meca, 2008, 2010; Sánchez-Meca y Botella, 2010; Sánchez-Meca *et al.*, 2011). La comprobación del grado de acuerdo se puede llevar a través de la obtención de índices de acuerdo como Kappa de Cohen para las variables cualitativas, o la correlación intra-clase para las variables continuas (Sánchez-Meca, *et al.*, 2011).

### **D. Valoración de la calidad metodológica de los estudios primarios**

El objetivo es controlar el posible sesgo de los estudios primarios que afectaría a la calidad de los resultados de la revisión sistemática. Un buena revisión sistemática debe incorporar algún sistema de valoración de la calidad de los estudios empíricos primarios, para comprobar su posterior relación con los resultados de la revisión sistemática (Kline, 2013; Perestelo-Pérez, 2013; Sánchez-Meca, 2010; Wright *et al.*, 2007).

La calidad metodológica de las publicaciones seleccionadas para formar parte de la revisión sistemática debe ser valorada de forma independiente al menos por dos revisores. Los desacuerdos en la valoración de la calidad metodológica pueden ser superados por consenso entre los dos revisores o por un tercer revisor.

No hay un único procedimiento para medir la validez o calidad de los estudios primarios de las investigaciones. Generalmente se utiliza una plantilla o un listado de revisión previamente elaborado en la fase de desarrollo del protocolo de la revisión sistemática. No obstante, también existen escalas o listados de comprobación ya elaborados para evaluar la validez y calidad de los estudios empíricos primarios experimentales, cuasi-experimentales y ex post facto; así como de estudios secundarios como las revisiones sistemáticas (v. gr. AMSTAR, Shea et al., 2007, Shea et al., 2007; Shea et al., 2009).

Se recomienda excluir de la revisión sistemática los estudios de baja calidad metodológica (Kline, 2013; Wright et al., 2007).

### ***E. Extracción de los datos de las publicaciones***

La extracción de los datos supone obtener información relevante de cada una de las publicaciones con el objetivo de elaborar la revisión sistemática. El tipo de información que será recogida de cada estudio primario debe ser definido previamente en el protocolo de la revisión sistemática, asegurando que los revisores extraerán toda la información necesaria de cada estudio de una manera uniforme. Por tanto, los datos a extraer no son iguales en todas las revisiones sistemáticas. Sin embargo, se suelen registrar: autor y año de publicación, tipo y características del diseño del estudio, número y características de los participantes, ámbito del estudio, medidas de resultado, fuentes de financiación, conflicto de intereses, etc. (Perestelo-Pérez, 2013).

Si el trabajo de revisión sistemática es cuantitativo (meta-análisis) entonces será necesario extraer datos numéricos de las publicaciones con el objetivo de computar los tamaños del efecto individuales y/o tasas de prevalencia que serán combinados en un tamaño del efecto medio y/o prevalencia media que aporta el trabajo de meta-análisis como resumen de la evidencia (Frías-Navarro, 2011).

En una revisión sistemática es muy importante evitar la inclusión de publicaciones que se han publicado varias veces utilizando los mismos datos. La inclusión de datos duplicados en un trabajo de meta-análisis podría conducir a un sesgo en la estimación del efecto medio o prevalencia media, sobrevalorando su tamaño. En estos supuestos se recomienda combinar en una sola estimación el tamaño del efecto (o prevalencia) e introducirse en el estudio de meta-análisis como un solo estudio. Esta solución rescata la información contenida en los estudios duplicados, mientras que al mismo tiempo, preserva la independencia estadística entre las entradas en el meta-análisis (Schmidt, 2008). Esta es la misma recomendación hecha por

Hunter y Schmidt (2004) para el caso de "la replicación conceptual" donde hay varias medidas del mismo constructo tomado en la misma muestra y en el mismo estudio. Estos autores recomiendan combinar estas medidas en una sola estimación del tamaño del efecto, de tal manera que los datos de la muestra se computan en el meta-análisis una sola vez, evitando así una violación de la hipótesis de independencia estadística.

Pero, ¿por qué la literatura contiene estudios duplicados? Probablemente debido a la presión por publicar, sobre todo entre los profesores de universidades sin plaza fija. Esta presión también ha tentado a algunos investigadores a ir más allá y publicar datos falsos (Schmidt, 2008).

Por otra parte, para que la extracción de los datos de los estudios empíricos primarios seleccionados sea objetiva y sistemática se requiere un control externo mediante una plantilla de revisión. Además, se recomienda que al menos dos revisores efectúan de forma independiente la extracción de los datos de los estudios (o de una muestra aleatoria de éstos). Los desacuerdos se pueden resolver por consenso o por la participación de un tercer revisor.

Se debería comprobar la fiabilidad del proceso de codificación verificando el grado de acuerdo entre los revisores (Marín-Martínez, et al, 2009; Sánchez Meca, 2010; Sánchez-Meca y Botella, 2010). La comprobación del grado de acuerdo se puede llevar a través de la obtención de índices de acuerdo como Kappa de Cohen para las variables cualitativas, o la correlación intra-clase para las variables continuas (Sánchez-Meca, et al, 2011).

#### ***F. Sintetizar la evidencia: meta-análisis***

El objetivo final de una revisión sistemática es ofrecer un resumen de los efectos detectados en los estudios primarios analizados. El resumen puede ser descriptivo o puede ofrecer una síntesis cuantitativa de los datos de las publicaciones cuando se trata de un trabajo de meta-análisis.

La gran mayoría de los meta-análisis se realizan computando el tamaño del efecto de diferencia estandarizada de medias, *odds ratio*, coeficientes de correlación y proporción de varianza explicada (Frías-Navarro, 2011b). Los estudios de meta-análisis también pueden ser útiles para computar una estimación más precisa de datos de frecuencia como proporciones de prevalencia (Barendregt, Doi, Lee, Norman, y Vos, 2013).

La estimación del tamaño del efecto medio o la prevalencia media en los estudios de meta-análisis puede realizarse mediante dos modelos estadísticos: el modelo de efectos fijos y el modelo de efectos aleatorios (Hedges y Vevea, 1998; Hunter, y Schmidt, 2000).

Conviene tener en cuenta la diferencia entre tamaño del efecto verdadero o prevalencia verdadera y el tamaño del efecto observado o prevalencia observada. El tamaño del efecto verdadero o prevalencia verdadera es el tamaño del efecto o la prevalencia en la población y es el tamaño del efecto o la prevalencia que se podría observar si el estudio tuviese un tamaño de muestra infinito y por lo tanto no existiese error muestral. El tamaño del efecto observado es el tamaño del efecto que se observa en la muestra de estudios. En consecuencia, en un estudio de meta-análisis se obtienen los tamaños del efecto observados o prevalencias observadas y se trata de estimar el efecto o prevalencia poblacional.

### ***Modelo de efecto fijo y modelo de efectos aleatorios***

El '*modelo de efecto fijo*' asume que hay un único tamaño del efecto verdadero o prevalencia verdadera (desconocido en la realidad) y los tamaños del efecto o prevalencias obtenidos con los estudios primarios proceden de una única población con un tamaño del efecto verdadero común o una prevalencia común, de ahí el término de efecto fijo. Es decir, se asume que los estudios incluidos en el trabajo de meta-análisis están estimando a un mismo y único tamaño del efecto poblacional ( $\theta$ ) o a una misma y única prevalencia poblacional. Si todos los estudios estiman al mismo efecto verdadero o la misma prevalencia verdadera entonces los tamaños del efecto observados o prevalencia observadas varían de estudio a estudio únicamente por el error de muestreo aleatorio inherente en cada estudio que se cuantifica estimando la varianza intra-estudio. Si cada estudio tuviese un tamaño de muestra infinito entonces el error muestral sería cero y el efecto o prevalencia de cada estudio sería el mismo que el del verdadero efecto o prevalencia verdadera. Por ello, las diferencias en los tamaños del efecto observados o en las prevalencia observadas en los estudios son debidas al error muestral o error provocado por utilizar muestras de participantes diferentes. En este sentido, todos los factores que podrían influir sobre el tamaño del efecto o prevalencia son los mismos en todos los estudios y el tamaño del efecto verdadero o la prevalencia verdadera es el mismo en todos los estudios. En la realidad el tamaño de la muestra de cada estudio no es infinito y por lo tanto hay error muestral y los efectos observados o las prevalencias observadas en el estudio no son iguales al efecto verdadero o a la prevalencia verdadera.

En el modelo de efecto fijo se asume que el tamaño del efecto verdadero o prevalencia verdadera es el mismo para todos los estudios pero esta asunción no siempre es posible. Cuando se ejecuta un trabajo de meta-análisis es necesario que la muestra de estudios sea homogénea sobre un determinado constructo o cuestión permitiendo combinar la información en un tamaño del efecto medio o prevalencia media pero generalmente no se asume que el tamaño del efecto verdadero o prevalencia verdadera sea exactamente el mismo en todos los estudios. En estas ocasiones es más adecuado plantear un trabajo de meta-análisis con un modelo de efectos aleatorios.

En el '*modelo de efectos aleatorios*' se asume que el efecto verdadero o prevalencia verdadera puede variar de estudio a estudio. Por ejemplo, el tamaño del efecto o la prevalencia podría ser más altos (o más bajos) en estudios con participantes adultos o con mayor nivel educativo. Es decir, podría haber diferentes tamaños del efecto o prevalencias poblacionales. En este caso se asume que los estudios primarios estiman a una distribución de tamaños del efecto o prevalencias en la población. Es decir, cada tamaño del efecto individual o prevalencia individual estima a un tamaño del efecto o prevalencia poblacional diferente. Por lo tanto, en el modelo de efectos aleatorios además de la variabilidad debida al error de muestreo o intra-estudio hay que estimar la variabilidad entre los estudios. Por ello, el modelo matemático tiene dos términos de error: uno debido a la variabilidad de cada estudio (variabilidad intra-estudio), que coincide con la variabilidad del modelo de efectos fijos, y otro relacionado con la variabilidad entre-estudios o desviación de cada estudio respecto del tamaño del efecto medio o prevalencia media. Si se pudiese realizar un número infinito de estudios los tamaños del efecto verdaderos o prevalencias verdaderas de dichos estudios se distribuirían en torno a una media. Por lo tanto, los tamaños del efecto o prevalencias de los estudios incluidos en el trabajo de meta-análisis representan una muestra aleatoria de esos tamaños del efecto o prevalencias y de ahí el término de efectos aleatorios. En el modelo de efecto fijo la generalización del tamaño del efecto medio o prevalencia media se limita a la población de estudios que tengan características muy similares a las incluidas en el trabajo de meta-análisis mientras que en el modelo de efectos aleatorios los resultados pueden generalizarse a una población mayor de posibles resultados o estudios con características similares aunque no necesariamente idénticas.

En un estudio de meta-análisis se obtienen los tamaños del efecto observados y se trata de estimar el efecto poblacional. Tanto en el modelo de efecto fijo como en el modelo de efectos aleatorios se pondera cada estimación del tamaño del efecto por la inversa de la varianza.

Para estimar de la manera más precisa el efecto población (es decir, minimizando la varianza) se computa una media ponderada de los tamaños del efecto observados o prevalencias observadas donde el peso asignado a cada estudio es la inversa de la varianza de dicho estudio. En el caso del modelo de efecto fijo se pondera por la inversa de la varianza intra-estudio. Cuanto menor sea la varianza más preciso es el tamaño del efecto estimado o prevalencia estimada y por lo tanto contribuye más en la estimación del efecto verdadero o prevalencia verdadera. En el modelo de efectos aleatorios se computa la varianza intra estudio más la varianza entre los estudios dado que se asume que los estudios estiman diferentes tamaños del efecto verdaderos o prevalencias verdaderas y por lo tanto los resultados de los estudios son combinados ponderando por la inversa de la varianza intra-estudio más la varianza entre los estudios.

En resumen, en el modelo de efecto fijo se asume que el tamaño del efecto verdadero o prevalencia verdadera es el mismo en todos los estudios y el tamaño del efecto medio o prevalencia media que se computa es una estimación de dicho tamaño del efecto común o prevalencia común y la única razón de que los tamaños del efecto o prevalencias varíen entre los estudios es por el error muestral (varianza intra-estudio). En cambio, cuando se ejecuta un modelo de efectos aleatorios se asume que el tamaño del efecto verdadero o prevalencia verdadera varía de un estudio a otro y los estudios incluidos en el trabajo de meta-análisis representan una muestra aleatoria de tamaños de efecto observados o prevalencias observadas. En el modelo de efectos aleatorios el objetivo no es estimar un único efecto verdadero o prevalencia verdadera sino estimar la media de una distribución de efectos o prevalencias (Figura 1 y Figura 2).

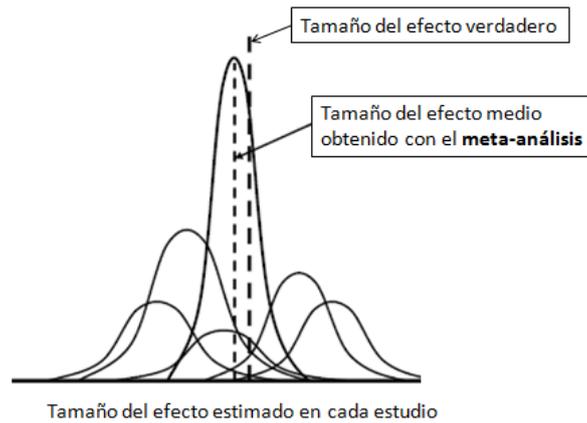


Figura 1. Modelo de efecto fijo

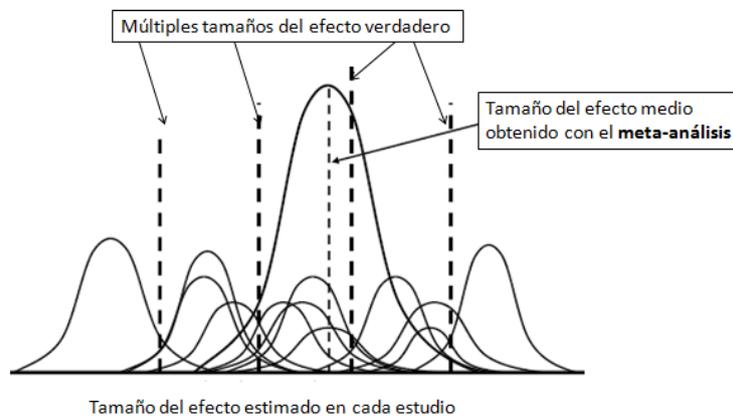


Figura 2. Modelo de efectos aleatorios

Decidir qué modelo estadístico es el más adecuado para realizar el trabajo de meta-análisis no es sencillo pues se desconocen los parámetros poblacionales que se pretenden estimar y por lo tanto no se conoce qué modelo se ajustará mejor a los datos. Se trata de una cuestión más conceptual que analítica donde el investigador debe valorar el grado de generalización que pretende otorgar a sus resultados y el grado de heterogeneidad que espera encontrar en los tamaños del efecto o prevalencias de los estudios primarios (Sánchez-Meca, Marín y Huedo, 2006).

En un trabajo de meta-análisis el tamaño del efecto medio o prevalencia media se acompaña de su intervalo de confianza. La amplitud entre el límite superior y el límite inferior del intervalo de confianza ofrece información sobre el grado de precisión de la estimación del efecto medio o prevalencia media en la población. Cuanto más estrecho es el intervalo de confianza, mayor es la precisión de la estimación puntual del efecto o de la prevalencia en la

población. Si se tiene en cuenta que en el modelo de efecto fijo la única fuente de incertidumbre es el error o varianza intra-estudio mientras que en el modelo de efectos aleatorios hay más variabilidad ya que se incorpora la varianza intra-estudio más la varianza entre los estudios entonces la varianza, el error estándar y el intervalo de confianza del tamaño del efecto medio o prevalencia media siempre será mayor (más amplio) en el modelo de efectos aleatorios para unos mismos datos. Exceptuando el caso donde la varianza entre los estudios fuese cero.

La estimación del tamaño del efecto medio o de la prevalencia media se acompaña de estadísticos que informan del grado de heterogeneidad estadística que se detecta entre los estudios o variación entre los resultados de los estudios más allá del error aleatorio (Song, Sheldon, Sutton, Abrams y Jones, 2001). La prueba Q de Cochran permite comprobar la heterogeneidad y su grado de significación estadística (Cochran, 1954). El estadístico Q tiene una distribución Ji Cuadrado con  $k-1$  grados de libertad, siendo  $k$  el número de estudios, facilitando una decisión estadística acerca de la hipótesis de homogeneidad de los tamaños del efecto computados o prevalencias computadas. Valores de  $Q > k-1$  sugieren heterogeneidad estadística. El estadístico Q se acompaña de un valor  $p$  de probabilidad que si es menor a 0.05 se sugiere presencia de heterogeneidad entre los tamaños del efecto o prevalencias. Otro estadístico de homogeneidad es el estadístico  $I^2$  que completa la información de la prueba Q e informa del porcentaje de variación que se detecta entre los tamaños del efecto o prevalencias de los estudios que componen el trabajo de meta-análisis (Takkouche, Cadarso-Sures y Spiegelman, 1999). El estadístico  $I^2$  puede interpretarse en términos de porcentaje indicando el grado de variabilidad total de los tamaños del efecto o prevalencias que se debe a la variabilidad entre los estudios y no al error de muestreo aleatorio. Una ventaja del estadístico  $I^2$  es que no está afectado por el tamaño de la muestra mientras que la prueba Q sí se ve afectada. Higgins y Thompson (2002) proponen interpretar el valor de  $I^2$  como heterogeneidad baja (25%), heterogeneidad media (50%) y heterogeneidad alta (75%).

Además, como ya se ha comentado, la descripción de los resultados se suele acompañar de una gráfica (*forest plot o diagrama de bosque*) donde se representan los tamaños del efecto o prevalencias de los estudios individuales junto con el tamaño del efecto medio computado o prevalencia media computada en el estudio de meta-análisis.

La figura 3 muestra un ejemplo de forest plot extraído de un reciente estudio de meta-análisis de prevalencia de la violencia de pareja en mujeres lesbianas auto-identificadas llevado a cabo

por Badenes-Ribera, Frías-Navarro, Bonilla-Campos, Pons-Salvador y Monterde-i-Bort (2014). En la parte izquierda de la imagen se detallan los estudios empíricos primarios que han formado parte del estudio de meta-análisis. En la parte derecha se detalla los resultados en prevalencia de cada uno de los estudios primarios y su peso en el estudio de meta-análisis. Los cuadrados de la representación gráfica muestran las prevalencias de cada uno de los estudios primarios, mientras que el diamante representa la prevalencia media (0,48 o 48%).

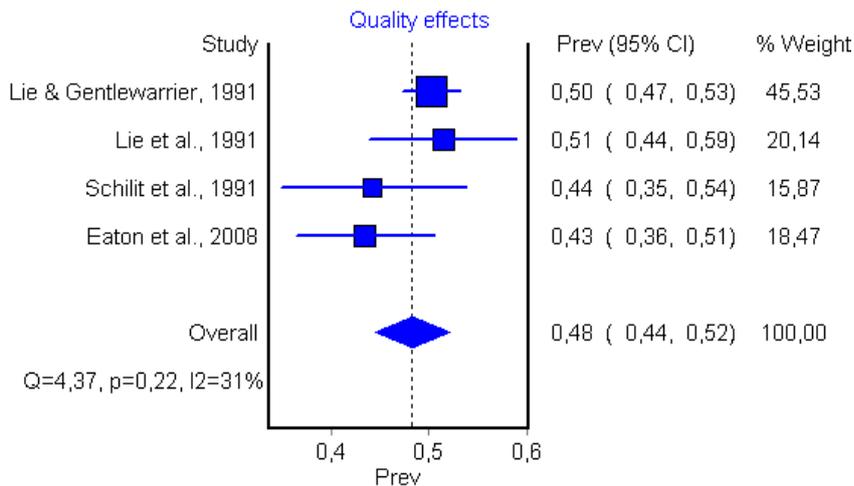


Figura 3.- Forest Plot: victimización en violencia de pareja

En este estudio para combinar estadísticamente los resultados de los estudios primarios se optó por un modelo de efectos aleatorios donde a priori no se asume homogeneidad en las magnitudes de las prevalencias, atribuyendo la variabilidad a error de muestreo y a la variabilidad entre los estudios. Es decir, se considera que cada estudio estima su propio tamaño del efecto poblacional (Hedges, 1994; Hedges & Olkin, 1985). Además este modelo se recomienda cuando el número de estudios es pequeño (Brockwell & Gordon, 2011). Para el cómputo de la prevalencia media se utilizó el programa MetaXL que incluye la posibilidad de ejecutar un modelo ponderando por la calidad de los efectos estimados de los estudios primarios, aspecto recomendable cuando se detecta heterogeneidad en los tamaños del efecto individuales (Barendregt, et al, 2013; Doi, & Thalib, 2008). El diamante del gráfico representa la prevalencia media.

## Valoración de la revisión sistemática

La elaboración de una revisión sistemática no garantiza de forma directa la calidad de sus hallazgos. No garantiza que se hayan incluido todos los trabajos sobre una determinada

cuestión, no garantiza que la validez de los estudios primarios haya sido valorada de forma apropiada ni tampoco garantiza que los datos hayan sido combinados adecuadamente en un trabajo de meta-análisis.

La valoración crítica de un estudio de revisión sistemática supone valorar la calidad de los estudios primarios incluidos en la revisión y también la calidad de la revisión sistemática como producto elaborado a partir de dichos estudios primarios.

En principio la calidad de los estudios primarios es valorada por los autores de la revisión quienes deben señalar en su informe la calidad científica de sus estudios primarios para posteriormente otorgar un determinado grado de evidencia a la revisión sistemática efectuada. Cuestiones relacionadas con el diseño de investigación de los estudios primarios y su validez y cuestiones vinculadas con el posible sesgo de las revisiones sistemáticas afectan a la calidad de las revisiones sistemáticas. Cuanto mayor sea la calidad de las pruebas aportadas por el estudio de revisión sistemática menos probabilidades habrá de que nuevos resultados de estudios posteriores afecten a los hallazgos detallados en la revisión sistemática. Dada la dificultad de las tareas de evaluación es necesario que en ciertas ocasiones al menos dos evaluadores externos lleven a cabo el proceso de valoración y toma de decisiones, calculándose el acuerdo entre-jueces.

Egger, Dickersin y Smith (2001) señalan los siguientes problemas o sesgos que podrían ocurrir en una revisión sistemática: sesgo de publicación, sesgo de retraso en la publicación, sesgo de publicación duplicada, sesgo de citación, sesgo del lenguaje y sesgo del informe de resultados. Realizar una revisión sistemática de la literatura sin tener en cuenta diferentes idiomas (sesgo del lenguaje) o los estudios no publicados (sesgo de publicación) puede sobreestimar el tamaño del efecto.

Cuando la publicación o no publicación de los hallazgos de una investigación depende de la naturaleza y la dirección de los resultados se trata del sesgo de publicación. La exclusión de ensayos clínicos no publicados podría introducir un sesgo en la revisión sistemática sobreestimando el efecto del tratamiento (Hopewell, McDonald, Clarke, y Egger, 2007). De ahí la importancia de localizar los estudios publicados y los no publicados cuando se efectúa una revisión sistemática.

Cuando la rápida o la lenta publicación de los resultados de la investigación dependen de la naturaleza y dirección de los hallazgos se trata del sesgo de retraso en la publicación. Hopewell et al., (2007) también comprobaron que los ensayos clínicos con resultados

estadísticamente significativos suelen publicarse un año o dos antes que los estudios con resultados no estadísticamente significativos.

Cuando la publicación múltiple o única de los hallazgos depende de la naturaleza o dirección de los resultados se trata del sesgo de duplicación de las publicaciones. En ocasiones algunos ensayos clínicos son publicados repetidamente e incluso con diferentes autores. El revisor debe valorar esas duplicidades ya que si los resultados de un mismo estudio son incorporados al estudio de meta-análisis entonces podría afectar al tamaño del efecto medio estimado, sesgando los resultados hacia los hallazgos de dicho estudio.

Si la citación o no citación de los hallazgos de la investigación depende de la naturaleza y la dirección de los resultados se trata del sesgo de citación. Los estudios con resultados estadísticamente significativos suelen citarse más que los estudios con resultados no estadísticamente significativos.

Si la publicación de los hallazgos en un sólo idioma depende de la naturaleza y dirección de los resultados se trata del sesgo del lenguaje.

Y si el informe selectivo de ciertos resultados y no de otros depende de la naturaleza y la dirección de los resultados se trata del sesgo del informe de resultados. Si el informe de investigación no informa de todos los resultados de los estudios o no informa con suficiente detalle entonces podría tratarse de un sesgo del informe. Por ejemplo, el informe podría estar sesgado si tiene en cuenta la significación estadística de los hallazgos, de manera que los resultados estadísticamente significativos son detallados pero los resultados estadísticamente no significativos no son informados o son informados de forma incompleta.

Así pues, el lector de una revisión sistemática deben realizar una lectura crítica de cada uno de los pasos de dicha revisión y también debe valorar si los resultados hallados puede ser útiles en su práctica profesional.

La valoración de la consistencia o no de los hallazgos obtenidos con una revisión sistemática es importante ya que en principio la revisión sistemática supone acumular información de diversas fuentes y de este modo mejorar la estimación de los efectos y su precisión. Si los resultados de las investigaciones no son consistentes entonces la revisión sistemática debe presentar una posible explicación de la heterogeneidad de los resultados de los estudios primarios.

También conviene reflexionar sobre los resultados que aporta la revisión sistemática. Si se trata de un trabajo de meta-análisis entonces se ofrecen datos del tamaño del efecto medio,

prevalencia media y sus intervalos de confianza, ayudando con ello a la interpretación de la precisión de la estimación puntual del tamaño del efecto medio o de la prevalencia media de un fenómeno.

Por último, el profesional o investigador debe valorar la aplicabilidad de los hallazgos de una revisión sistemática a su propio contexto de trabajo. De este modo, las actuaciones de los profesionales estarán siempre basadas en la mejor evidencia que exista sobre una cuestión, actuando desde los principios de la Práctica Basada en la Evidencia (Frías-Navarro y Pascual-Llobell, 2003; Pascual-Llobell, et al, 2004).

En definitiva las revisiones sistemáticas cualitativas y los trabajos de meta-análisis (revisiones sistemáticas cuantitativas) también deben ser sometidos a una lectura crítica que valore el rigor del proceso de investigación y la calidad de sus pruebas. Se ha comprobado que las deficiencias en la ejecución de los trabajos de meta-análisis podrían explicar las discrepancias encontradas en algunas áreas de investigación cuando se comparan sus resultados con ensayos controlados aleatorios que utilizan tamaños de muestra grandes (LeLorier, Gregoire, Benhaddad, Lapierre, y Derderian, 1997). Los autores y los lectores de las revisiones sistemáticas deben ser conscientes de los posibles sesgos que podrían cometerse en la revisión. Los autores de los revisiones deben ser cuidadosos para identificar todas las publicaciones sobre el tema objeto de estudio y deben valorar de forma crítica la calidad de los estudios primarios, facilitando un análisis de los efectos en función de la calidad metodológica del estudio. Los lectores de revisiones sistemáticas deben ser conscientes que es necesario realizar una lectura crítica ya que las revisiones sistemáticas no están exentas de sesgo. Teniendo en cuenta estas cuestiones, disponer de revisiones sistemáticas que incluyen las pruebas con mejor calidad supone trabajar con documentación válida, facilitando la toma de decisiones de los profesionales e investigadores (Cumming, 2012; Kline, 2013; Sánchez-Meca, et al., 2002).

---