

La revisión sistemática cuantitativa en enfermería

¹Rafael del Pino Casado

²Antonio Frías Osuna

³Pedro A. Palomino Moral

¹ Diplomado en Enfermería. Licenciado en Antropología y Doctor por la Universidad de Jaén. Profesor Contratado Doctor. Departamento de Enfermería. Universidad de Jaén.

² Diplomado en Enfermería. Licenciado en Antropología y Doctor por la Universidad de Granada. Profesor Titular de Universidad. Departamento de Enfermería. Universidad de Jaén.

³ Diplomado en Enfermería. Licenciado en Psicopedagogía y Doctor por la UNED. Profesor Titular de Universidad. Departamento de Enfermería. Universidad de Jaén.

Dirección de contacto: rdelpino@ujaen.es

Resumen

La finalidad del presente artículo es la de encuadrar el concepto de revisión sistemática cuantitativa, establecer su utilidad en el marco de la ciencia enfermera y proponer recomendaciones para su elaboración. Para alcanzar dicha finalidad, el artículo se estructura en tres partes: una primera de tipo conceptual, una segunda donde se describen los procedimientos para la elaboración de una revisión sistemática de tipo cuantitativo (con y sin metaanálisis) y una tercera en la que se incluyen aspectos relacionados con la evidencia científica aportada por las revisiones, a modo de reflexión final.

Palabras clave: revisión sistemática; meta-análisis; búsqueda bibliográfica; evidencia científica.

Abstract

Quantitative systematic reviews in nursing

The present paper is aimed at framing the concept of a quantitative systematic review, establish its usefulness for nursing science, and suggesting recommendations for making such reviews. The article is divided in three parts, as follows: a first part showing its conceptual basis; a second part describing procedures to perform a quantitative systematic review (with or without a meta-analysis); and a third part discussing some aspects related to scientific evidence provided by review, as a final reflection.

Key words: systematic review; meta-analysis; literature search; scientific evidence.

Introducción

La investigación, tanto cuantitativa como cualitativa, puede desarrollarse a dos niveles: 1) a nivel de datos originales, recogidos y analizados por el investigador (estudios originales o primarios), y 2) a nivel de datos de otras investigaciones ya realizadas (investigación secundaria).

Dentro de la investigación secundaria podemos distinguir entre el análisis secundario de datos y las revisiones. Los estudios de análisis secundario analizan de nuevo datos de otra investigación con una finalidad distinta a la original. Las revisiones analizan los estudios disponibles sobre un determinado tema con el fin de integrar sus resultados. En función de su grado de sistematización, exhaustividad y explicitación de los criterios de inclusión y análisis de los estudios, se pueden clasificar en revisiones narrativas (bajo grado) y revisiones sistemáticas (alto grado). En el presente artículo analizaremos las revisiones, ya que los estudios con análisis secundario exceden del interés del mismo. Dentro de las revisiones, analizaremos las revisiones sistemáticas, ya que la utilidad de las revisiones narrativas es bastante discutible, y dentro de las revisiones sistemáticas, nos centraremos en las que tiene como objeto de estudio la investigación original de tipo cuantitativo.

Así, la finalidad concreta del presente artículo será la de encuadrar el concepto de revisión sistemática cuantitativa, establecer su utilidad y proponer recomendaciones para su elaboración. Dichas recomendaciones están orientadas al lector que se inicia en el tema de las revisiones pero que dispone de conceptos básicos sobre metodología de investigación. La limitación de espacio hace que aspectos más especializados de la temática tratada en este artículo deban verse en obras más amplias y especializadas. No obstante, a lo largo del artículo serán frecuentes las referencias donde poder profundizar en los conceptos y procedimientos.

Para alcanzar el anterior fin, el artículo se estructura en tres partes: una primera de tipo conceptual, una segunda donde se describen los procedimientos para la elaboración de una revisión sistemática de tipo cuantitativo y una tercera en la que se incluyen aspectos relacionados con la evidencia científica aportada por las revisiones, a modo de reflexión final.

La revisión sistemática. Concepto

La revisión sistemática, como ya se ha comentado en la introducción, es un tipo de investigación secundaria que consiste en la búsqueda, localización, recuperación, análisis e integración del conocimiento científico sobre una determinada pregunta de investigación, utilizando para ello una aproximación sistemática que permite minimizar sesgos (1) y posibilitar la replicación de la revisión a cuantos investigadores les parezca oportuno. Dicha aproximación sistemática, que debe de ser documentada en el apartado de material y métodos del correspondiente informe de investigación (1), se materializa en el carácter exhaustivo de la búsqueda de los estudios existentes sobre el tema y en la explicitación de los criterios de búsqueda, inclusión y análisis de dichos estudios.

Las revisiones en general permiten aglutinar el conocimiento existente sobre una determinada cuestión, en una época en la que la proliferación y difusión de estudios en ciencias de la salud complica el análisis de lo que se publica sobre una determinada cuestión. La revisión sistemática en particular permite superar las limitaciones de las revisiones narrativas, limitaciones que son debidas al alto grado de subjetividad con el que se realizan, aportando una mayor transparencia al proceso de revisión. Además, la revisión sistemática posibilita la superación de las limitaciones de los estudios originales, que pocas veces son capaces de dar respuestas adecuadas y definitivas a las cuestiones que analizan, y que frecuentemente arrojan resultados contradictorios. Esta ventaja es particularmente importante en el caso de los estudios experimentales y cuasi-experimentales con grupo control, los cuales se realizan frecuentemente con muestras de conveniencia y donde el escaso tamaño muestral puede alterar la comparabilidad de los grupos.

La revisión sistemática puede o no incluir un metaanálisis. El metaanálisis es un procedimiento estadístico que permite integrar los resultados de los estudios analizados mediante la determinación de una medida promedio del tamaño del efecto de interés, es decir, de la magnitud de la asociación entre la exposición y el efecto o entre la intervención y el resultado. El metaanálisis aporta mayor validez interna y externa a la revisión sistemática, al permitir una valoración más objetiva y precisa del efecto de interés, posibilitar la cuantificación de la heterogeneidad de los resultados y de la varianza debida al error de muestreo, y reducir la probabilidad de falsos negativos al aumentar la potencia del análisis estadístico (1,2). La alternativa más común al metaanálisis es el análisis "cualitativo" de los distintos estudios, analizando si son coincidentes o discrepantes en sus conclusiones. Según Hunter y Smidcht (2) hay que ser cautos a la hora de interpretar las conclusiones de una revisión sistemática sin metaanálisis, ya que las variaciones entre los resultados de los estudios pueden deberse al error de muestreo.

Cuando incluyen metaanálisis, las revisiones sistemáticas suelen analizar la relación entre dos variables (por ejemplo, una intervención y un determinado resultado) o la relación existente en varias parejas de variables (es decir, de forma bivariente). No obstante, en la actualidad existen metodologías de análisis que permiten estudiar conjuntamente las relaciones entre más de dos variables de forma simultánea (3). Estas metodologías permiten analizar las correlaciones semiparciales y los efectos moderadores de una serie de covariables (como es el caso de la meta-regresión) e incluso analizar los efectos directos, indirectos y mediadores de unas variables con otras (como es el caso del *model-driven meta-analysis*). La meta-regresión será introducida más adelante. Aquellos lectores interesados en *model-driven meta-analysis* pueden consultar el texto de Cooper y cols. (3).

Pasemos seguidamente a analizar los pasos necesarios para la elaboración de una revisión sistemática.

Elaboración de una revisión sistemática

La realización de una revisión sistemática necesita de una serie de pasos que se detallan a continuación (1) y que se encuentran resumidos en el Cuadro 1.

Formulación del problema, la pregunta de búsqueda y los criterios de inclusión de estudios

El problema de investigación debe de ser relevante desde el punto de vista enfermero. Eso significa que el problema debe ser importante desde la óptica de la salud comunitaria (frecuencia y trascendencia) y que la investigación de dicho problema aportará resultados de interés para la formulación de teorías o marcos conceptuales, la práctica clínica y/o la investigación. Además, la evidencia disponible sobre dicho problema de investigación, como es lógico, debe ser insuficiente. Lo anterior se traduce en la no existencia de revisiones sistemáticas sobre dicho problema de investigación, en que existan revisiones pero no sean concluyentes, sean antiguas o no incluyan estudios de una determinada zona de interés.

La pregunta de búsqueda deviene del problema de investigación y es la base sobre la que se definen los términos de la búsqueda bibliográfica. La elaboración de la pregunta de búsqueda debe de ser lo más precisa y específica posible. Cuando la revisión sistemática versa sobre estudios de intervención, se ha propuesto el sistema PICO (Paciente, Intervención, Comparación y resultado [*Outcome* en inglés]), si bien dicho sistema se puede utilizar en estudios etiológicos, de diagnóstico o de pronóstico. No obstante, los autores de este capítulo recomendamos un sencillo método que puede utilizarse en cualquier situación. Dicho método consiste en definir la población de estudio (por ejemplo, personas cuidadoras de familiares mayores dependientes), el problema de estudio (por ejemplo, sobrecarga subjetiva) y la cuestión concreta que se quiere estudiar (por ejemplo, la relación entre afrontamiento disfuncional y sobrecarga subjetiva). El lector podrá encontrar más recomendaciones sobre cómo elaborar preguntas de búsqueda en el documento de Rotaache y cols. (4).

Mediante los criterios de inclusión se especifican las características que deben reunir los estudios para ser incluidos en la revisión. Dichas características hacen referencia a: los participantes, el evento de interés (exposición, intervención o procedimiento diagnóstico que se quiere evaluar) y con qué se compara, el resultado a evaluar, y si se estima necesario, el tipo de diseño de los estudios a incluir.

Búsqueda y localización de estudios

La estrategia de búsqueda debe estar claramente descrita (términos de búsqueda y cadena de búsqueda si procede, fuentes a consultar, filtros, etc.), y tiene que plantearse de manera que se evite el sesgo de selección por no inclusión de estudios relevantes, el sesgo de idioma (incluir sólo estudios publicados en un determinado idioma) y el sesgo de publicación.

Cuadro 1. Pasos de una revisión sistemática

1. Formulación del problema
2. Búsqueda y localización de estudios
3. Selección de estudios
4. Recuperación de la información
5. Análisis de la información
6. Interpretación de los resultados
7. Escritura del informe de investigación

Este último es un tipo de sesgo que se origina como consecuencia de la menor probabilidad que tienen de ser publicados los estudios que no encuentran asociación estadística entre variables.

Para poder conseguir los objetivos anteriores, la estrategia de búsqueda debe incluir criterios amplios (búsqueda sensible), no ha de estar limitada por idioma y tiene que abarcar no solamente la búsqueda en las bases de datos relacionadas con el tema en cuestión, sino también la búsqueda manual en revistas clave en dicho tema, la búsqueda de literatura gris (resúmenes de congresos, tesis doctorales, etc.) y la consulta a expertos para la localización de estudios no publicados. La búsqueda puede ampliarse a partir de la bibliografía de los estudios que se vayan seleccionando.

Selección de los estudios que cumplen los criterios de inclusión

La aplicación de los criterios de inclusión ha de realizarse por, al menos, dos observadores; debiendo estar prevista la estrategia para resolver desacuerdos (consenso, tercer revisor, etc.). Es una buena práctica documentar el proceso de inclusión/exclusión para poder detallar en el informe de la investigación el número de estudios excluidos por cada motivo de exclusión.

Valoración de la calidad metodológica de los estudios incluidos

Es necesario que dicha valoración sea realizada también por más de un observador, debiendo definirse también el método de resolución de desacuerdos. Es recomendable cegar a los evaluadores, los autores, la institución y la revista de cada estudio.

La valoración de la calidad metodológica puede utilizarse en una doble dimensión (y de forma no excluyente): bien para calificar a los distintos estudios que se incluyen en el metaanálisis, bien para excluir aquellos estudios que no ofrezcan suficientes garantías metodológicas. El primer caso permite clasificar a los estudios por grupos de calidad, lo cual puede ser utilizado para análisis de sensibilidad (ver más adelante) y/o análisis de la heterogeneidad.

Existen multitud de herramientas para la evaluación de la calidad metodológica. De entre todas ellas son recomendables las listas simples de chequeo frente a las escalas (1,5), ya que no se suele evaluar la validez y fiabilidad de dichas escalas y la puntuación asignada a cada ítem varía considerablemente entre las distintas escalas (5). Dentro de las listas de chequeo, las herramientas de lectura crítica del CASP (*Critical Appraisal Skills Programme*; Universidad de Oxford) son bastante recomendables por su facilidad de uso, su accesibilidad (están en red traducidas al castellano (6)) y porque aportan guías específicas para cada tipo de diseño.

En el caso de utilizar criterios de exclusión podemos definir cuatro situaciones en las que estaría justificada la exclusión de un estudio: presencia de no asociación con insuficiente potencia estadística, presencia de sesgo de clasificación, no control del sesgo de confusión y análisis estadístico inadecuado. La potencia estadística deja de ser un problema al realizar el metaanálisis y el sesgo de confusión deja también de plantear problemas cuando se realiza un metaanálisis de ensayos controlados aleatorizados (ECA), ya que este tipo de diseños controla *per se* dicho sesgo. Así, en un metaanálisis de ECA's se deberían plantear, como criterios de exclusión, la presencia de sesgo de clasificación y el análisis estadístico inadecuado; en los metaanálisis que incluyan otros tipos de diseños se debería de incluir el no control del sesgo de confusión; y en las revisiones sin metaanálisis deberían de incluirse los cuatro criterios.

Por último, es necesario no confundir las listas de chequeo de evaluación de la calidad con las listas de comprobación editorial, las cuales han sido desarrolladas para evaluar si un artículo tiene todas las partes necesarias, con la finalidad última de garantizar una adecuada comunicación de los resultados de la investigación. Entre ellas se encuentra la declaración CONSORT (7) (para ECA's), TREND (8) (para ensayos no aleatorios) o STROBE (9) (para estudios observacionales). Dichas listas de verificación pueden utilizarse para evaluar la calidad, pero hay que tener en cuenta que suelen ser poco manejables por su extensión y que en ellas se refleja si un determinado aspecto aparece o no en un informe de investigación, no si es adecuado o no, por lo que si se utilizan para medir calidad deben ser reconvertidas a tal efecto. Por ejemplo, un ítem que suele aparecer en las anteriores listas de comprobación es si se ha reflejado el cálculo del tamaño muestral. Si se quiere utilizar dicho ítem para medir calidad metodológica, habría que transformarlo en valorar la adecuación de la precisión y/o la potencia estadística que confiere el tamaño muestral analizado.

El proceso de inclusión/exclusión de estudios debe quedar documentado en el informe final de la revisión y plasmado en un diagrama de flujo. En la Figura 1 aparece un esquema básico de diagrama de flujo. La declaración PRISMA (10) reco-

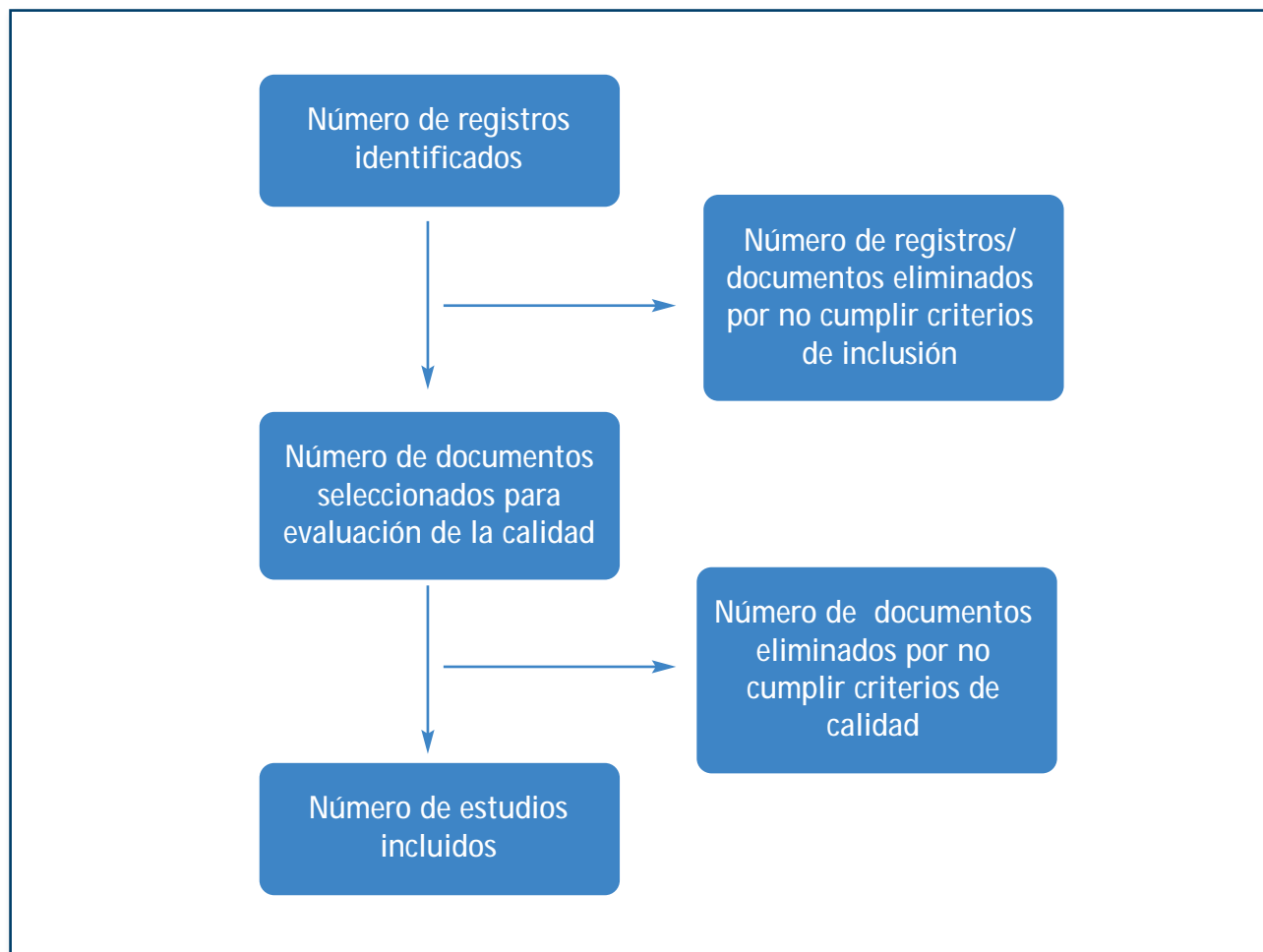


Figura 1. Esquema de un diagrama de flujo

mienda informar de los siguientes datos: 1) registros identificados en las búsquedas, 2) registros identificados en otras fuentes, 3) registros duplicados, 4) registros únicos identificados, 5) registros eliminados, 6) documentos a texto completo analizados, 7) documentos a texto completo excluidos por causa de exclusión, 8) artículos incluidos en la revisión. A este esquema se le puede añadir el número de estudios eliminados por redundantes (mismo artículo o misma muestra), el número de estudios excluidos por deficiente calidad metodológica y el número final de artículos incluidos.

Recuperación y codificación de la información

Una vez que se ha finalizado la inclusión de estudios, el siguiente paso a dar consiste en definir la información a extraer de los distintos estudios y realizar dicha extracción. Todo el proceso debe plasmarse en un protocolo donde aparezcan instrucciones para conseguir una adecuada extracción, codificación y registro de los datos. También es conveniente la elaboración de una hoja de extracción de datos a tal efecto, preferiblemente en formato electrónico de hoja de cálculo, lo cual facilita la gestión de los datos. En dicha hoja, los estudios suelen identificarse mediante el apellido del primer autor y el año. Es recomendable, así mismo, que la extracción se realice por al menos dos personas, con las mismas condiciones de observación por pares establecidas para la valoración de la calidad metodológica (resolución de desacuerdos y cegamiento).

La información que es necesario extraer de los distintos estudios para la realización de la revisión sistemática puede clasificarse en dos categorías: las características de los estudios que supuestamente puedan afectar a los resultados de la revisión y el efecto principal o resultado que se pretende analizar.

Tabla 1. Tamaño del efecto según tipo de variables

Variable independiente (exposición o intervención)	Variable dependiente (efecto o resultado)	Tamaño del efecto
Cuantitativa	Cuantitativa	<i>r</i> de Pearson
Cualitativa dicotómica	Cuantitativa	<i>d</i> de Cohen
Cualitativa policotómica	Cuantitativa	Eta cuadrado (η^2)
Cualitativa dicotómica	Cualitativa dicotómica	Prevalencia relativa u <i>odds ratio</i> en descriptivos transversales <i>odds ratio</i> en casos y controles Riesgo relativo en cohortes, experimentos y cuasi-experimentos
Cualitativa policotómica	Cualitativa dicotómica	Phi de Cramer
Cualitativa dicotómica	Cualitativa policotómica	
Cualitativa policotómica	Cualitativa policotómica	
Cuantitativa	Cualitativa dicotómica	<i>Odds ratio</i>

Para cualquier tipo de revisión sistemática es necesario registrar las características del estudio que pueden afectar a los resultados de la revisión. En una revisión sistemática sin metaanálisis, dichas características son necesarias para clasificar los resultados y para su posterior interpretación, sobre todo si dichos resultados son heterogéneos o contradictorios. En este caso, se intenta ver si existen resultados homogéneos por subgrupos de estudios formados por alguna de las mencionadas características de los estudios. En un metaanálisis, las mencionadas características permiten realizar estudios de sensibilidad y análisis de fuentes de heterogeneidad (ver más adelante ambos aspectos). La definición de estas características va a depender del objeto de estudio de la revisión, si bien pueden aportarse algunas propuestas: 1) características de la población del estudio: sociodemográficas (edad, género, etc.), localización del estudio y contexto del estudio (domicilio, institución, etc.); 2) características metodológicas del estudio: tipo de muestreo (probabilístico o no probabilístico), tamaño muestral, tipo de diseño (si hay más de un tipo de diseño en los estudios incluidos en la revisión), cuestionarios utilizados (si procede) para la medición de la exposición y el efecto (en estudios etiológicos) o el resultado (en estudios de intervención) y 3) características de la intervención, en caso de estudios de intervención; por ejemplo, en intervenciones educativas: tipo (individual, familiar y/o grupal), intensidad (informativas o educativas) y duración (una sesión puntual o varias).

Además de lo anterior, se debe registrar el efecto principal o resultado que se esté investigando, es decir, la asociación estadística o independencia entre la/s variable/ independiente/s y la/s variable/s dependiente/s analizadas. En las revisiones sin metaanálisis debe recogerse si hay o no asociación y, en caso de haber, el tipo (positiva o negativa) y el tamaño del efecto de dicha asociación (ver Tabla 1). En las revisiones con metaanálisis únicamente es necesario el tamaño del efecto, pero debe decidirse previamente con qué medida se va a recoger dicho tamaño del efecto y convertir posteriormente a esa medida las que aparezcan en cada estudio.

Los datos identificativos (autor y año), las principales características y el efecto de cada uno de los estudios incluidos en la revisión sistemática se deben presentar en forma de tabla en el manuscrito o el informe final de resultados, bien en el texto o bien como anexo, en función de su extensión. En caso de metaanálisis, es además conveniente la representación gráfica del efecto principal mediante un gráfico tipo *Forest Plot* (Figura 2).

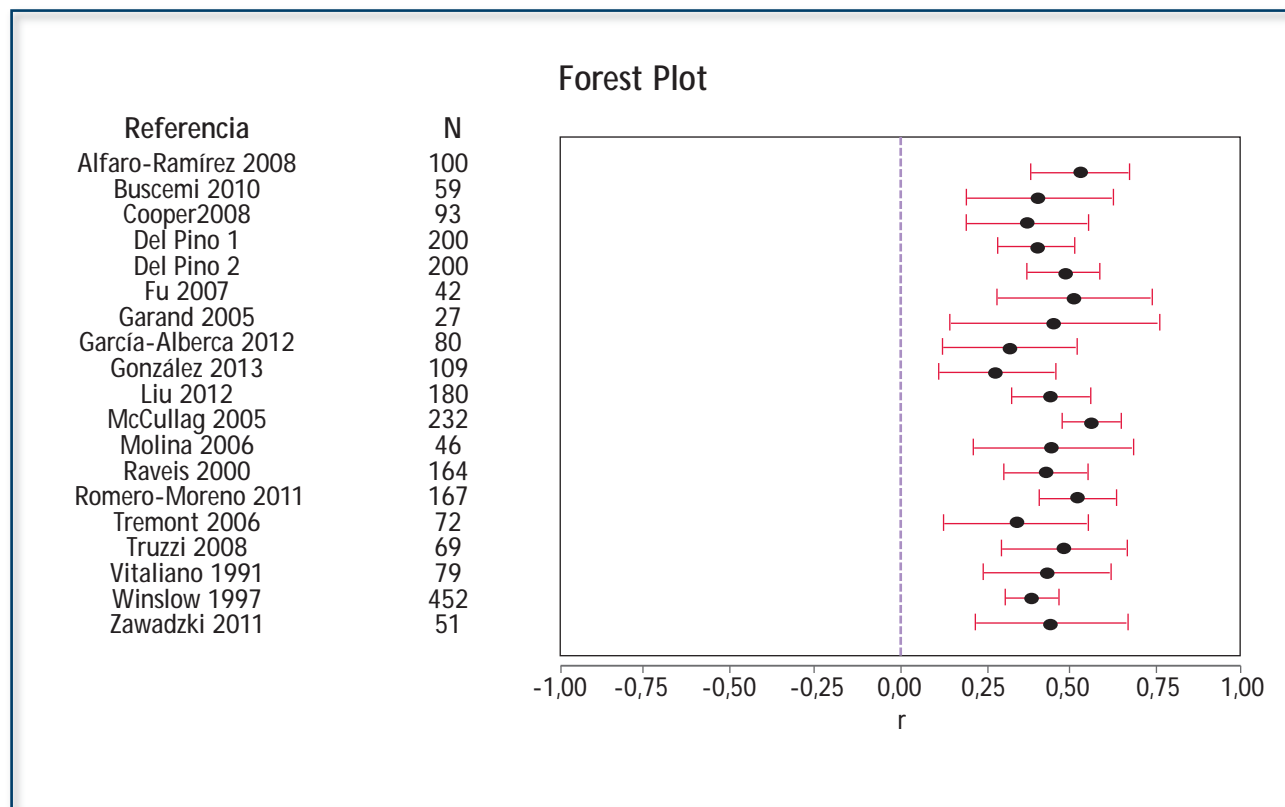


Figura 2. Ejemplo de *Forest Plot*

Análisis de la información

Obviamente, los procedimientos de análisis de los datos difieren en función de la realización o no de metaanálisis, por lo que veremos por separado la síntesis cualitativa y la cuantitativa.

Revisiones sistemáticas sin metaanálisis (síntesis cualitativa)

En las revisiones sin metaanálisis se analiza el grado de coincidencia o discrepancia de las conclusiones de los distintos estudios. El método de análisis en este caso es bastante simple: el recuento de votos, consistente en ver el número de estudios que presentan asociación positiva, negativa y/o ausencia de asociación entre las variables de interés.

A partir de la información de la dirección del efecto en los estudios (daño o beneficio), y sin tener en cuenta la significación estadística ni el tamaño del mencionado efecto, es posible generar otro método de análisis algo más sofisticado denominado *sign test* (3) (se podría traducir como test de signos). Dicho test permite demostrar que el número de estudios con un resultado determinado (por ejemplo, beneficioso) es significativamente mayor que el resultado contrario, es decir, es mayor que el esperado por azar. Para ello se testa la hipótesis nula consistente en que el número de estudios con un determinado efecto es igual al número de estudios con el efecto contrario, es decir, que la probabilidad de que aparezca un determinado efecto es 0,5. Es un test no paramétrico bastante simple, que se basa en la función de probabilidad binomial para el caso concreto de una probabilidad de 0,5. Para su realización, se identifican el número de estudios con efecto positivo y el número de estudios con efecto negativo, sin tener en cuenta la significación estadística ni el tamaño del efecto, de manera que existirá una cantidad mayor (M) y otra menor (m), con independencia del tipo de efecto (positivo o negativo) al que hagan referencia dichas cantidades. Posteriormente se calcula la probabilidad asociada de que existan al menos m estudios y la probabilidad de que existan M o más estudios, en una distribución binomial $B(n, 0,5)$, donde n es el número total de estudios. Al sumar las dos probabilidades se obtiene el valor de p para el test (probabilidad de error al rechazar la hipótesis nula de igualdad de efectos).

Tanto el método de recuento de votos como el *sign test* plantean dos importantes problemas: no tener en cuenta ni la falta de potencia estadística ni el tamaño del efecto. Ambos problemas hacen que los resultados de este tipo de análisis se puedan deber al error de muestreo y que la utilidad clínica de las conclusiones sea limitada. En general, la revisión sistemática sin metaanálisis, y sus métodos de análisis aquí descritos, se reservan para alguno de los siguientes casos: los estudios no proporcionan resultados numéricos pero sí informan de la dirección del efecto, o los estudios son muy diversos en sus poblaciones u otras características pero analizan una cuestión lo suficientemente similar como para poder comparar la dirección de los efectos.

Con independencia del método de análisis empleado, si los datos son heterogéneos puede intentarse buscar hipótesis explicativas de dicha heterogeneidad, ya que un análisis más definitivo de dicha heterogeneidad sólo es posible en metaanálisis, como se verá con posterioridad. Para buscar posibles fuentes de heterogeneidad se pueden agrupar los resultados de los distintos estudios en función de distintas características (calidad metodológica, tamaño muestral, zona geográfica de realización, características socio-demográficas de la población estudiada, etc.); si los resultados de los estudios son homogéneos dentro de cada grupo, la característica utilizada para formar los grupos de estudios podría explicar la heterogeneidad y ser una variable moderadora o de interacción. La existencia de una variable moderadora significa que el efecto estudiado es distinto en cada categoría de dicha variable moderadora. Es decir, la moderación no es un error o artefacto estadístico, sino una relación que aumenta la explicación de cómo se produce un determinado efecto.

Revisiones sistemáticas con metaanálisis (síntesis cuantitativa)

En las revisiones que incluyen metaanálisis, el primer paso es valorar si verdaderamente se puede realizar dicho metaanálisis. Para ello se analiza si existe heterogeneidad, es decir, si existen muchas diferencias entre los distintos estudios en función de un tipo de diseño, características de la población estudiada, métodos estadísticos utilizados, etc. Dicha heterogeneidad se manifiesta en la variabilidad de resultados, y puede determinarse por varios métodos estadísticos, como después veremos. Si la heterogeneidad es muy alta se desaconseja la realización del metaanálisis. Si va de moderada a alta se puede realizar el metaanálisis, pero con las condiciones que veremos más adelante.

Una vez comprobada su adecuación, son necesarios los siguientes pasos para llevar a cabo el mencionado metaanálisis: efectuar la corrección por artefactos del tamaño del efecto de cada estudio, decidir el método de análisis, valorar el sesgo de publicación, calcular el tamaño efecto promedio entre los distintos estudios, realizar análisis de sensibilidad, evaluar la heterogeneidad de los resultados y, en caso de que se confirme dicha heterogeneidad, formular un modelo explicativo de la misma.

Procedimientos de corrección por artefactos

La medida del efecto o resultado de los distintos estudios está afectada por posibles errores o artefactos. Entre los principales errores que pueden afectar a los datos de los estudios incluidos en un metaanálisis podemos destacar (2): 1) el error de muestreo: es decir, el error debido al error aleatorio y al sesgo de selección y 2) los errores de medición: fiabilidad (consistencia interna y fiabilidad test-retest), dicotomización de variables continuas y variación del rango (diferentes desviaciones típicas en la variable independiente) en variables continuas.

Hunter y Schmidt (2) proponen la corrección por errores de medición de las medidas del tamaño del efecto en los distintos estudios antes de realizar la síntesis de sus resultados, así como corregir el efecto promedio por el error de muestreo. Sánchez-Meca (11) aporta una práctica descripción de estos procedimientos para el caso de un metaanálisis con coeficientes de correlación de Pearson, para aquellos lectores interesados en el tema. No obstante, la corrección por error de muestreo se verá más adelante en este artículo, para ese mismo tipo de metaanálisis.

Aunque no se emplee el anterior método propuesto por Hunter y Schmidt, cada vez suele ser más habitual corregir por fiabilidad el tamaño del efecto de cada estudio.

Establecer el método de análisis

En el metaanálisis existen dos métodos de análisis: el método de efectos fijos y el método de efectos aleatorios. El primero considera que los distintos estudios provienen de la misma población, por lo que hay un solo efecto verdadero y la va-

riabilidad entre los estudios se debe al error de muestreo. El segundo método considera que los estudios provienen de poblaciones diferentes, por lo que existe un rango de efectos verdaderos y una fuente de variabilidad añadida, además del azar.

La elección de uno de los dos métodos anteriores ha sido una importante cuestión de debate entre expertos en metaanálisis. Clásicamente se ha mantenido que el método de efectos fijos se debe utilizar cuando no existe heterogeneidad en los resultados del metaanálisis, mientras que el de efectos aleatorios tiene que emplearse en presencia de heterogeneidad.

No obstante, cada vez son más los expertos (3) que opinan que la cuestión principal que determine la elección no debe ser la presencia o no de heterogeneidad, sino el alcance de la generalización de los resultados que pretenda el meta-análisis. En un reciente estudio, Schmidt y cols. (12) apoyan esta idea, argumentando que, dado que la mayoría de metaanálisis se realizan siguiendo un modelo de efectos fijos (por seguir el anterior criterio de ausencia de homogeneidad), la precisión de sus resultados está sobreestimada. En relación con el criterio del alcance de la generalización, se definen dos tipos de inferencias: condicional, cuando se plantea limitar la inferencia de resultados a las situaciones específicas de participantes, localizaciones, intervenciones o factores de riesgo y efectos estudiados, e incondicional, cuando se plantea hacer inferencias más generales, es decir, ampliar el ámbito de inferencia a otras situaciones o instancias más amplias que las estudiadas. En el primer caso estaría indicado el modelo de efectos fijos; en el segundo, el de efectos aleatorios. El criterio de heterogeneidad se emplearía por tanto de forma secundaria, ya que cuando existe heterogeneidad y no se puede explicar (bien porque falta información para analizarla o bien porque no se encuentra una explicación) solo se puede utilizar un modelo de efectos aleatorios. Para completar la regla de decisión habría que tener en cuenta el número de estudios incluidos en el metaanálisis, ya que cuando dicho número es pequeño no está indicado el método de efectos aleatorios (Figura 3).

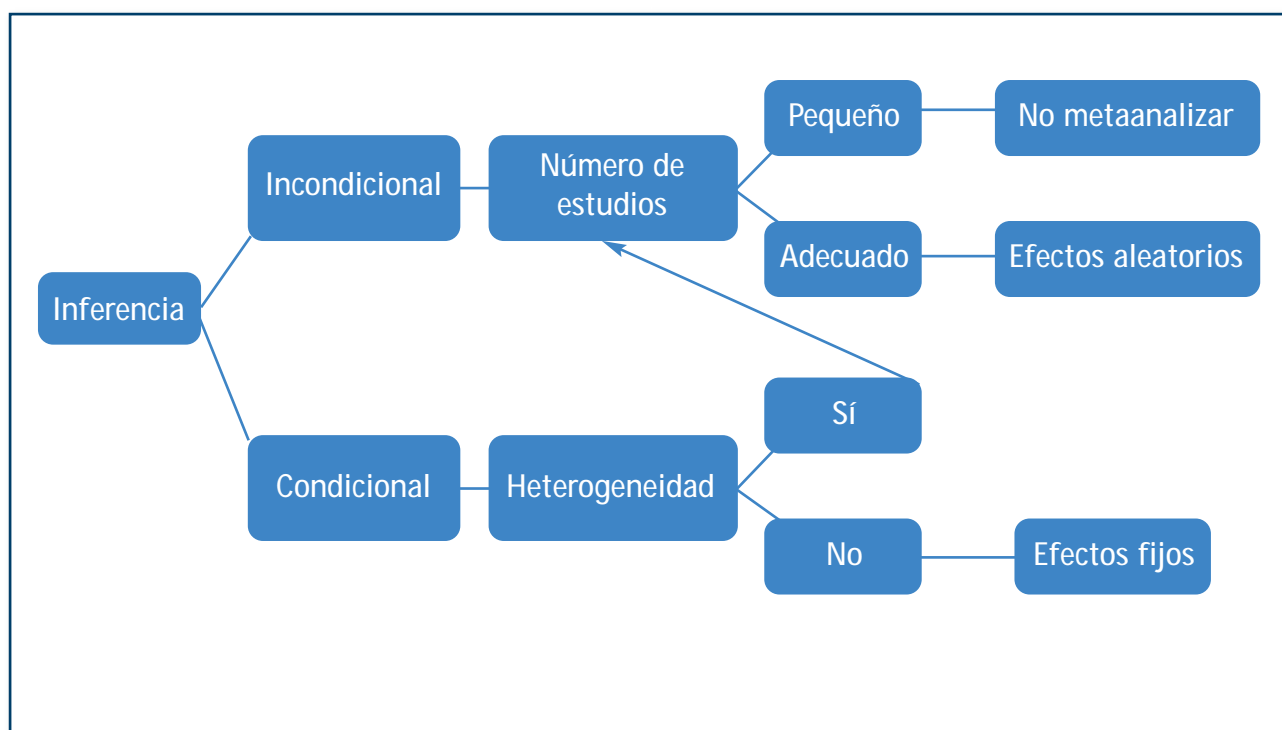


Figura 3. Criterio de decisión para la elección del tipo de modelo de análisis

Valoración del sesgo de publicación

El sesgo de publicación es un tipo de sesgo de selección que es específico de las revisiones sistemáticas. Su presencia, como ya se ha comentado, significa que los estudios incluidos en la revisión difieren del total de estudios realizados en el área en cuestión y su causa radica en que los estudios que no encuentran significación estadística tienen menos probabilidad de ser publicados.

Existen diversos procedimientos para explorar su existencia. Uno de los primeros métodos propuestos ha sido el índice de tolerancia a resultados nulos (*fail-safe N*), el cual cuantifica el número de estudios no publicados necesario para hacer no significativos los resultados del metaanálisis. Es un método bastante criticado y desaconsejado por diversos autores (3,13), fundamentalmente por cuestionar las asunciones en las que se basa y hacer énfasis en el valor de p en vez del tamaño del efecto.

Entre los métodos de detección del sesgo de publicación más utilizados está la elaboración de un gráfico de embudo (*Funnel Plot*, en inglés). Dicho gráfico se construye representando el parámetro o efecto estudiado en el eje de ordenadas y su error estándar en el eje de abscisas, consiguiéndose una figura en forma de embudo que se ensancha hacia la derecha. Si se intercambian los parámetros de los ejes, se obtiene un gráfico en forma de árbol (conocido como gráfico de *árbol de Navidad*; Ver Figura 4). En ambos tipos de gráficos se analiza la heterogeneidad trazando una línea horizontal (en el primer gráfico) o vertical (en el segundo) que pase por el valor del efecto ponderado (ver más adelante) y valorando la simetría a ambos lados de dicha línea.

Dado el alto grado de subjetividad que plantea el anterior método gráfico (13) se recomienda completar el análisis del sesgo de publicación con diversos métodos estadísticos que existen al respecto, entre los cuales tenemos (14): el método de Begg, el método de Egger y la regresión sobre el gráfico de embudo de Macaskill y cols. (18). Siguiendo a Palma y Delgado (14), el método de Begg (15) no debe utilizarse, ya que tiene menor sensibilidad y especificidad que los otros dos. En relación a los dos últimos métodos, el de regresión sobre el gráfico de embudo podría presentar menos falsos positivos (error alfa) que el de Egger, si bien éste presenta mayor potencia estadística y discriminatoria a igualdad de error tipo I (16).

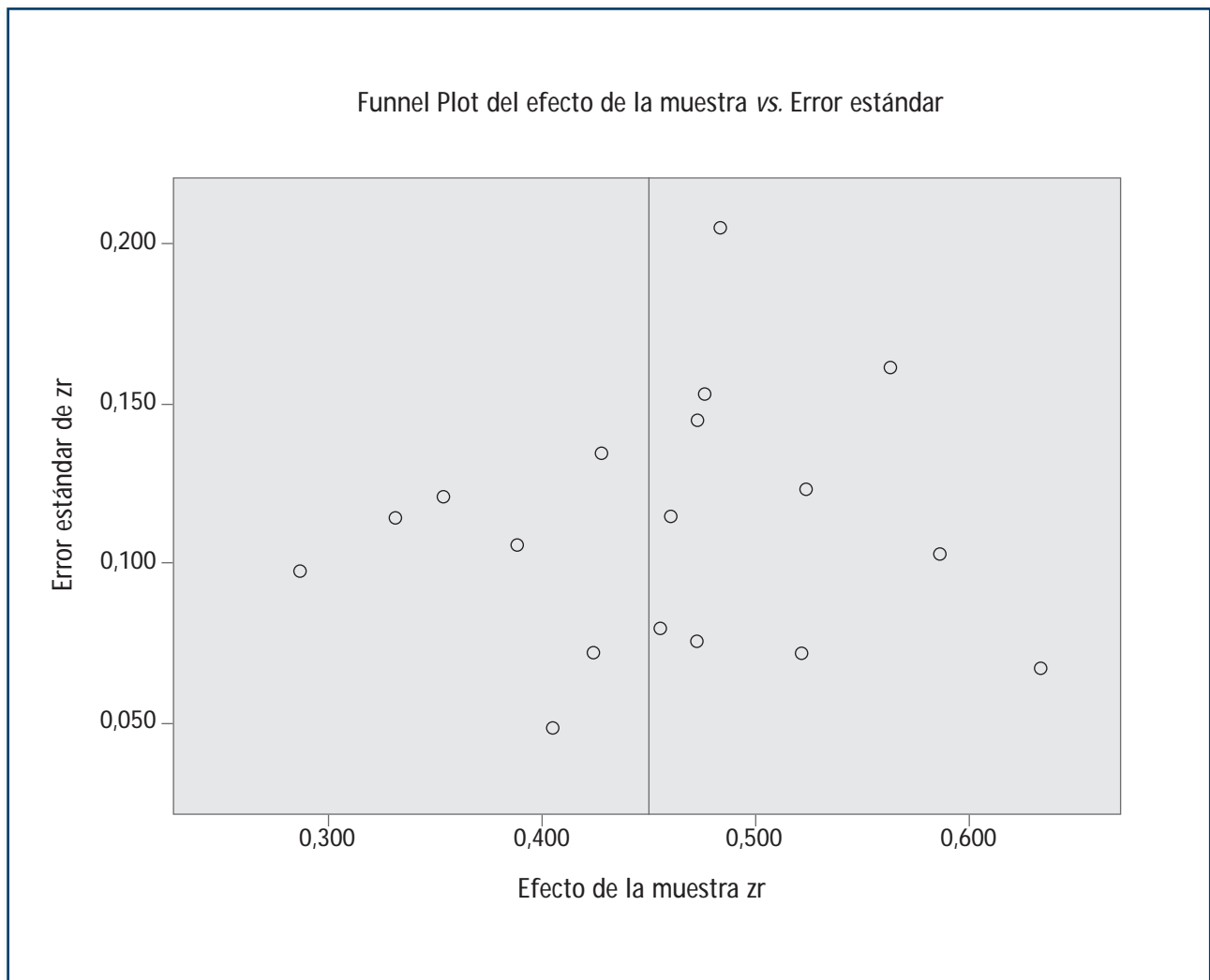


Figura 4. Ejemplo de gráfico de árbol de Navidad

El método de Egger (17) consiste en una regresión lineal simple en la que la variable dependiente es el valor z del tamaño del efecto de cada estudio y la variable independiente su error estándar. Si el valor de la ordenada en el origen (β_0) se acerca a 0 y su valor de p es no significativo, se considera que no hay sesgo de publicación. El caso contrario pone de manifiesto que los estudios pequeños solamente se publican cuando sus resultados son significativos, forzando a la recta de regresión a cruzar el eje de ordenadas alejada del valor 0 (14).

La regresión sobre el gráfico de embudo (18) se lleva a cabo realizando una regresión lineal simple en la que la variable dependiente es el tamaño del efecto de cada estudio (ponderado con el inverso de la varianza) y la variable independiente es su tamaño muestral. En dicha regresión se analiza el valor de la pendiente de la recta (β_1), de manera que si dicho valor se acerca a 0 y su valor de p es no significativo, se considera que no hay sesgo de publicación. Una pendiente de la recta de regresión próxima al valor 0 sugiere que la influencia de los estudios por encima del valor ponderado del efecto es similar a la de los situados por debajo de dicho valor para los distintos tamaños muestrales (14).

Hay que tener en cuenta que el sesgo de publicación está relacionado con la heterogeneidad de los resultados del metaanálisis, ya que a veces el sesgo de publicación puede ser el responsable de la ausencia de heterogeneidad en los resultados. Por tanto, la ausencia de heterogeneidad únicamente puede demostrarse en ausencia de sesgo de publicación.

Existen diversos métodos que podríamos denominar de corrección del sesgo de publicación. Estos métodos permiten calcular el valor de efecto estudiado en condiciones hipotéticas de ausencia de sesgo de publicación. Por supuesto, estos cálculos no son definitivos, pero sí permiten, como mínimo, analizar cuál ha sido el impacto del sesgo de publicación en los resultados del metaanálisis. Entre dichos métodos cabe destacar por su utilidad el método *Trim and Fill* (19) y el método de Vevea y Woods (13). Por su grado de especialización no vamos a describir dichos métodos en este artículo, pero el lector interesado puede consultar a Palma y Delgado (14) para el método *Trim and Fill* y a Vevea y Woods (13) para el método de dichos autores.

Cálculo del tamaño del efecto promedio

El cálculo del tamaño del efecto promedio dependerá de la medida del efecto que se haya elegido para integrar los resultados de los diferentes estudios y del método de análisis. En la Tabla 2 aparecen los distintos métodos disponibles para las medidas de efecto más utilizadas.

De todos ellos nos centraremos en los métodos relacionados con el coeficiente de correlación lineal simple de Pearson (r), por ser frecuente en la literatura enfermera y estar poco representados en los manuales sobre metaanálisis de ciencias de la salud. Los métodos de Rosenthal y Rubin, Hedges y Olkin y Hedges y Vevea son similares, por lo que describiremos el último. De entre los métodos para efectos aleatorios, Brannick y cols. (20), en un análisis comparativo de diversos métodos; recomienda el método de Hunter y Schmidt, que será el que describiremos.

Tabla 2. Métodos de integración de los resultados en un metaanálisis

Medida	Efectos fijos	Efectos aleatorios
RR, OR, RAR	Mantel-Haenzel Inverso de la varianza	DerSimonian-Laird
d	Inverso de la varianza, con ajustes de Cohen, Glass o Hedges	DerSimonian-Laird, con ajustes de Cohen, Glass o Hedges
r	Transformaciones de Fisher con ponderación por el inverso de la varianza (Rosenthal y Rubin, Hedges y Olkin y Hedges y Vevea)	Hunter y Schmidt Hedges y Vevea

Abreviaturas: RR: riesgo relativo, OR: *odds ratio*, RAR: reducción absoluta del riesgo (también denominado diferencia de riesgos), d : d de Cohen (diferencia estandarizada de medias), r : coeficiente de correlación de Pearson

El método de Hedges-Vevea para efectos fijos (21) (los métodos de Rosenthal y Rubin y de Hedges-Olkin son prácticamente similares a éste), se basa en el cálculo de la transformación de Fisher para cada valor de r (z_r) y en el cálculo del valor promedio de las transformaciones de Fisher de los coeficientes r de cada estudio (\bar{z}_r), ponderando por el inverso de la varianza. Una vez calculado este valor, se calcula su intervalo de confianza (IC) al 95% y se transforman el valor promedio y su IC en valores r . En el Cuadro 2 se muestran las fórmulas para dichos cálculos.

El método de Hunter y Schmidt (2) se basa en el cálculo del coeficiente r promedio ponderando por tamaño muestral (\bar{r}) y el posterior cálculo de sus intervalos de credibilidad, los cuales están basados en la varianza real de las correlaciones poblacionales, es decir, la varianza corregida por el error de muestreo (σ_p^2). Dicha varianza se calcula restando la varianza debida al error de muestreo a la varianza total observada en las correlaciones empíricas. En el Cuadro 3 aparecen las fórmulas para la realización de los mencionados cálculos.

Field y Gillett (22) han desarrollado varias sintaxis de SPSS para muchos de los procedimientos descritos en la Tabla 2. Concretamente, podemos encontrar sintaxis que nos ayuden a la realización de metaanálisis básicos para coeficientes r y d de Cohen. Dichas sintaxis, junto con las explicaciones pertinentes de cómo utilizarlas, pueden encontrarse en la página web de los autores (23).

Analizar la sensibilidad

Este análisis consiste en la realización de varios metaanálisis con subgrupos de estudios o con distintos tipos de análisis. Entre las posibilidades del análisis por subgrupos de estudios tenemos: estudios de alta calidad metodológica frente a estudios de baja calidad, estudios de muestras grandes frente a estudios de muestras pequeñas, estudios publicados frente a no publicados. También es posible analizar la muestra total de estudios frente al resultado que se obtiene al excluir los estudios *outliers* (fuera de la nube de puntos del gráfico de embudo o de árbol de Navidad), al excluir los estudios con muestras muy grandes, etc. A este último tipo de análisis de sensibilidad se le llama también análisis de la influencia. Entre las posibilidades de análisis por subgrupos de análisis tenemos, por ejemplo, efectos fijos frente a efectos aleatorios.

La coincidencia de los resultados de las distintas posibilidades de metaanálisis anteriormente mencionadas aporta robustez a las conclusiones de la revisión sistemática.

Cuadro 2. Fórmulas para la aplicación del método de efectos fijos para la r

Transformación de Fisher para la r :	$z_r = 0,5 \times \ln \left(\frac{1+r}{1-r} \right)$
z_r promedio ajustada por el inverso de la varianza:	$\bar{z}_r = \frac{\sum(N_i-3)z_{ri}}{\sum(N_i-3)}$
Error estándar de \bar{z}_r :	$SE(\bar{z}_r) = \sqrt{\frac{1}{\sum(N_i-3)}}$
Intervalo de confianza al 95%:	$IC = \bar{z}_r \pm 1,96 \times SE(\bar{z}_r)$
Conversión de valores z_r en r :	$r = \frac{e^{(2 \times z_r)} - 1}{e^{(2 \times z_r)} + 1}$

Abreviaturas: ln: logaritmo neperiano, Ni: tamaño muestral de cada estudio, Zri: valor de la transformación de Fisher del coeficiente de correlación de Pearson de cada estudio, r: coeficiente de correlación de Pearson de cada estudio, e: número e

Cuadro 3. Fórmulas para la aplicación del método de efectos aleatorios de Hunter y Schmidt para metaanálisis con coeficiente r

Coefficiente r promedio:	$\bar{r} = \frac{\sum N_i r_i}{\sum N_i}$
Varianza observada:	$\sigma_r^2 = \frac{\sum N_i (r_i - \bar{r})^2}{\sum N_i}$
Varianza debida al error de muestreo:	$\sigma_e^2 = \frac{(1 - \bar{r}^2)^2}{(\sum N_i / k) - 1}$
Varianza real:	$\sigma_p^2 = \sigma_r^2 - \sigma_e^2$
Intervalos de credibilidad de r al 95%:	<i>Intervalos de credibilidad de</i> $\bar{r} = \bar{r} \pm 1,96 \sqrt{\sigma_p^2}$
Abreviaturas: Ni: tamaño muestral de cada estudio, ri: coeficiente de correlación de Pearson de cada estudio, \bar{r} : coeficiente de correlación promedio ponderado por tamaño muestral, k: número de estudios	

Análisis de la heterogeneidad de los resultados

La valoración de la heterogeneidad de los resultados de los distintos estudios se lleva a cabo mediante diversos procedimientos estadísticos. Probablemente el método más utilizado sea el test basado en el estadístico Q desarrollado por Cochran (24), el cual analiza la hipótesis nula de igualdad de los efectos de los distintos estudios incluidos en el metaanálisis. Dicho estadístico sigue una distribución χ^2 y se computa calculando las diferencias cuadráticas del efecto de cada estudio respecto del efecto promedio y ponderando dichas diferencias por el mismo método de ponderación utilizado en el cálculo del efecto promedio (tamaño muestral o inverso de la varianza). La interpretación del test, dado que lo que se contrasta es la hipótesis nula de igualdad de efectos es: si $p \geq 0,05$, se acepta la igualdad de efectos y por tanto se afirma ausencia de heterogeneidad; si $p < 0,05$, se rechaza la igualdad de efectos y se afirma la heterogeneidad. El test tiene el inconveniente de su escasa potencia estadística (25), lo cual implica importantes exigencias en cuanto a tamaño muestral y puede llevar a falsas conclusiones de ausencia de heterogeneidad cuando no se cumplen dichas exigencias. Para evitar esto, diversos autores han desarrollado métodos para calcular la potencia estadística de dicho test, como es el caso de Hedges y Pigott (26).

Higgins y cols. (27) han propuesto una alternativa al anterior método que han denominado I2 y que suele conocerse como grado de inconsistencia. Dicho índice se calcula en base al estadístico Q ($Q - [k - 1] / Q * 100$, siendo k el número de estudios) y se interpreta como ausencia de heterogeneidad cuando posee el valor 0 (o valores negativos) y máxima heterogeneidad cuando alcanza el valor 100. Los autores han propuesto los siguientes puntos de corte para la interpretación del índice (27): 25% heterogeneidad pequeña, 50% media y 75% alta.

Por otro lado, Hunter y Schmidt (2) han propuesto un método de valoración de la heterogeneidad basado en su anteriormente comentado modelo de efectos aleatorios, de forma que si al menos el 75% de la varianza observada está explicada por la varianza debida al error de muestreo, es decir, si la varianza debida al error de muestreo supone al menos el 75% de la varianza observada ($100 * \sigma_e^2 / \sigma_r^2 \geq 75$) los efectos se consideran homogéneos, mientras que si es inferior al 75% se concluye que existe heterogeneidad entre los efectos estudiados.

Como ya se ha dicho anteriormente, en caso de que exista alta heterogeneidad está desaconsejada la realización del metaanálisis. Si dicha heterogeneidad es media-alta, es necesario explorar las fuentes de heterogeneidad, es decir, la existencia de variables moderadoras que expliquen dicha heterogeneidad. Dicho de otro modo, hay que identificar variables que expliquen la variabilidad en los resultados de los estudios. Como ya se ha comentado anteriormente, dichas variables pue-

den ser las características de los estudios analizados (diseño, calidad, tamaño muestral, tipo de muestreo, etc.) o de la población estudiada (género, edad, etc.).

Para analizar las fuentes de heterogeneidad pueden segmentarse los estudios por las anteriores variables, realizar metaanálisis en los distintos subgrupos de estudios y ver si los resultados coinciden o no (al estilo del análisis de la sensibilidad). En caso de los resultados de los grupos fuesen homogéneos dentro de cada grupo, la variable utilizada para formar los subgrupos explicaría la heterogeneidad.

Otra opción posible para el estudio de la heterogeneidad es la anteriormente mencionada meta-regresión. Esta técnica estadística funciona igual que una regresión múltiple, donde la variable dependiente es el efecto estudiado y las variables independientes son las posibles variables moderadoras o explicativas de la heterogeneidad. Aquellas variables que demuestren asociación estadística con la variable independiente serán las variables moderadoras que permiten explicar la heterogeneidad. Hoy en día cada vez son más los paquetes estadísticos de metaanálisis que incluyen esta opción, amén de existir una amplia oferta de macros diseñadas para ser ejecutadas con los paquetes estadísticos más habituales (ver a modo de ejemplo las sintaxis para SPSS propuestas por Field y Gillett para el análisis de la heterogeneidad (23)).

Interpretación de los resultados

Este apartado debe incluir el análisis de las limitaciones de la revisión sistemática, la valoración de la fuerza de la evidencia que aportan sus conclusiones, la aplicabilidad de las mismas, el ámbito de generalización de las conclusiones y las implicaciones para futuras investigaciones.

Elaboración del informe final

Para ver las partes que debe tener un informe de investigación sobre una revisión sistemática podemos apoyarnos en las listas de comprobación para revisiones sistemáticas de ensayos clínicos (PRISMA (10)) o de estudios observacionales (MOOSE (28)). De cualquier forma, hay algunas recomendaciones que no se deben olvidar al elaborar el informe de investigación: 1) presentar los resultados de los estudios incluidos y excluidos en un diagrama de flujo, 2) presentar una estadística de estudios excluidos por cada causa de exclusión (no cumplimiento de criterios de inclusión y de calidad), 3) incluir la información de los estudios incluidos en forma de tabla y 4) incluir una lista de referencias de los estudios excluidos (al menos por no cumplir los criterios de calidad). En los artículos de revistas científicas, el punto 3 (en el caso de revisiones con metaanálisis) y el punto 4 (en cualquier caso) no suelen incluirse en el texto, siendo conveniente que aparezcan como material complementario al artículo.

Revisiones y evidencia científica

Para finalizar este apartado de revisiones, interesa exponer algunas ideas sobre el nivel de evidencia científica aportado por las revisiones sistemáticas, ya que de ello va a depender la utilidad de sus conclusiones en la práctica clínica.

En este sentido, es necesario comentar que las revisiones sistemáticas con metaanálisis realizadas a partir de ECA's aportan la máxima evidencia posible, pero no siempre se disponen de suficientes ECA's para hacer una revisión adecuada. Otras veces el tema de estudio es un factor de riesgo y por tanto no es posible la experimentación. En ambos casos la alternativa es la revisión sistemática de estudios observacionales.

Dichas revisiones de estudios observacionales plantean el problema, ya comentado con anterioridad, de que los diseños observacionales no previenen *per se* el sesgo de confusión y de que los sesgos de selección son más frecuentes en diseños de casos y controles y descriptivos transversales que en el resto de diseños. Por tanto, en este tipo de revisiones el análisis de la calidad de los estudios debe incluir el control del sesgo de confusión y potenciar el control del sesgo de selección. Egger (1) plantea que, en las revisiones de estudios observacionales, la realización de un metaanálisis no debe ser el objetivo principal, ya que la suma de estudios sesgados produciría resultados muy precisos pero erróneos. Por lo tanto, una estrategia adecuada para este tipo de revisiones consistiría en aplicar de forma muy estricta los criterios de calidad y realizar una valoración "cualitativa" de los datos de los estudios, con especial énfasis en el análisis de las posibles fuentes de heterogeneidad entre estudios.

En el caso concreto de las revisiones de estudios descriptivos transversales, a las anteriores precauciones hay que añadir que las conclusiones de dichas revisiones son tentativas, ya que, como se ha comentado anteriormente, dicho tipo de estudios no puede establecer causalidad. En este sentido, Melnyk y Fineout-Overholt (29) clasifican el nivel de evidencia aportado por las revisiones de estudios descriptivos transversales por debajo de la aportada por los estudios analíticos.

Bibliografía

- [1] Egger M, Smith GD, Altman D. Systematic reviews in health care: meta-analysis in context. 2nd ed. London: BMJ Books; 2001.
- [2] Hunter JE, Schmidt FL. Methods of meta-analysis: Correcting error and bias in research findings. 2nd ed. Thousand Oaks, CA: SAGE Publications; 2004.
- [3] Cooper HM, Hedges LV, Valentine JC. The handbook of research synthesis and meta-analysis. New York: Russell Sage Foundation Publications; 2009.
- [4] Rotaeche del Campo R, Pérez Irazusta I, Etxeberria Agirre A, Balagué Gea A. Cómo formular preguntas clínicas contestables. Guías Clínicas Fisterra 2010; 10(18). [En línea] [fecha de acceso: 11 de noviembre de 2013]. URL disponible en: http://www.cochrane.es/files/Recursos/Formulacion_Preguntas_Clinicas_FISTERRA.pdf
- [5] Centre for Reviews Dissemination. Systematic reviews: CRD's guidance for undertaking reviews in health care. York: Centre for Reviews and Dissemination, University of York; 2009.
- [6] Critical Appraisal Skills Programme. Herramientas de lectura crítica. Oxford: Institute of Health Sciences, Oxford University; 2012. [En línea] [fecha de acceso: 11 de noviembre de 2013]. URL disponible en: <http://redcaspe.org/drupal/?q=node/29>
- [7] Moher D, Schulz K, Altman D. La Declaración CONSORT: Recomendaciones revisadas para mejorar la calidad de los informes de ensayos aleatorizados de grupos paralelos. Rev Sanid Milit 2002; 56(1):23-28.
- [8] Vallvé C, Artés M, Cobo E. Estudios de intervención no aleatorizados (TREND). Medicina Clin 2005; 125:38-42.
- [9] Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. Declaración de la iniciativa STROBE (Strengthening the Reporting of Observational Studies in Epidemiology): Directrices para la comunicación de estudios observacionales. Gac Sanitaria 2008; 22(2):144-150.
- [10] Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. J Clin Epidemiol 2009; 62(10):1006-12.
- [11] Sánchez-Meca J. Metaanálisis para la investigación científica. En: Sarabia Sánchez F (ed.). Metodología para la investigación en marketing y dirección de empresas. Madrid: Pirámide; 1999. p. 173-201.
- [12] Schmidt FL, Oh IS, Hayes TL. Fixed- versus random-effects models in meta-analysis: model properties and an empirical comparison of differences in results. Br J Math Stat Psychol 2009; 62(Pt 1):97-128.
- [13] Vevea JL, Woods CM. Publication bias in research synthesis: sensitivity analysis using a priori weight functions. Psychol Methods 2005; 10(4):428-443.
- [14] Palma Pérez S, Delgado Rodríguez M. Consideraciones prácticas acerca de la detección del sesgo de publicación. Gac Sanitaria 2006; 20:10-16.
- [15] Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics 1994; 50(4): 1088-101.
- [16] Hayashino Y, Noguchi Y, Fukui T. Systematic evaluation and comparison of statistical tests for publication bias. J Epidemiol 2005; 15(6):235-243.
- [17] Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997; 315(7109):629-634.
- [18] Macaskill P, Walter SD, Irwig L. A comparison of methods to detect publication bias in meta-analysis. Stat Med 2001; 20(4):641-654.
- [19] Duval S, Tweedie R. Trim and Fill: A Simple Funnel Plot-Based Method of Testing and Adjusting for Publication Bias in Meta-Analysis. Biometrics 2000; 56(2):455-463.

- [20] Brannick MT, Yang L-Q, Cafri G. Comparison of weights for meta-analysis of r and d under realistic conditions. *Organizational Research Methods* 2011; 14(4):587-607.
- [21] Hedges LV, Vevea JL. Fixed-and random-effects models in meta-analysis. *Psychological methods* 1998; 3(4):486-504.
- [22] Field AP, Gillett R. How to do a meta-analysis. *British Journal of Mathematical and Statistical Psychology* 2010; 63(3):665-694.
- [23] Field AP, Gillett R. How to do a meta-analysis. 2009. [En línea] [fecha de acceso: 11 de noviembre de 2013]. URL disponible en: http://www.statisticshell.com/meta_analysis/how_to_do_a_meta_analysis.html
- [24] Cochran WG. The combination of estimates from different experiments. *Biometrics* 1954; 10(1):101-129.
- [25] Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; 327(7414):557.
- [26] Hedges LV, Pigott TD. The power of statistical tests in meta-analysis. *Psychological methods* 2001; 6(3):203.
- [27] Higgins J, Thompson S, Deeks J, Altman D. Statistical heterogeneity in systematic reviews of clinical trials: a critical appraisal of guidelines and practice. *Journal of health services research & policy* 2002; 7(1):51-61.
- [28] Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000; 283(15):2008-12.
- [29] Melnyk B, Fineout-Overholt E. Evidence-based practice in nursing & healthcare: A guide to best practice. Philadelphia, PA: Lippincott Williams & Wilkins; 2005.