

BJPsych

Advances

Continuing professional development in psychiatry

Spanish translation of:

The usefulness and interpretation of systematic reviews

Katharine A. Smith, Andrea Cipriani & John R. Geddes

BJPsych Advances 2016, 22: 132–141

doi: 10.1192/apt.bp.114.013128

Reprints/permissions To obtain reprints or permission to reproduce material from this article,
please write to permissions@rcpsych.ac.uk

© The Royal College of Psychiatrists 2016

ARTÍCULO

Interpretación y utilidad de las revisiones sistemáticas[†]

Katharine A. Smith, Andrea Cipriani & John R. Geddes

Katharine Smith es especialista honoraria en Psiquiatría en el National Institute for Health Research [Instituto Nacional de Investigación en Salud] (NIHR) Oxford cognitive health Clinical Research Facility [Servicio de Investigación Clínica en salud cognitiva de Oxford] (CRF) y el Department of Psychiatry at Oxford University [Departamento de Psiquiatría de la Universidad de Oxford]. **Andrea Cipriani** es Profesor Adjunto en el Departamento de Psiquiatría de la Universidad de Oxford, especialista honorario en Psiquiatría en el Oxford Health NHS Foundation Trust [Servicio Nacional de Salud británico en Oxford] y editor en jefe de Evidence-Based Mental Health (<http://ebmh.bmj.com>). Su interés clínico principal son los trastornos del ánimo y su investigación se centra en la evaluación de los tratamientos en Psiquiatría, tratando de desarrollar la metodología de síntesis de la evidencia para informar mejor a los clínicos en su toma de decisiones diaria. **John Geddes** es Jefe del Departamento de Psiquiatría de la Universidad de Oxford y especialista honorario en Psiquiatría en el Oxford Health NHS Foundation Trust, en donde provee atención clínica a pacientes con trastornos del ánimo y se especializa en trastorno bipolar. Es también Director de Investigación y Desarrollo en el Oxford Health NHS Foundation Trust y Director del NIHR Oxford cognitive health CRF. **Correspondencia** Dr Katharine Smith, NIHR Oxford cognitive health Clinical Research Facility, Warneford Hospital, Oxford OX3 7JX, UK. Email: katharine.smith@psych.ox.ac.uk

RESUMEN

Una parte esencial de la práctica clínica es mantenerse al día con la mejor evidencia sobre intervenciones de tratamiento, pero esto parece ser una tarea abrumadora para los clínicos ocupados. Las revisiones sistemáticas y los metanálisis proveen un resumen útil y conveniente del conocimiento y forman una parte esencial de un enfoque de la práctica clínica basado en la evidencia. Sin embargo, estas revisiones varían en metodología y por lo tanto en la calidad de las recomendaciones que proveen. Los clínicos necesitan sentirse confiados en sus habilidades de apreciación crítica, de modo que puedan evaluar los méritos relativos de las revisiones sistemáticas. En este artículo discutimos las fortalezas y limitaciones de los distintos tipos de síntesis de evidencias, para posibilitar que el lector se sienta más confiado en valorar la información científica para utilizarla en la práctica clínica.

OBJETIVOS PEDAGÓGICOS

- Entender qué es una revisión sistemática y cómo realizar una evaluación crítica de sus fortalezas y limitaciones, incluyendo la identificación de las fuentes potenciales de sesgo
- Entender qué es un metanálisis y cuando usarlo, cómo evaluar su validez interna y externa, y las dificultades de la heterogeneidad clínica y estadística
- Apreciar las metodologías avanzadas (por ej. metanálisis de datos de pacientes individuales y metanálisis de red) utilizadas para individualizar la respuesta al tratamiento y evaluar la efectividad comparativa

DECLARACIÓN DE CONFLICTOS DE INTERÉS

Ninguna

diagnóstica y terapéutica que optimice la calidad de vida y las evoluciones clínicas de los pacientes. Los médicos tienen la obligación profesional de ‘mantener actualizados su conocimiento profesional y habilidades’ (General Medical Council 2013). Sin embargo, mantenerse al día con la evidencia actual es una tarea hercúlea. Más de 2000000 de artículos se publican cada año en 20000 revistas biomédicas, y aun si un clínico restringiera su lectura a revistas psiquiátricas de alto rendimiento, necesitaría leer más de 5000 artículos por año (Geddes 1999), una tarea que simplemente no es factible para el clínico ocupado.

Para mantenerse eficientemente actualizado, el clínico necesita un sistema que resuma los hallazgos de la investigación primaria en una forma que provea una síntesis confiable y fácil de leer del conocimiento actual. Sin embargo, como cualquier otra forma de investigación, estos resúmenes (o revisiones) varían en calidad y son susceptibles de varias formas de error sistemático, o sesgo. Para utilizar las revisiones efectivamente se debe estar consciente de las ventajas y limitaciones potenciales de los diferentes tipos de revisión, de manera de poder sopesar los resultados y los méritos relativos de la metodología, y de esta manera evaluar críticamente las conclusiones de la revisión.

¿Qué es una revisión sistemática?

Las revisiones sistemáticas sintetizan estudios de investigación primaria utilizando estrategias metodológicas específicas para limitar el riesgo de sesgos. En una revisión sistemática los autores hacen una pregunta claramente formulada y utilizan métodos sistemáticos y explícitos para identificar, seleccionar y evaluar críticamente toda la evidencia relevante para considerar dicha pregunta (Higgins 2011a). Las revisiones sistemáticas difieren significativamente de las revisiones más tradicionales (o narrativas) (Tabla 1).

En una revisión sistemática todos los métodos son descriptos y claramente especificados en un protocolo de revisión, de modo que el lector entienda exactamente cuáles estrategias han sido usadas. Los métodos deben ser descriptos en

Una parte esencial del logro de buenas decisiones clínicas durante la práctica cotidiana es mantenerse actualizado con el conocimiento de la eficacia comparativa de los diferentes tratamientos (Cipriani 2013a). Para tomar las mejores decisiones clínicas, los médicos deben combinar su propia pericia y entrenamiento clínico con evidencia científica de alta calidad y con las opiniones del paciente (Guyatt 2000). Esta combinación puede crear una poderosa alianza

[†]Título original: ‘The usefulness and interpretation of systematic reviews’, *BJPsych Advances*, 2016, 22, 132–141. Traducido por: Rodolfo Zaratiegui. © The Royal College of Psychiatrists.

suficiente detalle para permitir que cualquiera que utilice la misma metodología reproduzca los mismos resultados. Esto mejora la confiabilidad y precisión de las conclusiones.

El primer paso es identificar una pregunta específica a ser considerada por la revisión. El rango de esta pregunta necesita ser estrecho ya que no es posible ni útil recoger toda la evidencia disponible en un tópico que es demasiado general o amplio. La naturaleza de la pregunta determina el tipo de evidencia de investigación que será revisado (por ejemplo, estudios aleatorizados versus observacionales) y cuáles estudios serán incluidos o excluidos de acuerdo con criterios explícitos, predefinidos en el protocolo de revisión. Por ejemplo, una pregunta relacionada con la eficacia de los tratamientos (¿Cuál tratamiento es mejor? ¿Qué dosis es más efectiva?) generalmente está mejor respondida revisando evidencia de ensayos controlados randomizados (ECR), ya que la aleatorización protege contra el sesgo de selección (Tabla 2). Sin embargo, los ECRs no son el diseño de ensayo apropiado para todas las cuestiones. Preguntas de etiología (¿El ACV predispone a un trastorno depresivo posterior?) están mejor respondidas por estudios de cohorte o de casos y controles. Las preguntas diagnósticas (¿Qué tan bien una herramienta de tamizaje detecta casos de psicosis temprana?) están mejor estudiadas con investigaciones transversales y prospectivas de pacientes en riesgo. Estos estudios son llamados estudios de validación diagnóstica cuando un método diagnóstico nuevo es comparado con un método de referencia aceptado.

Una vez que se ha identificado la pregunta, la revisión procede a la identificación sistemática de todos los estudios relevantes que responden a esa pregunta (de acuerdo con los métodos descritos en el protocolo). Los datos publicados se acceden a menudo vía bases de datos electrónicas tales como PubMed, Embase, PsycINFO y CINAHL. Se necesita tener cuidado en la elección y ordenamiento de las palabras claves utilizadas en la búsqueda, ya que esto va a tener un efecto significativo en las investigaciones que se identifiquen. Los revisores deben buscar no sólo los estudios publicados, sino también datos no publicados y ‘literatura gris’ (material escrito publicado informalmente, como reportes técnicos o documentos de trabajo de grupos de investigación). Los revisores deben practicar todos los esfuerzos necesarios para contrarrestar cualquier sesgo de publicación que pueda existir (ver ‘Métodos para reducir las consecuencias del sesgo’ más abajo).

Luego de la identificación de los estudios los revisores evalúan críticamente cada uno. Hasta que punto una revisión sistemática puede

TABLA 1 Comparación de las características de las revisiones sistemáticas y narrativas

Revisiones sistemáticas	Revisiones narrativas
Sin sinonimia en castellano	También llamadas ‘revisiones tradicionales’
Recogen todos los estudios que abarcan una pregunta clínica definida	Seleccionan estudios sobre la base de las opiniones de los autores de la revisión, usualmente expertos en el tema
Usan estrategias metodológicas explícitas para identificar todos los estudios sobre un tópico específico	Seleccionan los estudios por criterios implícitos, más que por estrategias metodológicas explícitas
Se consideran e incluyen todos los estudios que cumplen con criterios predefinidos, aunque se les pueda dar un peso diferente en la conclusión final, dependiendo de las fortalezas y debilidades de la metodología de cada estudio	La selección y la síntesis de los resultados están basadas principalmente en la experiencia y opinión del autor/es de la revisión
Los métodos utilizados en la evaluación crítica y síntesis de los datos están claramente definidos	Generalmente no incluyen una sección que describa los métodos utilizados para sintetizar los resultados
Se describen claramente las faltas de datos/debilidades	Las faltas de datos/debilidades se describen según la opinión de los autores de la revisión
La metodología explícita disminuye el riesgo de sesgo, aunque no lo excluye completamente	La falta de una metodología clara y explícita aumenta la posibilidad de sesgo y la interpretación incorrecta de los hallazgos de los estudios. Los resultados pueden ser engañosos, pero la magnitud de la falta de fiabilidad es difícil de juzgar
Pueden ser publicadas en una base de datos de revisiones sistemática (por ej. la Cochrane Library, www.thecochranelibrary.com/view/0/index.html) y se actualizan regular y sistemáticamente de acuerdo con los criterios originales	Generalmente no son actualizadas a medida que nuevos estudios están disponibles Representan un resumen de estudios seleccionados por un experto en un punto particular del tiempo Pueden ser útiles como una herramienta descriptiva para resumir aspectos diferentes de una cuestión compleja

TABLA 2 Tipos de sesgo y las estrategias utilizadas en los ECRs para minimizarlos

Sesgo	Estrategias para prevenirlo
Sesgo de selección (diferencias sistemáticas entre las características basales de los grupos)	Aleatorización (por ej. una tabla de números al azar generada por computadora) Ocultamiento de la asignación (ocultar la secuencia de asignaciones de modo que no pueda ser descubierta)
Sesgo de procedimiento (diferencias sistemáticas entre los grupos en el cuidado o exposición a otros factores)	Participantes y personal del estudio ciegos a cuál intervención ha sido asignada
Sesgo de detección (diferencias sistemáticas entre los grupos en cómo los resultados son determinados)	Participantes y evaluadores ciegos a cuál intervención ha sido asignada
Sesgo de deserción (diferencias sistemáticas entre los grupos en los abandonos del estudio)	Reporte completo de los datos de resultados que incluya los abandonos y exclusiones, con las respectivas razones
Sesgo de reporte (reporte preferencial de sólo los resultados favorables de un estudio)	Reporte completo de los datos

sacar conclusiones depende de si los datos y los resultados de los estudios primarios incluidos son válidos. La validez de un estudio se relaciona con que resuelva la pregunta de investigación ‘correctamente’, es decir sin sesgo (Higgins 2011b). La evaluación de la validez de los estudios

incluidos es así un componente esencial de una revisión sistemática, y debe influenciar el análisis, interpretación y conclusiones de la misma. No siempre está disponible una evidencia de alta calidad para todos los resultados de interés. En ese caso, el resumen de la evidencia puede ser todavía presentado, junto con una medida de calidad para guiar al lector, por ejemplo usando el enfoque llamado Grading of Recommendation Assessment, Development and Evaluation (GRADE, gradiente de recomendaciones de investigación, desarrollo y evaluación) (Guyatt 2008). Este provee un sistema para un orden jerárquico de la calidad de la evidencia y la fortaleza de las recomendaciones que es abarcativo y pragmático, y está siendo crecientemente adoptado a nivel global. Puede ayudar a asegurar que los juicios sobre el riesgo de sesgo, así como otros factores que afectan la calidad de la evidencia (tales como imprecisión, heterogeneidad y sesgo de publicación), sean considerados cuando se interpretan los resultados de las revisiones sistemáticas.

La Cochrane Library (Biblioteca Cochrane) (www.thecochranelibrary.com) es posiblemente la base de datos de revisiones sistemáticas más conocida. El sitio web contiene varias bases de datos diferentes. Estas incluyen la Cochrane Database of Systematic Reviews (CDSR), Cochrane Central Register of Controlled Trials (CENTRAL), Cochrane Methodology Register (CMR), Database of Abstracts of Review of Effects (DARE) y Health Technology Assessment Database (HTA).

Metanálisis

El término se refiere específicamente al uso de técnicas estadísticas para resumir

cuantitativamente datos como parte de una revisión sistemática (Higgins 2011a). Sin embargo, el término se usa a menudo más liberalmente para referirse a cualquier revisión sistemática que use métodos estadísticos para combinar, sopesar y resumir los resultados de varios estudios (Cook 1995). Los resultados de los estudios originales (ej. resultados primarios y secundarios, porcentajes de efectos adversos) se extraen, se colocan juntos y se analizan estadísticamente en un estimativo final combinado. Varios programas de computación están disponibles para efectuar estos análisis, como el RevMan (<http://tech.cochrane.org/revman>) y Meta-DisC (www.hrc.es/investigacion/metadisc_en.htm), ambos de uso libre, Stata (www.stata.com) y Comprehensive Meta-Analysis (www.meta-analysis.com/index.php). Un metanálisis debe tomar en cuenta las características de cada uno de los estudios primarios, ya que la calidad metodológica de los ensayos individuales afectará la calidad de las recomendaciones que cada metanálisis puede proveer. Es importante consignar que los métodos estadísticos de metanálisis deben ser llevados a cabo en el marco de una revisión sistemática (sólo una revisión sistemática puede garantizar una colecta transparente y abarcativa de toda la evidencia disponible, para evitar sesgos sistemáticos en la selección de estudios a analizar). Por el contrario, el metanálisis no es una parte esencial de cada revisión sistemática: en algunos casos puede ser inapropiado combinar los resultados de estudios, por ejemplo si los estudios originales son demasiado diferentes entre sí.

Los resultados generales del metanálisis brindan los efectos principales del tratamiento y los relacionan con la respuesta promedio en un paciente promedio. La práctica clínica, sin embargo, abarca el estudio y tratamiento de un individuo, por lo que los resultados de un análisis de subgrupo (de acuerdo con características clínicas o sociodemográficas diferentes) pueden en principio parecer más relevantes para las decisiones de los clínicos. Los análisis de subgrupo pueden ser efectuados combinando datos de subgrupos específicos de cada estudio. No obstante, los resultados de subgrupos no están siempre reportados en las publicaciones originales y la aleatorización de tratamientos en los estudios primarios pudo no haber sido estratificada de acuerdo con los mismos subgrupos. Además, cuanto más subgrupos se analicen, más probable será que un resultado estadísticamente significativo pero incorrecto sea hallado puramente por azar, como se muestra en el cuadro 1. Como regla general, cualquier análisis de subgrupo dentro de un metanálisis debe ser tratado cuidadosamente y es

CUADRO 1 Los efectos del azar en un análisis de subgrupos

Counsell *et al* (1994) condujeron una investigación sobre los efectos del azar en los resultados de una revisión sistemática que contenía un análisis de subgrupo de un tratamiento ficticio llamado DICE ('dado' en inglés, así mismo se sugiere como un acrónimo de 'Don't Ignore Chance Effects', es decir 'no ignore los efectos del azar' (nota del T.)):

- Se simularon 44 ensayos aleatorizados tirando dados – cada tirada dando el resultado para un 'paciente'
- Cada investigador llevó a cabo dos ensayos para simular el efecto de la adquisición de experiencia con la intervención

- Se especificó previamente que serían realizados análisis de subgrupo para distinguir el primer ensayo de cada investigador del segundo

En general, el azar mostró que el 'tratamiento DICE' no era significativamente mejor que el 'control', según se midieron las tasas de mortalidad. Sin embargo, en un análisis de subgrupo buscando sólo los ensayos 'publicados' (utilizando un modelo de sesgo de publicación tomado de ensayos reales) llevados a cabo por operadores 'experimentados' (sólo los segundos ensayos), hubo una reducción significativa del 23% en la mortalidad. De este modo, se pueden encontrar efectos significativos debidos sólo al azar

mejor entendido como generador de hipótesis para examinar en el futuro, que como proveedor de evidencia confiable sobre un subgrupo particular.

Fortalezas y debilidades potenciales de los metanálisis

Fortalezas

El metanálisis es una herramienta estadística con grandes fortalezas. La ‘magnitud del efecto’ es la estimación del efecto del tratamiento en un estudio (por ej. la relación de riesgo o la razón de probabilidades para los resultados dicotómicos y la diferencia promedio o diferencia media estandarizada para las evoluciones de variables continuas (Nikolakopoulou 2014), y las técnicas de metanálisis combinan datos de investigación de un número de estudios para proveer una estimación general de la magnitud del efecto de una manera fácilmente digerible. Los resultados de un metanálisis son presentados generalmente en un diagrama de bosque^a o ‘manchograma’^b (Cipriani 2006). La columna de la izquierda del diagrama enlista los nombres de los estudios (usualmente en orden cronológico) y la columna de la derecha muestra la magnitud del efecto para cada uno de ellos (a menudo representado por un cuadrado), incorporando intervalos de confianza representados por líneas horizontales. La medición metanalítica del efecto general se dibuja generalmente como un diamante, cuyos puntos laterales enmarcan los intervalos de confianza para dicha medición. Además, los métodos usados en el análisis evalúan la calidad de los estudios incluidos y de esta forma los revisores pueden indicar la fortaleza de la evidencia sumaria que reportan (Higgins 2011b).

Debilidades potenciales

La metodología de la revisión sistemática

Los resultados de un metanálisis deben interpretarse siempre con cuidado ya que su validez es dependiente de la metodología de la revisión sistemática original. Si ésta no fue adecuadamente conducida los resultados del metanálisis estarán sesgados. Cuando se lee una revisión sistemática es importante ser capaz de evaluar sus méritos, ya que no todas utilizan la misma metodología (cuadro 2). La intensidad con la que el sesgo ha sido controlado da una medida de la validez interna del estudio. La validez externa (o capacidad de generalización) provee una medida sobre hasta qué punto los resultados confieren una base correcta para la generalización a otras circunstancias.

CUADRO 2 Cómo evaluar lo méritos de una revisión sistemática y metanálisis

- ¿Cuáles son las afiliaciones de los autores y el apoyo financiero?
- ¿Cuáles son los métodos usados para identificar y seleccionar los estudios primarios sobre los que se basa la revisión?
- ¿Cuál fue la calidad de los estudios primarios?
- ¿El análisis y síntesis fueron apropiados?
- Se tomaron en cuenta fuentes probables de sesgo?
- ¿Cuál fue la significación estadística y clínica de los resultados?
- ¿Ha habido una actualización de la búsqueda de literatura?

La calidad de los estudios primarios

Los resultados del análisis también serán afectados por la calidad de los estudios primarios. Si la calidad es pobre, en consecuencia no será posible lograr resultados significativos del metanálisis: ‘lo que pones es lo que recibes’. Un metanálisis requiere determinar hasta qué punto las variaciones en la calidad de los estudios afectan la decisión de combinar los datos.

Se han propuesto muchas herramientas para evaluar la calidad de los estudios en una revisión sistemática y metanálisis. La mayoría de los instrumentos son ya sea escalas, en las que se puntúan varios componentes de calidad y se combinan para dar un puntaje general, o listas de verificación en las que se responden preguntas específicas (Jüni 2001). Muchos instrumentos contienen no sólo ítems basados en los criterios generalmente aceptados para la calidad de la metodología (aleatorización, ocultamiento de la asignación, enmascaramiento/procedimiento ciego), sino también ítems que no están directamente relacionados con la validez interna, tales como la presencia de un cálculo de poder estadístico (que se relaciona más bien con la precisión de los resultados) o si los criterios de inclusión y exclusión están claramente descritos (que se relaciona más con la aplicabilidad que con la validez) (Moher 1995). Posiblemente el mejor ejemplo de métodos usados para evaluar calidad en los ECAs es CONSORT (www.consort-statement.org), pero hay métodos diferentes para otros diseños de investigación. Estos incluyen a QUADAS (www.bris.ac.uk/quadas) para estudios de precisión de métodos diagnósticos, STROBE (www.strobe-statement.org) para estudios

a. ‘forest plot’ en el original, a menudo incorrectamente traducido como ‘diagrama de Forest’ (nota del T.)

b. ‘blobbogram’ en el original (nota del T.)

observacionales y TREND (www.cdc.gov/trendstatement) para estudios no aleatorizados.

Estas herramientas varían y algunas se focalizan más en la calidad de los reportes que en la metodología subyacente de los estudios. Para abordar este problema, la Colaboración Cochrane recomienda evaluar la calidad de los estudios utilizando su instrumento ‘riesgo de sesgo’, que no es ni una escala ni una lista de verificación. Es una evaluación basada en dominios, en la que se hacen valoraciones críticas para diferentes cuestiones vinculadas a los estudios en forma separada: generación de la secuencia aleatoria; ocultamiento de la asignación; enmascaramiento/procedimiento ciego de los participantes y personal; enmascaramiento/procedimiento ciego de la evaluación del resultado; datos del resultado incompletos; reporte selectivo y otras fuentes de sesgo (Higgins 2011b).

Abordando la heterogeneidad clínica y estadística de los estudios

Los estudios siempre varían, por ejemplo en términos de los tipos de participantes involucrados, los métodos utilizados, los tipos de intervención usados como comparadores, la longitud del seguimiento y los resultados que se miden. En consecuencia, será necesaria una selección para incluir los estudios. Para evitar el sesgo, antes de comenzar la revisión es muy importante especificar en el protocolo los principales criterios de selección de los estudios. Los revisores necesitan evitar una inclusión excesiva de estudios dispares, pero también evitar un exceso de exclusión de estudios que tengan datos relevantes. No obstante, aún si los criterios de inclusión y exclusión son claros y coherentes, a veces los estudios incluidos difieren significativamente. Esta ‘heterogeneidad’ puede presentar desafíos. Puede no ser posible combinar los resultados y realizar un metanálisis; la presencia de una heterogeneidad significativa ha sido asemejada al error de ‘combinar naranjas con manzanas’ (Eysenck 1994). Aun si ha sido posible combinar los estudios, puede también hallarse heterogeneidad durante el análisis. Si así sucede, se recomienda un modelo de efectos aleatorios, ya que este reconoce que las diferencias observadas en las magnitudes de los efectos entre los diferentes estudios reflejan tanto heterogeneidad verdadera como errores aleatorios (Nikolakopoulou 2014). Por esta razón, las estimaciones combinadas de tal análisis tiene intervalos de confianza más amplios y los resultados son más conservadores que el análisis de efectos fijos.

Un ejemplo de las dificultades que involucra la consideración de la heterogeneidad de los estudios

en Psiquiatría es el asunto de la efectividad del tratamiento comunitario, sea intensivo o estándar, para mejorar la evolución de los pacientes. Las revisiones sistemáticas que abarcan esta cuestión lidian todas con asuntos similares. Por ejemplo, las definiciones de ‘tratamiento comunitario’ y ‘tratamiento de control’ varían significativamente entre los centros que conducen los ensayos y han cambiado a lo largo del tiempo en que dichos estudios se condujeron. Los servicios e intervenciones complejas en Salud Mental son difíciles de estandarizar, también los términos ‘cuidado estándar’ y ‘servicios usuales’ que se utilizan como tratamientos de control están a menudo vagamente definidos y se pueden superponer con el tratamiento activo. Además, los estudios han diferido en la elección del mejor indicador de resultados, utilizando distintas medidas (Dieterich 2010). Un enfoque (ej. Marshall 2000a,b) es respetar los términos (tales como tratamiento comunitario asertivo, manejo de casos, cuidado estándar) usados para cada arma de tratamiento por los investigadores de los estudios originales. Esta es una solución práctica, pero puede también enmascarar una ‘heterogeneidad clínica’ en las diferentes armas de tratamiento. Algunas revisiones (por ej. Murphy 2012) sólo han encontrado un número pequeño de estudios que cumplieran con sus criterios. Además, muchos estudios en esta área tienen muestras pequeñas (por ej. Malone 2007), otorgando un poder insuficiente para detectar diferencias estadísticamente significativas en la evolución, confirmando ‘heterogeneidad estadística’. Catty et al (2002) utilizaron criterios de inclusión más amplios para incluir más estudios y aumentar el tamaño general de la muestra. Incluyeron todos los estudios de ‘tratamiento domiciliario’, que abarcaban cualquier tratamiento fuera del hospital. A pesar de estos amplios criterios de inclusión y la elección de una sola medida de resultados (días hospitalizado), y un seguimiento intensivo de los autores de los estudios primarios, encontraron que solamente 57% de los estudios brindaban datos que eran utilizables en su metanálisis. Esto es típico de las revisiones sistemáticas en esta área y limita severamente su generalización.

Resumen

Los resultados de un metanálisis dependen no sólo de la metodología usada en la revisión sistemática y metanálisis sino también de la calidad de las fuentes primarias de datos. Las revisiones sistemáticas y metanálisis sobre el mismo tópico pueden producir resultados conflictivos. Por ejemplo, desde la publicación del estudio pionero

de Caspi y colegas (Caspi 2003) que sugirió que el gen del transportador de serotonina modifica la relación entre los eventos vitales estresantes y la depresión, un número de estudios individuales sobre el tema fueron conducidos. Los metanálisis de dichos estudios han sido contradictorios, con algunos (ej. Risch 2009) que no confirmaron tal relación gen-ambiente y otros que sí (Ej. Karg 2011; Miller 2013). Así, aunque el metanálisis es probablemente la herramienta más robusta actualmente disponible para resumir la evidencia, los resultados raramente son inequívocos y siempre necesitan una cuidadosa valoración e interpretación.

Sesgos en las revisiones sistemáticas y metanálisis

El sesgo puede ocurrir durante la selección, evaluación o síntesis de los datos y debe ser evitado, dado que brinda resultados inadecuados o equívocos. Los tipos de sesgos se resumen en el cuadro 3.

Una fuente importante de sesgo en las revisiones sistemáticas es el sesgo de publicación, que ocurre como resultado de la tendencia de los autores, revisores y editores de publicar preferencialmente los estudios que tienen un resultado claramente definido y estadísticamente significativo (Mavridis 2014). Los estudios en los que el tratamiento tiene un efecto similar o menor que el placebo, o que el tratamiento actual bien establecido, tienen menor probabilidad de ser publicados. La investigación con financiación pública es más probable que sea publicada con cualquiera sea el resultado, mientras que la investigación pagada comercialmente muestra un sesgo significativo hacia la publicación cuando los hallazgos son positivos (Dickersin 1990). Un metanálisis basado puramente en los resultados publicados bien puede ser engañoso ya que el conjunto de datos publicados pueden no ser representativos de la evidencia total (Higgins 2011b). Por ejemplo, Turner et al (2008) obtuvieron revisiones de estudios no publicados de antidepresivos presentados para la aprobación regulatoria de la Food and Drug Administration de Estados Unidos (FDA). Los autores confrontaron los resultados de los reportes no publicados con las publicaciones correspondientes, si estaban disponibles. De interés, el 31% de los estudios no estaban publicados. Era mucho más probable que los resultados positivos fueran publicados y, de los estudios negativos que fueron publicados, la mayoría fueron presentados de una manera que transmitía un resultado positivo. Como resultado del reporte selectivo, la literatura publicada confería una magnitud del efecto cerca de un tercio

CUADRO 3 Tipos de error y sesgos en las revisiones sistemáticas y metanálisis

- Calidad pobre de los estudios primarios (lo que tiende a sobre representar un resultado favorable)
- Reporte selectivo de los estudios primarios (generalmente con resultados significativos y favorables)
- Sesgo en la selección de los estudios a incluir para la revisión:
 - sesgo de publicación – los grandes estudios positivos es más probable que estén publicados
 - sesgo de lenguaje – es más probable que los revisores seleccionen artículos en inglés
 - Es más probable que se identifiquen los estudios listados en las bases de datos electrónicas
 - Las preferencias de los revisores para seleccionar los estudios a incluir
- Sesgo en los métodos estadísticos para extraer y combinar los datos
- Sesgo en los supuestos/simplificaciones de los autores al extraer y/o combinar los datos
- Sesgo de financiación/patrocinio (que tiende a favorecer la rama de tratamiento financiada por el patrocinante)

mayor que la que se derivaba de los datos de la FDA. Whittington et al (2004) también resaltaron las diferentes recomendaciones para la práctica de la prescripción que pueden ser deducidas de analizar sólo los datos publicados. Utilizaron el ejemplo de estudios de los inhibidores selectivos de la recaptación de serotonina (ISRSs) versus placebo en el tratamiento de la depresión de niños y jóvenes de entre 5 y 18 años. Como la mayoría de las decisiones clínicas balancean eficacia contra riesgo de daño/efectos colaterales, el no reporte de estudios negativos puede engendrar una diferencia significativa. Cuando se lee un metanálisis, es importante comprobar si los autores buscaron estudios y datos suplementarios no publicados.

Para abordar el reporte subóptimo en metanálisis, de problemas metodológicos tales como el sesgo potencial de publicación, un grupo internacional desarrolló una guía o declaración llamada Calidad de los Reportes de Metanálisis (Quality of Reporting of Meta-Analysis Statement, QUOROM), que se centra en los reportes de metanálisis de ECAs (Moher1999). Más recientemente estas guías han sido revisadas y renombradas Ítems Preferenciales de Reporte para las Revisiones Sistemáticas y Metanálisis (Preferred Reporting Items for Systematic Reviews and Meta-Analyses, PRISMA) (Moher

2009). Las guías PRISMA contienen una lista de comprobación para evaluar los elementos diversos de calidad de una revisión sistemática (con o sin metanálisis) y también para guiar a los autores cuando comunican sus hallazgos. Luego de la publicación de la declaración PRISMA, el Centro para Revisiones y Disseminación del Reino Unido (CRD por sus siglas en inglés) de la Universidad de York desarrolló PROSPERO, o Prospective Register of Systematic Reviews with Health-Related Outcomes (Registro prospectivo de revisiones sistemáticas con resultados relacionados a la salud) (Booth 2012; www.crd.york.ac.uk/PROSPERO). Los objetivos son reducir la duplicación no planeada de revisiones y proveer transparencia en el proceso de revisión, tendiendo a minimizar el sesgo de reporte.

Métodos para reducir los efectos de los sesgos

El ejemplo del sesgo de publicación

Probablemente la estrategia más efectiva contra el sesgo de publicación es la prevención (un método prospectivo). Una orientación es crear registros de ensayos en los que los detalles de los mismos se registran antes de que comiencen, para capturar datos de todos los estudios, sean eventualmente publicados o no (De Angelis 2004). Otra sugerencia es una amnistía, mediante la cual los investigadores son alentados a someter a publicación los reportes de ensayos previamente no publicados (Horton 1997). Sin embargo, estos sistemas son difíciles de implementar y muchos ensayos son previos a los registros en varios años y sus datos no están disponibles al público (Goldacre 2013). En general, aunque las estrategias prospectivas pueden reducir el problema del sesgo de publicación en el futuro, probablemente permanezca como un asunto que será necesario abordar con una extensión mayor o menor en todos los metanálisis en un futuro próximo.

Los métodos retrospectivos tratan de compensar el sesgo de publicación luego del evento. Por ejemplo, los revisores deben hacer todos los esfuerzos para encontrar todos los datos disponibles, incluyendo estudios no publicados y estudios publicados en otros idiomas. Además de las búsquedas electrónicas, deben buscar manualmente, revisar referencias y abstractos de conferencias y comunicarse directamente con los autores. Otras fuentes de datos incluyen los sitios web de las agencias regulatorias (ej. la FDA (www.fda.gov), la Agencia Europea de Medicinas (www.ema.europa.eu)), compañías farmacéuticas (por ejemplo el GlaxoSmithKline Clinical

Study Register (www.gsk-clinicalstudyregister.eu) y organizaciones independientes como la Organización Mundial de la Salud (http://apps.who.int/trialsearch), ClinicalTrials.gov (https://clinicaltrials.gov) y la European Union Clinical Trials Register (www.clinicaltrialsregister.eu). Sin embargo, los esfuerzos para incluir datos no publicados pueden resultar un arma de doble filo. Los datos pueden ser difíciles de recuperar o pueden estar incompletos, no representativos de la muestra bajo estudio y pueden no haber sido objeto de revisión por pares. Los métodos de un metanálisis deben reconocer que, a pesar de los mejores esfuerzos de los revisores, es probable que haya un grado de sesgo de publicación en los estudios seleccionados para la revisión sistemática.

Los investigadores tratan de detectar el sesgo de publicación utilizando un número de pruebas estadísticas (ej. test de Egger y gráficos de embudo) que se basan en la teoría subyacente que los estudios con un tamaño pequeño de la muestra están más predispuestos al sesgo de publicación, mientras que los estudios más grandes es más probable que sean publicados no importa su hallazgo (Egger 1997). En un gráfico de embudo las magnitudes de los efectos se dibujan en el eje horizontal contra una medida del tamaño de cada estudio (ej. error estándar o tamaño de la muestra) en el eje vertical. De esta manera, es común que un metanálisis muestre dicho gráfico y lleve a cabo procedimientos tales como el método 'podar y completar'^c para identificar asimetría y ajustarla (Duval 2000). La asimetría es interpretada a menudo como revelando evidencia directa de la presencia del sesgo de publicación. Sin embargo, esto es demasiado simplista: la simetría puede resultar también de una diferencia esencial (o heterogeneidad) entre estudios grandes y pequeños (Lau 2006). Por ejemplo, los estudios pequeños pueden centrarse en pacientes de alto riesgo, para los que el tratamiento puede ser más efectivo; o pueden tener un seguimiento menor. Variaciones en la calidad también afectan la forma del gráfico de embudo, con estudios más pequeños y de menor calidad mostrando beneficios mayores del tratamiento.

Ejemplos de metodología avanzada

Metanálisis de datos de pacientes individuales

Como ha sido mencionado (ver 'metanálisis' más arriba), los análisis de subgrupos dentro de un metanálisis estándar tienen limitaciones significativas. El metanálisis de datos de pacientes individuales (MADPI) es un abordaje potencialmente útil en el que el metanálisis es

c. 'trim and fill' (nota del T.)

conducido usando individualmente los datos de los pacientes de los estudios primarios (Clarke 2005). Esto permite un análisis de subgrupos más adecuado ya que éstos pueden estar basados en clasificaciones comunes de subgrupos entre los estudios. Es crucial que el metanálisis preserve el agrupamiento original de los pacientes en los estudios: es inapropiado analizar los datos de todos los pacientes como si todos hubieran participado en un mismo estudio. No obstante, un análisis apropiado puede producir resultados que informen la práctica basada en la evidencia, como una estimación del efecto del tratamiento a través de la combinación de todos los estudios, cómo el efecto del tratamiento varía entre los mismos (ej. con la dosis o la localización del estudio) y varía entre tipos de pacientes (ej. agrupados por edad o estadio de la enfermedad). El MADPI tiene muchas ventajas potenciales sobre los metanálisis que usan datos agregados, donde los datos están a veces pobremente reportados, no disponibles o se presentan en forma diferente según el estudio (Riley 2010). El uso de los datos individuales estandariza los métodos del estudio y provee a menudo de datos extra (por ej. seguimientos más largos, más medidas de resultados) que no han sido incluidos en la publicación original en donde estaban agregados. Sin embargo, el MADPI es un abordaje que consume mucho tiempo y recursos, tanto para los revisores como para los autores de los estudios originales; requiere métodos estadísticos avanzados y los datos originales bien pueden ser pobres o faltar. Todavía no ha sido ampliamente utilizado en Psiquiatría, aunque hay algunos ejemplos de cómo el MADPI puede ayudar a los clínicos a sopesar los beneficios del tratamiento psiquiátrico en el paciente individual (ej. Furukawa 2015). La propuesta de Tudur Smith y colegas (2014) de comenzar un repositorio central de datos de pacientes individuales de los ensayos reduciría sustancialmente el tiempo requerido para coleccionar los datos originales.

Metanálisis en red^d

Los metanálisis usan como técnica estadística estándar comparaciones de los tratamientos de a pares, o apareadas. Esto significa que cuando revisan los datos sobre la eficacia de todos los tratamientos disponibles para una condición particular, a los clínicos se les presenta un despliegue de comparaciones de a pares, mientras que ellos más bien compararían la eficacia relativa de todos los tratamientos simultáneamente. Además, algunas comparaciones entre tratamientos no han sido efectuadas directamente y por lo tanto no hay datos directos sobre los que

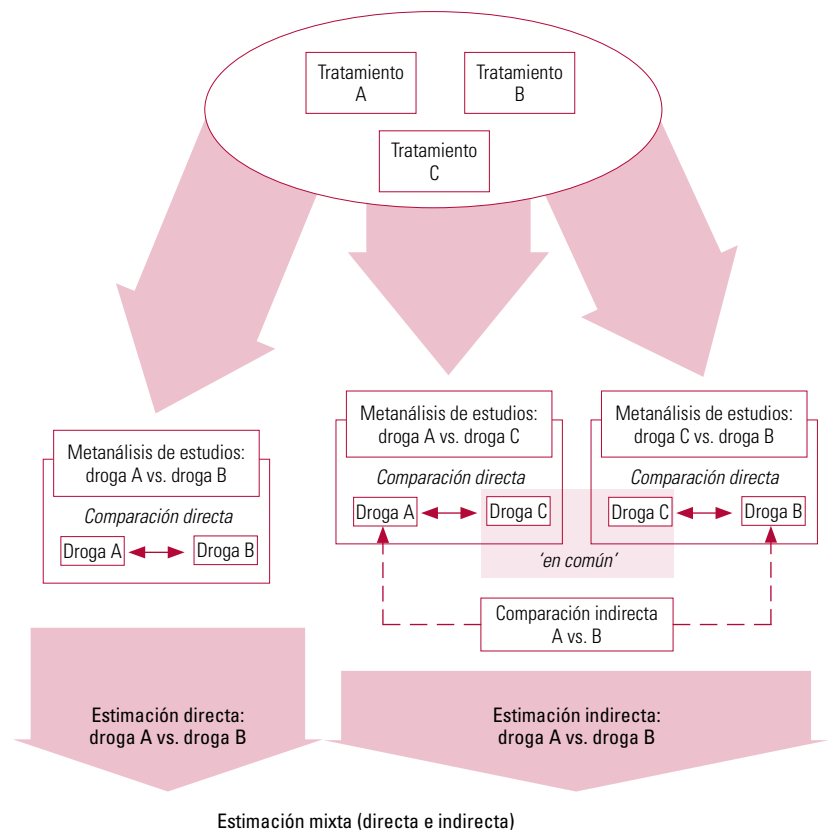


FIG 1 1 Combinación de evidencia directa e indirecta para estimar una magnitud del efecto única para tratamiento A vs. tratamiento B (estimación mixta).

basar una comparación apareada. El metanálisis en red (MER) (también llamado metanálisis de tratamientos múltiples o comparación mixta de tratamientos) es un método estadístico que puede llenar este vacío ya que permite que sean evaluados múltiples tratamientos al mismo tiempo, utilizando evidencia directa e indirecta de los datos disponibles de comparaciones (Caldwell 2005). La evidencia indirecta proviene de inferir la eficacia relativa de dos drogas que no han sido comparadas directamente entre ellas, pero que ambas han sido directamente comparadas con la misma droga de comparación. De esta forma, como en el ejemplo de la Fig. 1, si hay ensayos de droga A vs. droga B, esto nos da información comparativa directa de las respectivas eficacias. Los ensayos de droga A vs. C y B vs. C nos pueden proveer también de evidencia indirecta de la eficacia relativa de A vs. B. El uso de evidencia indirecta cumple dos funciones: provee datos de comparaciones para las que no existen ensayos y mejora la precisión de los datos directos al añadir indirectos (y de esta manera reducir la amplitud del intervalo de confianza de la estimación que se brinda) (Cipriani 2013b).

d. 'Network meta-analysis' en el original (nota del T.)

Los MERs tienen un rol útil no sólo en fortalecer la base de evidencias sino también en la provisión de un ranking de tratamientos para trastornos específicos de acuerdo con un resultado de interés, por ejemplo eficacia y aceptabilidad. Esto permite un resumen de todos los tratamientos para los que hay evidencia disponible, tanto indirecta como directa. Permite que sean ordenados jerárquicamente, generando una tabla (similar a la tabla de millas en un atlas caminero) que muestra la eficacia y tolerabilidad relativa de cada agente. Ejemplos de revisiones MER con metodología robusta incluyen: drogas antimaníacas en la manía aguda (Cipriani 2011; Yildiz 2015), tratamientos de mantenimientos para el trastorno bipolar (Miura 2014), antidepresivos para el tratamiento agudo del trastorno depresivo unipolar (Cipriani 2009), agentes potenciadores en la depresión resistente (Zhou 2015a), psicoterapias para la depresión en niños y adolescentes (Zhou 2015b), tratamientos para la ansiedad social (Mayo-Wilson 2014) y drogas antipsicóticas para el tratamiento agudo de la esquizofrenia (Leucht 2013). Las ventajas de este enfoque son claras y la información es fácil de entender y de aplicar en la práctica clínica. De esta manera, los MERs han sido crecientemente empleados para apoyar guías clínicas y abordajes de tecnología de la salud (Barbui 2011). Sin embargo, a pesar de las ventajas, no son todavía una práctica establecida. Se han expresado algunas preocupaciones sobre la validez de los métodos empleados. Aunque se utiliza evidencia aleatorizada y la evidencia indirecta preserva la randomización original, la evidencia indirecta no es por sí misma evidencia aleatorizada ya que los tratamientos han sido comparados originariamente entre estudios pero no a través de ellos, y tal comparación puede entonces estar sujeta a sesgos. De este modo, la evidencia directa es más robusta y la indirecta debe ser idealmente utilizada como un suplemento de la directa. No obstante, en la mayoría de los casos las evidencias directas e indirectas concuerdan (Song 2008).

Conclusiones

La Medicina basada en la evidencia se ha desarrollado sustancialmente en las últimas décadas. Inicialmente, el énfasis estuvo en proveer la mejor evidencia disponible para contestar preguntas terapéuticas específicas. Se ha empleado mucho tiempo y esfuerzo de la mejor manera, incorporando la metodología más rigurosa para proveer tal evidencia. Para establecer los beneficios y seguridad de las intervenciones es esencial generar, resumir y entender la mejor evidencia disponible. Las revisiones sistemáticas,

incluyendo a menudo metanálisis, se han tornado una herramienta valiosa a tales fines.

Sin embargo, las revisiones sistemáticas como diseño de estudios tienen limitaciones y es necesario abordar un número de asuntos antes de implementar la síntesis de evidencia en la práctica clínica (Berlin 2014). La heterogeneidad clínica de los pacientes psiquiátricos y la calidad a veces variable de los estudios primarios hacen que algunas revisiones sean difíciles de interpretar y de utilizar. A menudo las preguntas que surgen en la clínica son mucho más complejas que las que responden las revisiones sistemáticas. Los clínicos (pero también los investigadores, los que desarrollan guías, los editores de revistas y los lectores críticos de la literatura) deben ser conscientes de esto, porque el entendimiento de las limitaciones y del potencial de la evidencia metanalítica es crucial para desarrollar un cuidado mejor de los pacientes. Los clínicos necesitan desplegar las habilidades que se requieren para sentirse confiados al utilizar la práctica basada en la evidencia en su abordaje de las cuestiones clínicas cotidianas. Publicaciones tales como *Evidence-Based Mental Health* (<http://ebmh.bmj.com>) están cambiando de alguna manera su enfoque para encarar estas necesidades (Cipriani 2014), incluyendo por ejemplo ‘conferencias clínicas’ online en tiempo real vía Google Hangout, que usan prácticas basadas en la evidencia para abordar cuestiones clínicas complejas en forma práctica, y una sección estadística regular (un área a menudo dejada de lado cuando se leen investigaciones). La práctica basada en la evidencia continuará generando desafíos para los clínicos, pero a medida que ganen confianza en las técnicas requeridas y las incorporen en la práctica de rutina, estos desafíos los llevarán a cosechar frutos que bien valen el esfuerzo.

Agradecimientos

K.A.S. y A.C. agradecen el apoyo del National Institute for Health Research (NIHR) Oxford cognitive health Clinical Research Facility. J.G. es un Investigador Principal del NIHR. La preparación de este artículo fue financiada por la NIHR Collaboration for Leadership in Applied Health Research and Care Oxford. Las opiniones aquí expresadas son de los autores y no necesariamente del NHS, NIHR o del Department of Health.

Referencias

- Barbui C, Cipriani A (2011) What are evidence-based treatment recommendations? *Epidemiology and Psychiatric Sciences*, **20**: 29–31.
- Berlin JA, Golub RM (2014) Meta-analysis as evidence: building a better pyramid. *JAMA*, **312**: 603–5.

Respuestas

1 e 2 c 3 b 4 d 5 a

- Booth A, Clarke M, Dooley G, et al (2012) The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. *Systematic Reviews*, 1: 2.
- Caldwell DM, Ades AE, Higgins JP (2005) Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ*, 331: 897–900.
- Caspi A, Sugden K, Moffitt TE, et al (2003) Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science*, 301: 386–9.
- Catty J, Burns T, Knapp M, et al (2002) Home treatment for mental health problems: a systematic review. *Psychological Medicine*, 32: 383–401.
- Cipriani A, Barbui C (2006) What is a forest plot? *Epidemiologia e Psichiatria Sociale*, 15: 258–9.
- Cipriani A, Furukawa TA, Salanti G, et al (2009) Comparative efficacy and acceptability of 12 new-generation antidepressants: a multiple-treatments meta-analysis. *Lancet*, 373: 746–58.
- Cipriani A, Barbui C, Salanti G, et al (2011) Comparative efficacy and acceptability of antimanic drugs in acute mania: a multiple-treatments meta-analysis. *Lancet*, 378: 1306–15.
- Cipriani A (2013a) Time to abandon *Evidence Based Medicine? Evidence-Based Mental Health*, 16: 91–2.
- Cipriani A, Higgins JP, Geddes JR, et al (2013b) Conceptual and technical challenges in network meta-analysis. *Annals of Internal Medicine*, 159: 130–7.
- Cipriani A, Furukawa TA (2014) Advancing evidence-based practice to improve patient care. *Evidence-Based Mental Health*, 17: 1–2.
- Clarke MJ (2005) Individual patient data meta-analyses. *Best Practice & Research Clinical Obstetrics & Gynaecology*, 19: 47–55.
- Cook DJ, Sackett DL, Spitzer WO (1995) Methodologic guidelines for systematic reviews of randomized control trials in health care from the Potsdam Consultation on Meta-Analysis. *Journal of Clinical Epidemiology*, 48: 167–71.
- Counsell CE, Clarke MJ, Slattery J, et al (1994) The miracle of DICE therapy for acute stroke: fact or fictional product of subgroup analysis? *BMJ*, 309: 1677–81.
- De Angelis C, Drazen JM, Frizelle FA, et al (2004) Clinical trial registration: a statement from the International Committee of Medical Journal Editors. *Lancet*, 364: 911–2.
- Dickersin K (1990) The existence of publication bias and risk factors for its occurrence. *JAMA*, 263: 1385–9.
- Dieterich M, Irving CB, Park B, et al (2010) Intensive case management for severe mental illness. *Cochrane Database of Systematic Reviews*, 10: CD007906.
- Duval S, Tweedie R (2000) Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56: 455–63.
- Egger M, Davey Smith G, Schneider M, et al (1997) Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 315: 629–34.
- Eysenck HJ (1994) Meta-analysis and its problems. *BMJ*, 309: 789–92.
- Furukawa TA, Levine SZ, Tanaka S, et al (2015) Initial severity of schizophrenia and efficacy of antipsychotics: participant-level meta-analysis of 6 placebo-controlled studies. *JAMA Psychiatry*, 72: 14–21.
- Geddes JR, Wilczynski N, Reynolds S, et al (1999) Evidence-based mental health – the first year. *Evidence-Based Mental Health*, 2: 4–5.
- General Medical Council (2013) Duties of a doctor. In *Good Medical Practice*. GMC.
- Goldacre B (2013) Are clinical trial data shared sufficiently today? No. *BMJ*, 347: f1880.
- Guyatt GH, Haynes RB, Jaeschke RZ, et al (2000) Users' Guides to the Medical Literature: XXV. Evidence-based medicine: principles for applying the users' guides to patient care. *JAMA*, 284: 1290–6.
- Guyatt GH, Oxman AD, Vist GE, et al (2008) GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*, 336: 924–6.
- Higgins JPT, Green S (eds) (2011a) Glossary. In *Cochrane Handbook for Systematic Reviews of Interventions: Version 5.1.0*. The Cochrane Collaboration (<http://community.cochrane.org/glossary>).
- Higgins JPT, Altman DG, Gøtzsche PC, et al (2011b) The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*, 343: d5928.
- Horton R (1997) Medical editors' trial amnesty. *Lancet*, 350: 756.
- Jüni P, Altman DG, Egger M (2001) Assessing the quality of controlled clinical trials. *BMJ*, 323: 42–6.
- Karg K, Burmeister M, Shedden K, et al (2011) The serotonin transporter promoter variant (5-HTTLPR), stress, and depression meta-analysis revisited: evidence of genetic moderation. *Archives of General Psychiatry*, 68: 444–54.
- Lau J, Ioannidis JP, Terrin N, et al (2006) The case of the misleading funnel plot. *BMJ*, 333: 597–600.
- Leucht S, Cipriani A, Spineli L, et al (2013) Comparative efficacy and tolerability of 15 antipsychotic drugs in schizophrenia: a multiple-treatments meta-analysis. *Lancet*, 382: 951–62.
- Malone D, Marriott S, Newton-Howes G, et al (2007) Community mental health teams (CMHTs) for people with severe mental illnesses and disordered personality. *Cochrane Database of Systematic Reviews*, 3: CD000270.
- Marshall M, Gray A, Lockwood A, et al (2000a) Case management for people with severe mental disorders. *Cochrane Database of Systematic Reviews*, 2: CD000050.
- Marshall M, Lockwood A (2000b) Assertive community treatment for people with severe mental disorders. *Cochrane Database of Systematic Reviews*, 2: CD001089.
- Mavridis D, Salanti G (2014) How to assess publication bias: funnel plot, trim-and-fill method and selection models. *Evidence-Based Mental Health*, 17: 11–5.
- Mayo-Wilson E, Dias S, Mavranzeouli I, et al (2014) Psychological and pharmacological interventions for social anxiety disorder in adults: a systematic review and network meta-analysis. *Lancet Psychiatry*, 1: 368–76.
- Miller R, Wankerl M, Stalder T, et al (2013) The serotonin transporter gene-linked polymorphic region (5-HTTLPR) and cortisol stress reactivity: a meta-analysis. *Molecular Psychiatry*, 18: 1018–24.
- Miura T, Noma H, Furukawa TA, et al (2014) Comparative efficacy and tolerability of pharmacological treatments in the maintenance treatment of bipolar disorder: a systematic review and network meta-analysis. *Lancet Psychiatry*, 1: 351–9.
- Moher D, Olkin I (1995) Meta-analysis of randomized controlled trials: a concern for standards. *JAMA*, 274: 1962–4.
- Moher D, Cook DJ, Eastwood S, et al (1999) Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *Lancet*, 354: 1896–900.
- Moher D, Liberati A, Tetzlaff J, et al (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine*, 6: e1000097.
- Murphy S, Irving CB, Adams CE, et al (2012) Crisis intervention for people with severe mental illnesses. *Cochrane Database of Systematic Reviews*, 5: CD001087.
- Nikolakopoulou A, Mavridis D, Salanti G (2014) Demystifying fixed and random effects meta-analysis. *Evidence-Based Mental Health*, 17: 53–7.
- Riley RD, Lambert PC, Abo-Zaid G (2010) Meta-analysis of individual participant data: rationale, conduct, and reporting. *BMJ*, 340: c221.
- Risch N, Herrell R, Lehner T, et al (2009) Interaction between the serotonin transporter gene (5-HTTLPR), stressful life events, and risk of depression: a meta-analysis. *JAMA*, 301: 2462–71.
- Song F, Harvey I, Lilford R (2008) Adjusted indirect comparison may be less biased than direct comparison for evaluating new pharmaceutical interventions. *Journal of Clinical Epidemiology*, 61: 455–63.

Tudur Smith C, Dwan K, Altman DG, et al (2014) Sharing individual participant data from clinical trials: an opinion survey regarding the establishment of a central repository. *PLoS ONE*, **9**: e97886.

Turner EH, Matthews AM, Linardatos E, et al (2008) Selective publication of antidepressant trials and its influence on apparent efficacy. *New England Journal of Medicine*, **358**: 252–60.

Whittington CJ, Kendall T, Fonagy P, et al (2004) Selective serotonin reuptake inhibitors in childhood depression: systematic review of published versus unpublished data. *Lancet*, **363**: 1341–5.

Yildiz A, Nikodem M, Vieta E, et al (2015) A network meta-analysis on comparative efficacy and all-cause discontinuation of antimanic treatments in acute bipolar mania. *Psychological Medicine*, **45**: 299–317.

Zhou X, Ravindran AV, Qin B, et al (2015a) Comparative efficacy, acceptability, and tolerability of augmentation agents in treatment-resistant depression: systematic review and network meta-analysis. *Journal of Clinical Psychiatry*, **76**: e487–98.

Zhou X, Hetrick SE, Cuijpers P, et al (2015b) Comparative efficacy and acceptability of psychotherapies for depression in children and adolescents: A systematic review and network meta-analysis. *World Psychiatry*, **14**: 207–22.

Preguntas de elección múltiple

Seleccione la mejor opción, única para cada pregunta

1 ¿Cuál de estas afirmaciones sobre una revisión sistemática bien conducida no es cierta?

- a se identifican estudios sobre un tópico inespecífico
- b se incluyen los estudios que cumplen con criterios predefinidos
- c Los métodos utilizados para evaluar y sintetizar los datos están claramente definidos
- d la revisión se actualiza regularmente usando los criterios originales
- e los métodos sistemáticos y explícitos utilizados eliminan la posibilidad de sesgo.

2 ¿Cuál de las siguientes no es cierta en relación con los sesgos?

- a las revisiones narrativas tienen un riesgo de sesgo mayor que las sistemáticas
- b uno de los objetivos de guías tales como PRISMA es minimizar el sesgo de publicación
- c el uso de un procedimiento ciego por los evaluadores de los resultados puede ayudar a superar el sesgo de selección

d el metanálisis de una revisión sistemática sesgada también estará sesgado

e la Colaboración Cochrane recomienda una herramienta para sesgos basada en dominios.

3 ¿Cuál de las siguientes no es cierta en relación con los metanálisis?

- a combina estadísticamente datos de diferentes estudios para dar una estimación general de la magnitud del efecto con una muestra mayor
- b generalmente se aborda la heterogeneidad con un análisis de efectos fijos
- c es el uso de técnicas estadísticas para resumir datos cuantitativamente
- d los metanálisis sobre una misma cuestión puede dar conclusiones significativamente diferentes
- e sus resultados pueden ser representados en un diagrama de bosque.

4 ¿Cuál de las siguientes afirmaciones sobre los metanálisis de datos de pacientes individuales no es cierta?

- a pueden incluir datos obtenidos pero no reportados en los estudios originales
- b puede investigar cómo los efectos del tratamiento varían entre centros

c permiten análisis de subgrupos si los datos individuales están preservados en sus agrupaciones originales

d los métodos estadísticos son fáciles de usar y la recuperación de datos no ocupa mucho tiempo

e pueden analizar cómo los efectos del tratamiento varían en diferentes grupos de pacientes.

5 ¿Cuál de las siguientes no es cierta en relación con los metanálisis en red?

- a utilizan solamente evidencia indirecta para comparar tratamientos
- b puede hacerse un ranking de los tratamientos con relación a una variable específica, por ej. eficacia o tolerabilidad
- c los datos indirectos pueden proveer información cuando no existen comparaciones directas
- d los datos indirectos pueden agregarse a los directos para aumentar el tamaño de la muestra
- e los resultados pueden ser fácilmente entendidos por los clínicos y aplicados a la práctica clínica