

Revisiones sistemáticas de la literatura

Óscar A. Beltrán G. MD.

Ciertamente, se puede hacer una gran crítica a la profesión médica, por no haber organizado un resumen crítico por especialidad o subespecialidad, actualizando periódicamente, todos los ensayos clínicos controlados que sean relevantes (1).

Archie Cochrane, 1980

RESUMEN

Las Revisiones Sistemáticas son un diseño de investigación observacional y retrospectivo, que sintetiza los resultados de múltiples investigaciones primarias. Son parte esencial de la medicina basada en la evidencia por su rigurosa metodología, identificando los estudios relevantes para responder preguntas específicas de la práctica clínica. El término metanálisis se reserva para la combinación numérica de los datos. No todas las revisiones sistemáticas lo incluyen. Las revisiones sistemáticas son tan populares que han tenido un crecimiento vertiginoso en los últimos años. Este artículo revisará detenidamente la definición, las diferencias con las revisiones narrativas y el proceso para desarrollar e interpretar una revisión sistemática, incluyendo sus limitaciones.

Palabras Clave

Revisión Sistemática. Metanálisis.

SUMMARY

Systematic Review is a retrospective method of research and it has become increasingly accepted. Systematic Reviews are an essential part of evidence-based medicine as rigorous methodology can identify all relevant papers that address specific research questions. Metanalysis is a statistical technique for combining the findings from independent studies to provide a summary effect estimate. Systematic Review has become popular with a dramatic rise in papers over the last years. This paper will review the process that has led to the development of a systematic review.

Key Words

Systematic Review. Metanalysis.

INTRODUCCIÓN

En la práctica clínica se generan frecuentemente dudas en la toma de decisiones frente a los problemas

por los que consultan los pacientes. Pero dada la escasez de tiempo, la proliferación de información, la falta de conocimientos básicos de epidemiología y bioestadística, la Medicina Basada en Evidencia (MBE), como estrategia metodológica, ha desarrollado aspectos prácticos para la búsqueda de la evidencia y solución de nuestras preguntas, y uno de

Internista, Gastroenterólogo. Epidemiólogo Clínico (c), Clínica San Pedro Claver. Clínica Juan N. Corpas.

sus aportes es la posibilidad de trabajar con bases de datos.

Los sistemas de indización de los trabajos de investigación y las bases de datos bibliográficas (Medline, Embase, Lilacs, etc.) nos han permitido superar las dificultades de acceso a la literatura existente. Sin embargo, estos mismos sistemas nos han creado la falsa impresión de que estas fuentes contienen toda la evidencia existente, cuando en verdad no es así: por ejemplo Medline no referencia muchos de los excelentes trabajos publicados en otro idioma diferente del inglés, su indexación es confusa y se ha estimado que una búsqueda electrónica en Medline sólo conseguirá identificar el 60%-70% de los ensayos clínicos realmente incluidos en esta base de datos (2).

La imposibilidad de superar los problemas que se derivan del insuficiente acceso y manejo de la información científica ha generado que progresivamente hayan cobrado mayor importancia las revisiones de la literatura científica llevadas a cabo por personas ajenas al clínico o persona interesada. En el fondo, se delega la necesidad individual de disponer de la información existente en algún grupo o persona con mayor capacidad de obtenerla y de sintetizarla.

¿POR QUÉ SE REQUIEREN LAS REVISIONES SISTEMÁTICAS?

En primer lugar, cuando los clínicos consultamos la literatura publicada, frecuentemente, no nos conformamos con la lectura de un artículo, porque la práctica nos ha enseñado que, hallazgos reportados por un autor son cuestionados por otros o por el contrario pueden ser confirmados por repetidas investigaciones. Pero, rara vez, los estudios de investigación, aun los más rigurosos, nos dan respuestas definitivas a nuestras preguntas clínicas, y aun en muchos casos, muchos resultados, con el tiempo se demostraron que eran incorrectos por sesgos o insuficiente poder estadístico por tamaño de la muestra. Por tanto el clínico, para llevar a cabo una buena práctica clínica, debe estar basado en el análisis de la mejor eviden-

cia disponible, la cual puede ser proporcionada por una revisión sistemática, sobre la base del análisis sistemático y metódico de un número adecuado de informes de investigación relevantes (3).

En segundo lugar, existe una explosión de información. Más de dos millones de artículos son publicados anualmente, en más de 20.000 revistas médicas. Ante este volumen inmanejable de información, los médicos preferimos resúmenes no sesgados de investigaciones originales múltiples.

En tercer lugar, dentro de la investigación por revisión de publicaciones se conocen dos tipos diferentes: la revisión sistemática y la revisión no sistemática o revisión narrativa (update). El método tradicional de integración en la literatura ha sido la revisión narrativa, pero tiene dos debilidades básicas. En primer lugar, no hay norma sobre cómo conseguir los datos primarios, cómo integrar los resultados; lo que prima es el criterio subjetivo del revisor. En segundo lugar, el revisor narrativo no sintetiza cuantitativamente los datos hallados en las distintas publicaciones, por tanto, estas revisiones son muy susceptibles a imprecisiones y sesgos. La revisión sistemática por el contrario, exige un método riguroso y explícito para la identificación, evaluación crítica y síntesis de la evidencia obtenida (5). En la tabla 1 se resumen las diferencias. Por ejemplo, una revisión narrativa sobre enfermedad por reflujo gastroesofágico (RGE), incluiría apartados sobre la fisiopatología, clínica, diagnóstico y tratamiento. Una revisión sistemática plantearía concretamente la pregunta, si la erradicación del *helicobacter Pylori* en los pacientes con RGE aumentaría la severidad de los síntomas y la frecuencia de recaídas, cuando se compara con el tratamiento estándar: Inhibidores de bomba de protones.

En cuarto lugar, las revisiones sistemáticas pueden identificar las razones de las discrepancias o las contradicciones entre los resultados de las distintas investigaciones, impulsando a rediseñar los estudios con el objeto de mejorar la investigación clínica. Además, al incluir mayor número de pacientes

Tabla 1. Diferencias entre revisión sistemática y revisión narrativa.

CARACTERÍSTICAS	NARRATIVA	SISTEMÁTICA
Focalizada	Tema	Pregunta
Estrategia de búsqueda	No especificado	Claramente especificado
Criterios de selección	No especificado	Especificado y aplicado
Análisis de la información	Variable	Riguroso y crítico
Síntesis	Cualitativa	Cualitativa o cuantitativa (Metanálisis)

tienen mayor precisión en sus estimaciones que los estudios primarios.

¿QUÉ ES UNA REVISIÓN SISTEMÁTICA O METANÁLISIS?

Se define revisión sistemática como un estudio integrativo, observacional, retrospectivo, secundario, en el cual se combinan estudios que examinan la misma pregunta. A su vez, dentro de la revisión sistemática existen dos formas: “cuantitativa o metanálisis” y “cualitativa u overview”. Las diferencias están dadas fundamentalmente por el uso de métodos estadísticos, que permite la combinación y análisis cuantitativo de los resultados obtenidos en cada estudio. El término metanálisis fue introducido por Glass en 1976, quien lo definió en los siguientes términos: un análisis estadístico de una amplia serie de análisis de resultados de estudios individuales con el objeto de integrar sus hallazgos. Aunque a veces se utilicen indistintamente los dos términos, una revisión sistemática no es igual a metanálisis (5).

¿CUÁNDO SE NECESITA UNA REVISIÓN SISTEMÁTICA?

Las revisiones sistemáticas se necesitan cuando existe una pregunta puntual generalmente relacionada con efectividad clínica (¿este tratamiento confiere mayores beneficios cuando es usado en este grupo de pacientes?), varios estudios primarios (quizás con hallazgos divergentes) y una incertidumbre marcada.

Aunque la mayor parte de las revisiones sistemáticas evalúan ensayos clínicos para responder a preguntas sobre intervenciones terapéuticas, se empieza a extender su aplicación a otro tipo de estudios que responden preguntas sobre pruebas diagnósticas o de pronóstico, involucrando estudios observacionales (6).

Las revisiones sistemáticas proveen una síntesis racional de la investigación básica. Supera las limitaciones de las revisiones narrativas al aplicar estándares rigurosos a la investigación secundaria (donde la unidad del estudio son otros estudios de investigación) como si fueran aplicados a estudios de investigación primaria (estudios originales).

Es conocido el hecho de que pueden pasar años desde que se publica una intervención efectiva hasta que se aplica de forma rutinaria por los clínicos. Un ejemplo típico es el estudio de Antman y cols, publicado en 1992 (7) en el que se evidenció que transcurrieron más de diez años desde que comienza a recomendarse la trombolisis como intervención efectiva tras un infarto agudo del miocardio, hasta que se refleja en los libros de texto y en las revisiones narrativas, con el perjuicio que provoca en aquellos pacientes que pudieron haberse beneficiado de dicha intervención si la información se hubiese aplicado con celeridad (8).

Cuando abordamos una pregunta o una duda clínica, tenemos claro que el mejor tipo de estudio que puede resolver nuestra pregunta es una revisión sistemática, especialmente cuando se refiere a terapia.

Archie Cochrane acuñó la frase que aparece en la introducción, sobre la necesidad de un mayor número de revisiones de conjunto sistemáticas. La respuesta a este desafío lo constituyen las revisiones llevadas a cabo por la Colaboración Cochrane, una organización internacional sin ánimo de lucro que intenta ayudar a tomar decisiones clínicas y de salud pública, preparando, manteniendo y divulgando revisiones sistemáticas (9).

No se deben utilizar las revisiones sistemáticas y el metanálisis para aprobar o no un fármaco nuevo, hacer que un efecto pequeño se haga significativo al combinar estudios de baja calidad, frenar las investigaciones originales o con el único objeto de aumentar el número de publicaciones sin una necesidad y un objetivo claro para responder una pregunta de investigación relevante (8).

¿CUÁLES SON LOS PASOS BÁSICOS DE UNA REVISIÓN SISTEMÁTICA?

El desarrollo de una revisión sistemática requiere los siguientes pasos (10).

Tabla 2. Elementos para una correcta revisión sistemática.

- Claro enunciado de la pregunta a responder
- Definición de criterios de inclusión y exclusión
- Extensa identificación de la evidencia disponible
- Uniforme extracción de los datos
- Evaluación de la calidad de los estudios
- Clara presentación de los resultados
- Análisis estadístico (metanálisis)
 - . Estimación de la medida global del efecto con su IC 95%
 - . Determinación de la heterogeneidad
 - . Análisis de subgrupos y sensibilidad
 - . Determinación del sesgo de publicación

1. Definir una pregunta claramente

Para iniciar una revisión sistemática es necesario identificar y convertir el problema, la incertidumbre o “laguna del conocimiento” en una pregunta que pueda ser respondida. Formular una pregunta significa reducirla a términos claros y precisos, identificando sus componentes principales. La nemotecnia PICO nos ayuda a estructurar los cuatro componentes de una pregunta clínica. Los pacientes o grupos de población relevantes, la intervención (exposición o procedimiento diagnóstico) de interés, así como también contra quién se compara la intervención y los desenlaces (outcomes) apropiados. Por ejemplo: ¿En los pacientes con reflujo gastroesofágico, *Helicobacter Pylori* positivo que reciben IBP por períodos prolongados, la erradicación del *Helicobacter Pylori* (Intervención) disminuye el riesgo de gastritis atrófica (Outcome o Desenlace) cuando se compara con la no erradicación? Posteriormente deberán especificarse también los tipos de estudios que resultan apropiados para tratar de responder a la pregunta apropiada (2). En general, con cuanta más precisión se defina cada uno de los componentes, más exacta resultará la revisión.

2. Especificación de los criterios de inclusión y exclusión de los estudios

A priori se deben señalar las características de los estudios para ser incluidos o excluidos. Como base para los criterios de inclusión pueden exponerse los siguientes aspectos relevantes: el tipo de diseño metodológico del estudio (por ejemplo, sólo incluir estudios controlados y aleatorizados), el tamaño de las muestras de cada estudio, el tipo de tratamiento tanto experimental como de control, el año de publicación del estudio y sobre todo, que describa con suficiencia las características de los pacientes o las unidades de observación, así como los resultados obtenidos con las escalas y las unidades de medición que fueron utilizadas (11).

3. Formulación del plan de búsqueda de la literatura

La recopilación de la información debe ser exhaustiva, tanto con estudios publicados como no publicados para evitar incurrir en el sesgo de selección (no incluir estudios relevantes provocando un error sistemático en la estimación del efecto) o en el sesgo de publicación. (Se sabe que muchas revistas sólo aceptan estudios con resultados positivos y rechazan los estudios con resultados negativos). Otros sesgos habituales, además de los señalados son el sesgo de publicación en inglés, de publicación múltiple y de no inclusión en bases de datos y de citación (11). La búsqueda se inicia con la formulación adecuada de las palabras clave, la investigación en las bases de datos electrónicas (Medline, Embase, Lilacs, Colaboración Cochrane, etc.), búsquedas manuales en revistas no incluidas en las bases de datos, registros de ensayos clínicos existentes, consultas con expertos, resúmenes, comunicaciones a congresos (abstracts), consultas a Agencias financiadoras de investigaciones y la industria farmacéutica. La literatura “gris” u “opaca”, compuesta por informes técnicos, actas de congresos y tesis doctorales, también puede ser útil, ya que muchos estudios referenciados de esta manera no son publicados. Solo un 31% de los metanálisis publicados incluyen este tipo de evidencia, sobrestimando el efecto de las intervenciones evaluadas en un 33% (12, 13). La estrategia de búsqueda debe estar claramente descrita en el protocolo.

4. Registro de los datos y evaluación de la calidad de los estudios seleccionados

Es importante registrar con todo rigor y detalle las características relevantes mediante un formato predeterminado. El registro debe incluir escalas o criterios preestablecidos para definir la validez de los estudios. La evaluación de la calidad de los estudios individuales que se incluyen en las revisiones sistemáticas es necesaria para limitar los sesgos, formarse una idea más precisa de las potenciales comparaciones y guiar la interpretación de los resultados.

Tabla 3. Identificación de la evidencia disponible

- Predeterminada estrategia de búsqueda
- No limitada a una base de datos (Medline, Aidsline, Cancerlit, Embase, Lilacs, Colaboración Cochrane, Publicaciones secundarias ACP journal, etc.
- Registros de ensayos clínicos
- No limitada a un idioma
- Consulta con expertos
- Investigación manual de publicaciones específicas
- Inclusión de literatura “gris”.

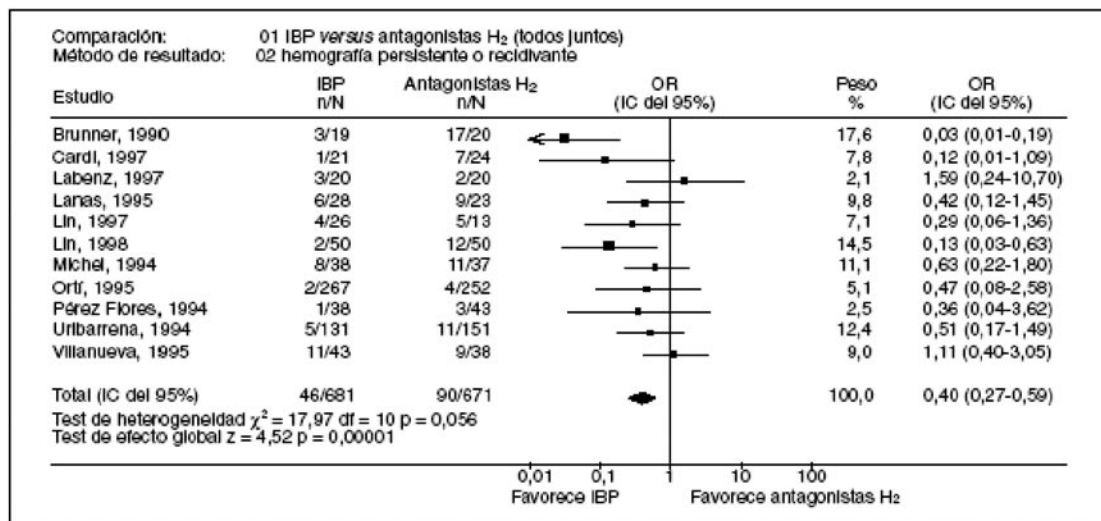
La calidad es difícil de definir. Una dimensión de la calidad está relacionada con la validez de los hallazgos generados por el estudio. Y se plantean dos distinciones: Validez interna: se refiere a la capacidad de evitar errores sistemáticos o sesgos mediante el diseño y la realización del estudio, mientras que la validez externa o general se relaciona con aplicabilidad de los resultados a una población específica. Las escalas más frecuentemente usadas son las de Jadad (14) y los criterios de Guyat y Sacket (15) que involucran en los experimentos clínicos controlados, tres criterios fundamentalmente: Aleatorización y ocultamiento en la asignación del tratamiento, el cegamiento para aquéllos que dan y reciben la atención y finalmente el sesgo de desgate o de pérdidas (attrition bias) que involucra la información sobre las pérdidas de los participantes (por abandono, bajas, violaciones de protocolo, etc.) Todos los pacientes deben ser incluidos en el análisis y mantenerlos en el grupo de asignación original, sin relación a su adherencia al protocolo. En otras palabras, el análisis debe ser realizado, de acuerdo al principio de la intención de tratar. David Moler y sus colaboradores han identificado 25 escalas y 9 listas utilizadas para evaluar la validez y la calidad de los ensayos clínicos aleatorizados y controlados (15). Existen otras guías como Consort, para experimentos clínicos, que evalúan más la calidad de los informes que la calidad metodológica (16).

5. Interpretación y presentación de los resultados

Los hallazgos de los estudios individuales deben ser agregados para producir una gráfica sobre la efectividad clínica de la intervención. Algunas veces esta agregación es cualitativa pero usualmente es cuantitativa, usando la técnica conocida como metanálisis. La forma usual de presentar los datos es a través de una grafica conocida como “forest plot” o gráfico de efectos. La gráfica 1 muestra los resultados de un metanálisis que compara el riesgo de hemorragia persistente o recidivante por úlcera péptica entre los IBP y los bloqueadores H-2 (17). El resultado de cada estudio se representa por uno de los índices habituales, por ejemplo el OR (odds ratio o razón de disparidad). El OR de cada estudio se representa en la grafica por un cuadrado (la medida resumen del tamaño del efecto) y con una línea horizontal (su intervalo de confianza es del 95%). El tamaño del cuadrado varía de acuerdo a la información en cada estudio individual y la longitud de la línea representa la incertidumbre o precisión del efecto de tratamiento de ese estudio. A una línea de mayor longitud,

menor precisión el tamaño del efecto, probablemente por pequeño número de pacientes y menor poder para detectar pequeñas diferencias significativas.

Recuerde que si no hay diferencia entre los tratamientos, el OR es 1. El convenio en las revisiones sistemáticas es construir el OR de modo que los valores menores de 1 estén a la parte izquierda de la gráfica correspondiendo a un efecto favorable del tratamiento en estudio. La gráfica contiene también una línea vertical de referencia en OR: 1, valor de no efecto. Si se realiza metanálisis, el OR combinado se representa por un rombo, cuya diagonal horizontal representa su IC del 95%. En la figura se observan 11 estudios. En la mayoría de los casos el cuadrado que representa el efecto resumen se encuentra situado a la izquierda de la línea media, sugiriendo un efecto preventivo de los IBP con relación al desenlace persistente o recidiva hemorrágica de la úlcera péptica. Sin embargo, en 9 de ellos, los intervalos cruzan la línea de no efecto, indicando que aunque existe una tendencia a favor de los IBP, las diferencias no alcanzan significancia estadística probablemente por el reducido tamaño de la muestra. Y es aquí el valor



Gráfica 1. Metanálisis que compara los inhibidores de bomba de protones con los bloqueadores H-2 en el tratamiento de la hemorragia digestiva por la úlcera péptica (17)

del metanálisis que aglutina los resultados de todos los estudios, proporcionándonos una estimación más precisa sobre el efecto de tratamiento: Un OR global de 0,4 (representado por el rombo en la parte inferior, con un intervalo de confianza del 95% (representado por los extremos de dicho rombo o diamante) que no incluye la unidad (0,27 a 0,59), por consiguiente, indica un efecto estadísticamente significativo (18, 19).

Análisis

Cuando se planea un análisis cuantitativo (metanálisis) aparecen los siguientes problemas:

1. Comprobar si los resultados del estudio se pueden combinar
2. Evaluar la heterogeneidad
3. Estimar la medida de resumen del efecto o tamaño del efecto
4. Evaluar los potenciales sesgos
5. Presentar los resultados.

Evaluación de la heterogeneidad. Es importante preguntarse si los resultados de los estudios son similares u homogéneos y se pueden combinar entre sí. Una manera de responder a esta pregunta es examinar una presentación gráfica de los resultados. Si los intervalos de confianza para los resultados de cada estudio (representados típicamente por líneas horizontales) no se superponen, esto indica que es probable que las diferencias entre los estudios sean estadísticamente significativas y no se podrán combinar porque probablemente las diferencias observadas no son explicadas por el azar. Las pruebas estadísticas que valoran heterogeneidad suelen tener una baja potencia estadística, por lo que es habitual situar el punto de corte para la p en 0,1 en lugar del habitual 0,05. Además, una prueba estadística no significativa no necesariamente descarta la existencia de heterogeneidad.

El test de homogeneidad se puede realizar mediante la prueba de χ^2 cuadrado. Si el resultado de esta prueba no es significativo, se interpretará como

que los distintos estudios son homogéneos y por lo tanto se pueden combinar. Si por el contrario, la prueba resulta significativa ($p < 0,01$), se llega a la conclusión de que los estudios son heterogéneos y no combinables entre sí. Aquí existen dos opciones. Algunos autores son partidarios de continuar el análisis estadístico, medir la magnitud del efecto eligiendo pruebas estadísticas que tengan en cuenta esa heterogeneidad. Otros autores sugieren que debe concluirse aquí el estudio, convertirse en una revisión sistemática cualitativa con la exposición separada de los estudios e intentando explicar la heterogeneidad.

Para medir el efecto ponderado de todos los estudios se elegirá una prueba estadística de acuerdo al test de homogeneidad. Se utilizará el modelo de efectos fijos, cuando el test de homogeneidad no ha sido significativo, es decir, se asume que la variabilidad entre los estudios es debida al azar. Pero si el test de homogeneidad ha sido significativo, se usa el modelo de efectos variables, asumiendo que cada estudio tiene un efecto diferente y que la muestra de los estudios es una muestra aleatoria de toda la población de estudios (19-21).

Medida de efecto para combinar los resultados. Se deben establecer las medidas de resumen de los resultados para cada uno de los estudios incluidos. Estos valores describirán los efectos de tratamiento observados. En una segunda etapa se calculan los efectos de tratamiento global, como un promedio ponderado de las medidas de resumen de los estudios. La Odds ratios (OR) –no hay una traducción satisfactoria, razón de disparidad o de momios–, es la principal medida de efecto usada en metanálisis, con desenlaces binarios (muerte o supervivencia). En datos continuos como los cambios de la tensión diastólica después de tratamiento antihipertensivo se pueden expresar como diferencias de medias. Para propósitos prácticos, el OR, que ofrece algunas ventajas prácticas, puede ser interpretado como si fuera el riesgo relativo. Un OR de 2, implica que el desenlace se espera que ocurra dos veces más frecuentemente en el grupo de intervención que en el grupo control.

Un OR de 0,5, implica un 50% de reducción en el evento definido en el grupo intervenido que en el grupo control. La interpretación del OR (como un aumento o una disminución), tiende a sobreestimar el efecto de tratamiento (10), aunque muy mínimo, especialmente cuando la tasa de eventos en los grupos es menor al 30%.

Otro estadístico de resumen disponible, es el “número necesario a tratar” NNT, que puede obtenerse fácilmente a partir de la diferencia de riesgos (es su inverso); es una forma útil e intuitiva, con mayor significancia clínica de expresar los resultados: nos dice cuántos pacientes es necesario tratar con el nuevo tratamiento para producir un evento de interés, en comparación con el placebo o con el tratamiento tradicional. Por ejemplo, en el metanálisis de la gráfica 1, cuando se compara la incidencia de resangrado en los pacientes con úlcera péptica tratados con IBP versus bloqueadores H-2 se observó un OR de 0,4 (con intervalos de confianza que no incluye la unidad). Este dato, que muestra un efecto protector, no es fácil de traducir a la clínica, pero si lo expresamos como un NNT de 3, indica que por cada 3 pacientes con úlcera péptica a quienes administremos IBP, (comparado con bloqueadores H-2), conseguimos adicionalmente prevenir un episodio de resangrado. Aunque no existe una regla para clasificar un NNT como alto o bajo, un valor de 3 produce un efecto terapéutico importante. Pero si el NNT es de 50, en un metanálisis que evalúa mejoría sintomática en pacientes con dispepsia no ulcerosa y terapia de erradicación del helicobacter, parece que no es beneficioso tratar 50 con terapia antibiótica para lograr la mejoría en un paciente con dispepsia funcional. (9, 22, 23).

Las medidas absolutas como el NNT o diferencias de riesgos pueden ser más informativas porque reflejan el riesgo basal y el cambio del riesgo con la intervención, pero sin embargo, como generalmente en los estudios, la tasa de eventos en el grupo control no es similar, deberán utilizarse medidas relativas como el OR o el riesgo relativo. En metanálisis,

el NNT debe acompañarse de su IC y solamente deben calcularse cuando los períodos de seguimiento son similares.

Análisis de sensibilidad y análisis de subgrupos. El siguiente paso es determinar si la estimación es robusta o sólida. El análisis de sensibilidad determina si la medida de resumen del efecto está influenciada por cambios en las asunciones y por las decisiones asumidas en el protocolo para las combinaciones de los datos. A menudo resulta interesante examinar una categoría especial de participantes en la revisión (mujeres, un grupo de edad o la gravedad de la enfermedad). Debe reducirse al mínimo el número de análisis de subgrupos porque pueden encontrarse diferencias únicamente por azar.

Evaluación de los sesgos. El sesgo de publicación (estudios positivos son más probables de que sean publicados que los estudios con resultados negativos) puede causar heterogeneidad en los resultados del metanálisis. La más común aproximación para evaluarlo es el “funnel plot”, herramienta mediante la cual se compara gráficamente algún efecto específico del estudio (OR o RR) con alguna medida de su precisión (habitualmente el número de participantes en el estudio). Si el gráfico es simétrico, como una V invertida, se interpreta como una demostración de que probablemente no existe sesgo de publicación. Si el gráfico es asimétrico, la interpretación es que probablemente existe sesgo de publicación. (24, 25).

¿CUÁLES SON LAS LIMITACIONES DE LAS REVISIONES SISTEMÁTICAS Y METANÁLISIS?

Existen varios errores acerca de las Revisiones Sistemáticas (RVS) y los metanálisis. La creencia de que es una simple combinación de múltiples estudios, una revisión ordinaria pero más grande. Como resultado de reducir los sesgos en la selección de los estudios teniendo en cuenta la calidad, una RVS puede incluir menos estudios que una revisión narrativa.

No se debe olvidar que la RVS es un estudio observacional, retrospectivo, susceptible a sesgos, y sigue siendo una técnica controvertida dado que los métodos estadísticos no pueden corregir o ajustar las discrepancias entre los estudios y mucho menos lograr que los estudios contradictorios se combinen en una única medida global. El metanálisis no debe sustituir a los ensayos clínicos de gran tamaño; sin embargo, cuando éstos no existen, constituyen una guía para la práctica clínica e incluso para el diseño de futuros ensayos. El metanálisis es un instrumento para la generación de hipótesis y no para la toma de decisiones. Es complementario, no sustituto del ensayo clínico aleatorizado de gran tamaño (26, 9).

Las RVS no son la panacea para convertir estudios malos en buenos, y el error más frecuente al realizar metanálisis es mezclar estudios buenos con malos, por eso, las guías comentadas para evaluar los ensayos clínicos controlados y los estudios observacionales deben estar disponibles y cumplirse. Tampoco se pueden mezclar “manzanas y naranjas”, si los investigadores ignoran las diferencias fundamentales de riesgo basal, determinantes de los resultados de los estudios y los combinan sin más (27).

En resumen, las RVS y los metanálisis, si se utilizan correctamente, constituyen la mejor herramienta posible para combinar los resultados de diferentes estudios en presencia de esta información desbordante, pero dado que son estudios retrospectivos de investigación, pueden llevarnos a conclusiones sesgadas a partir de estudios imperfectos que aisladamente no habrían sido capaces de confundirnos (28).

Referencias

- Sánchez, M.; Lifshitz, A.; Juárez Díaz, N., “Introducción y conceptos generales”, en *Medicina basada en evidencias*, México, McGraw-Hill Interamericana, 2002: 1-20.
- Centro Cochrane Español, “Las revisiones narrativas. Curso de Revisiones Sistemáticas”, en www.cochrane.es, diciembre de 2004.
- León Hernández, S.R.; Aguilera Cepeda, J.M., Metodología de los estudios de metanálisis, *Rev Mex Ortop Traum*, 2001, 15(2): 94-99.
- Thacker, S., Metanálisis: un enfoque cuantitativo para la integración de las investigaciones, *Bol Of Sanit Panam*, 1993, 115(4): 328-339.
- Ortiz, Z., “¿Qué son las revisiones sistemáticas?”, en www.centrocochrane.mx, diciembre de 2004.
- Davies, Huw; Crombie, Iain, “What is a systematic review?”, en www.evidence-based-medicine, marzo de 2005.
- Antman, EM; Lau, J; Kupelnick, B, et al., A comparison of results of metaanalysis of randomized controls trials and recommendations of clinical experts, *JAMA*, 1992, 268: 240-248.
- Guerra, J.A.; Martín Muñoz, P.; Santos Lozano, J.M., “Las revisiones sistemáticas. Niveles de evidencia y grados de recomendación”, en www.fisterra.com.
- Gisbert, J.P.; Bonfill, X., Cómo realizar, evaluar y utilizar revisiones sistemáticas y metanálisis, *Gastrenterol Hepatol*, 2004, 27(3): 129-149.
- Davies, Huw; Crombie, Iain, “What is metaanalysis?”, en www.evidence-based-medicine, marzo de 2005.
- Roque, M., Organización de una Revisión Sistemática para un grupo Cochrane de Revisiones, Adaptación de puntos básicos para revisores, Grupo Cochrane de Revisiones Neonatales, Centro Cochrane Iberoamericano, diciembre 2004.
- McManus, R.J.; Wilson, S.; Delaney, BC, et al., Review of the usefulness of contacting other experts when conducting a literature search for systematic reviews, *BMJ*, 1998, 317: 1562-1563.
- McAuley, L, et al., Does the inclusion of grey literature influence estimates of intervention effectiveness reported in metaanalyses? *Lancet*, 2000, 356: 1228-1231.
- Clarke, M.; Oxman, AD, “Manual de Revisores. 4.1. 6. Evaluación de la calidad de los estudios”, Colaboración Cochrane, 2000.
- Juni, P.; Altman, D; Egger, M., “Assessing the quality of randomized controlled trials”, en *Systematic Reviews in Health Care. Metanalysis in context*, 2ª edic., Londres, BMJ Publishing group, 2001, 85-105.
- Altman, D.; Schulz, KF; Moher, D; Egger, M.; Davidoff, F; Elbourne, D., for the CONSORT Group, “The revised CONSORT statement for Reporting Randomized Trials: Explanation and Elaboration”, *Ann Intern Med*, 2001,134: 663-694.
- Gisbert, JP; González, L; Calvet, X.; Roque, M.; Gabriel, R.; Pajares, JM., Proton pump inhibitors versus H2-antagonist: a metaanalysis of their efficacy in treating bleeding peptic ulcer, *Aliment Pharmacol Ther*, 2001, 15: 917-26.
- Abraira, V., Revisiones sistemáticas y metanálisis, *Semergen*, 2003, 29: 183-185.
- Cuervo, I.G.; Lozano, JM.; Pérez, A., “Revisiones sistemáticas de la literatura y metanálisis”, en Ruiz Morales, A; Gómez, R; Londoño, D., *Investigación clínica: Epidemiología clínica aplicada*, Centro Editorial Javeriano, 2001, 245-267.
- Moliner, Luis M., “Metanálisis. Asociación de la Sociedad Española de Hipertensión”, en www.seh-lelha.org. Diciembre de 2004.
- Greenhalgh, T., How to read a paper: Papers that summarize other papers (systematic review and metaanalyses), *BMJ*, 1997, 315: 672-675.
- McQuay, H.; Moore, A., Using numerical results from Systematic Reviews in clinical practice, *Ann Intern Med*, 1997, 126: 712-720.

-
23. Gisbert, JP; Calvet, X.; Gabriel, R., Pajares, JM, Infección por helicobacter Pylori y dispepsia funcional. Metanálisis de la eficacia del tratamiento erradicador, Barcelona, Med Clin, 2002, 118: 405-409.
 24. Moayyedy, P., Metanalysis: Can We Mix apples and Oranges?, Am J Gastroenterol, 2004, 2297-2301.
 25. Francisco García, Funnel plots y heterogeneidad, Bandolera, 2000, 81-85.
 26. Dennis, R., "Metanálisis de experimentos clínicos", en Ardila, E.; Sánchez, R.; Echeverri, J., Estrategias de investigación en Medicina Clínica, Manual Moderno, 2001, 169-179.
 27. Huang, J.Q.; Hunt, R., "Editorial: uses and Abuses of meta-analysis in gastroenterology", Chinese Journal of Digestive Diseases, 2002, 3: 39-42.
 28. Molinero, LM., Interpretación, utilidad y limitaciones del metanálisis, Gastroenterol Hepatol Continuada, 2002, 311-318.