

- dence.org/
32. CAT Bank [en línea] Michigan University [fecha de acceso 19 de Febrero de 2001]. URL disponible en: <http://www.ped.med.umich.edu/ebm/cat.htm>
  33. CAT Bank [en línea] Washington University [fecha de acceso 19 de Febrero de 2001]. URL disponible en: <http://depts.washington.edu/pedebm/>
  34. CAT Bank [en línea] Rochester University [fecha de acceso 19 de Febrero de 2001]. URL disponible en: <http://www.urmc.rochester.edu/medicine/res/CATS/ped.html>
  35. CAT Bank [en línea] North Carolina University [fecha de acceso 19 de Febrero de 2001]. URL disponible en: <http://www.med.unc.edu/medicine/edursrc/!catlist.htm>
  36. The PedsCCM Evidence- Based Journal Club [en línea] Web Pediatric Critical Care Medicine [fecha de acceso 19 de Febrero de 2001]. URL disponible en: [http://pedscm.wustl.edu/EBJournal\\_Club.html](http://pedscm.wustl.edu/EBJournal_Club.html)
  37. The Centre for Clinical Effectiveness [en línea] Monash University [fecha de acceso 19 de Febrero de 2001]. URL disponible en: <http://www.med.monash.edu.au/publichealth/cce/>
  38. Archivo de Temas Valorados Críticamente [en línea] web de la AEPap [fecha de acceso 19 de Febrero de 2001]. URL disponible en: <http://www.aepap.org/pedev/pedev-4.htm>
  39. PubMed [en línea] National Library of Medicine [fecha de acceso 19 de Febrero de 2001]. URL disponible en: <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=PubMed>
  40. Descriptores en Ciencias de la Salud (DeSC) [en línea] BIREME-Biblioteca Virtual en Salud [fecha de acceso 19 de Febrero de 2001]. URL disponible en: <http://www.bireme.br/decs/E/homepage-gee.htm>
  41. PubMed- Opción "Clinical Queries" [en línea] National Library of Medicine [fecha de acceso 19 de Febrero de 2001]. URL disponible en: <http://www.ncbi.nlm.nih.gov/entrez/query/static/clinical.html>
  42. Buscar en Medline con PubMed (guía de uso) [en línea] González C (Web Fisterra) [fecha de acceso 19 de Febrero de 2001]. URL disponible en: [http://www.fisterra.com/recursos\\_web/no\\_explor/pubmed.htm](http://www.fisterra.com/recursos_web/no_explor/pubmed.htm)
  43. EMBASE [en línea] Elsevier [fecha de acceso 19 de Febrero de 2001]. URL disponible en: <http://www.elsevier.nl/homepage/sah/spd/embase/menu.htm>
  44. Proyecto "3.000 revistas" [en línea] Infodocor [fecha de acceso 19 de Febrero de 2001]. URL disponible en: <http://www.infodocor.org/revis.htm>
  45. Freemedicaljournals [en línea] AMEDEO, The Medical Literature Guide [fecha de acceso 19 de Febrero de 2001]. URL disponible en: <http://www.freemedicaljournals.com/>
  46. Free Online Full-text Articles [en línea] HighWire Press [fecha de acceso 19 de Febrero de 2001]. URL disponible en: <http://highwire.stanford.edu/lists/freetart>
  47. Birken CS, Parkin PC. ¿En qué revistas encontrarán los pediatras las mejores pruebas para la práctica clínica? *Pediatrics (ed. esp.)* 1999; **47**: 313- 319.
  48. Archives of Diseases in Childhood [en línea] BMJ Publishing Group [fecha de acceso 19 de Febrero de 2001]. URL disponible en: <http://adc.bmjournals.com/>
  49. British Medical Journal [en línea] BMJ Publishing Group [fecha de acceso 19 de Febrero de 2001]. URL disponible en: <http://www.bmj.com/>
  50. Journal of the American Medical Association (JAMA) [en línea] American Medical Association [fecha de acceso 19 de Febrero de 2001]. URL disponible en: <http://jama.ama-assn.org/>
  51. The Pediatric Infectious Diseases Journal [en línea] Lippincott, Williams & Wilkins [fecha de acceso 19 de Febrero de 2001]. URL disponible en: <http://www.pidj.com/>
  52. The Lancet [en línea] The Lancet Publishing Group [fecha de acceso 19 de Febrero de 2001]. URL disponible en: <http://www.thelancet.com/>
  53. New England Journal of Medicine [en línea] Massachusetts Medical Society [fecha de acceso 19 de Febrero de 2001]. URL disponible en: <http://www.nejm.org/content/index.asp>
  54. Pediatrics [en línea] American Academy of Pediatrics & HighWire Press [fecha de acceso 19 de Febrero de 2001]. URL disponible en: <http://www.pediatrics.org/>
  55. Journal of Infectious Diseases [en línea] The University of Chicago Press, Journals Division [fecha de consulta 19 de Febrero de 2001]. URL disponible en: <http://www.journals.uchicago.edu/JID/home.html>

### 3. REVISIONES SISTEMÁTICAS. COLABORACIÓN COCHRANE

#### LEER E INTERPRETAR UNA REVISIÓN SISTEMÁTICA

**Antonio Sáenz**

*Centro de Salud Pozuelo 1. INSALUD. Madrid.*

#### EL COMIENZO

Es difícil mantenerse actualizado en la literatura médica. Cada año se publican 30.000 revistas. Un médico que intente mantenerse al día sólo en su especialidad, por ejemplo Vd., debe leer 140 artículos completos cada semana, o 600 el día de vuelta de sus vacaciones de verano. Las revisiones nos ofrecen la posibilidad de conseguir estar informados sin necesidad de invertir esa enormidad de tiempo.

Tradicionalmente se han realizado un tipo de revisiones que llamaremos narrativas. Para ello, podemos buscar la ayuda de expertos muy familiarizados con un tema al que

han dedicado buena parte de su vida profesional, como el tratamiento del niño con otitis<sup>(1)</sup>. El inconveniente es que en este tipo de revisiones utilizan métodos informales, no explícitos, y a menudo personales y subjetivos.

En ese caso, encaramos al menos 2 problemas. Primero, los autores no especifican el proceso que han seguido para buscar, reunir y evaluar la información que les ha llevado a ofrecernos sus conclusiones. En segundo lugar, y debido a la ausencia de esta información, nosotros los lectores no vamos a ser capaces de repetir y verificar los resultados y las conclusiones de las revisiones. Son conclusiones que se toman o se dejan, pero no podemos discutir las.

Una revisión sistemática es aquella en la que existe una búsqueda exhaustiva de estudios relevantes en un tema (p.e. ¿es beneficioso el ácido fólico para prevenir defectos en el tubo neural del niño?, o bien ¿es eficaz la intervención sobre los factores de riesgo cardiovascular en la prevención primaria de morbilidad coronaria?). Los estudios así identificados son obtenidos en texto completo y los resultados sintetizados de acuerdo con un método predeterminado y explícito. Esta forma de revisión da al lector una gran ventaja sobre otras revisiones: la posibilidad de replicarla y verificar si se llega a la misma conclusión. Aunque Vd. no disponga de suficiente tiempo para replicarla, le podemos asegurar que alguien lo hará.

La revisión sistemática (RS) es el proceso que nos lleva a obtener los estudios cuyos resultados pueden combinarse matemáticamente para poder ofrecer conclusiones. Ese método matemático se llama metaanálisis (MA) y por lo tanto el MA es sólo una parte de la RS. Un MA es la combinación estadística de, al menos, dos estudios para obtener una estimación o suma única del efecto de la intervención en salud que estamos evaluando (p.e. ¿es útil la pentoxifilina en la claudicación intermitente?: en este caso existen 17 ensayos clínicos randomizados y los podemos combinar)<sup>(2)</sup>.

En ocasiones al hacer una revisión sistemática sólo encontramos un estudio. Al no poder combinarlo con otro estudio, hemos realizado una RS, pero no podemos hacer MA, (p.e. ¿es útil el dipiridamol en la claudicación intermitente?: En este caso sólo existía un ensayo clínico randomizado)<sup>(3)</sup>.

El simple acto de combinar estadísticamente estudios no es garantía de una respuesta válida (validez: varios grupos de profesionales con los mismos métodos obtendrían los mismos artículos y resultados) y fiable (fiabilidad: varios gru-

pos de profesionales evaluando estos artículos y resultados llegarían a las mismas conclusiones). De hecho, no todas las revisiones sistemáticas y metaanálisis han sido conducidas rigurosamente. Para ser válido y fiable la RS y MA debe seguir un proceso sistemático y detallado, asegurar la ausencia de sesgos e imprecisiones, y que sea publicado con claridad y precisión, para que otros lo puedan replicar.

### Ventajas

Las RS realizadas y comunicadas apropiadamente pueden ayudar a resolver controversias entre estudios divergentes, guiar la investigación planteando nuevas hipótesis, identificando áreas insuficientemente investigadas, o por el contrario, áreas suficientemente investigadas y dónde un estudio adicional podría no ser ético. Además la RS puede identificar beneficio o daño de una intervención años antes de que un ensayo clínico de muestra suficientemente grande lo pueda hacer.

### Limitaciones

Las búsquedas electrónicas sólo identifican el 50-75% de todos los estudios relevantes existentes. La búsqueda manual de revistas no indexadas es laboriosa y no está al alcance de todos nosotros. También es laborioso buscar literatura "gris" (estudios no publicados, no indexados, tesis, *abstracts* de congresos, comunicaciones orales, publicaciones médicas no científicas). Incluso si encontramos todos los estudios relevantes su calidad es variable y puede ofrecernos resultados contradictorios, p.e. todos los estudios combinados parecen favorecer a la intervención, pero cuando combinamos sólo los de calidad superior el efecto beneficioso de la intervención tiende a ser menor o incluso inexistente.

Para superar estas limitaciones y maximizar los beneficios de las RSC en las decisiones en salud, un grupo internacional de voluntarios prepara, mantiene y disemina RS de calidad homologada. Este grupo se conoce como la Colaboración Cochrane.

### LEYENDO LA FORMULACIÓN DE LA PREGUNTA

Una pregunta confusa conduce a una respuesta confusa. Un posible método de hacer la pregunta principal es dividirla en 4 partes: intervención, población, diagnóstico y resul-

tados (IPDR): p.e. ¿Son útiles los corticoides en niños de 1 a 10 años con diagnóstico de Croup para disminuir la duración de los síntomas, la estancia hospitalaria, la tasa de ingresos,...?. Por cierto, la respuesta es sí<sup>(4)</sup>.

La pregunta principal, desarrollada a priori, determinará el enfoque principal de la revisión. Las preguntas secundarias, también formuladas a priori, forman la base de los análisis secundarios, como son los de subgrupos y de sensibilidad.

### Desarrollando los criterios de selección

Con una pregunta bien formulada, los criterios de selección salen solos, claros y concisos. Nos describen qué estudios serán elegibles para ser incluidos. Se basan en IPDR y en el diseño de los Estudios que vamos a incorporar. Además: años en los que los estudios fueron realizados, lenguajes de publicación, y si se incluirán estudios, tanto publicados, como no publicados. Se necesita una explicación de los criterios de selección.

Es importante para los revisores definir y justificar sus criterios de selección en base a la necesidad de responder a una pregunta clínica importante. P.e. será importante realizar una revisión sistemática sobre la utilidad de la metformina en el manejo de la hiperglucemia (y hemoglobina glicosilada) en diabéticos tipo 2 (hay unos 650 pacientes incluidos en varios ensayos clínicos randomizados, mientras que la acidosis láctica que puede producir la metformina se incluirá como pregunta importante, pero secundaria ya que afecta a sólo 1 de cada 30.000 pacientes-año.

### Validez y revisión sistemática

Validez: cuán cerca está de la verdad o está libre de sesgos. Cuanta mayor validez, más nos fiaremos de las conclusiones. ¿Tiene Vd. delante una revisión?. No mire a las conclusiones. Primero hágase las siguientes preguntas. La mejor revisión sistemática cumplirá los 6 puntos con alguna omisión menor:

- 1.- ¿Trata de una pregunta clínica bien enfocada?
- 2.- ¿Los criterios de inclusión utilizados para incluir los artículos fueron apropiados?
- 3.- ¿Existe posibilidad de que se haya perdido - no localizado algún estudio importante?
- 4.- ¿Se evaluó la calidad de los estudios incluidos?
- 5.- ¿La evaluación de los estudios es reproducible por

nosotros?

6.- ¿Los resultados de los estudios fueron similares entre los diferentes estudios?

### replicabilidad y revisiones sistemáticas

Ya que la RS es un proceso científico, los métodos de realización deben ser descritos en suficiente detalle como para facilitar su replicación por nosotros y obtención de iguales resultados. Dos de los más efectivos mecanismos de una RS para reducir los sesgos e imprecisión son incluir el máximo de ensayos clínicos relevantes y ofrecer una descripción detallada de los puntos fuertes y limitaciones de la revisión.

Cada una de las siguientes fuentes de ensayos clínicos tienen sus pros y sus contras. Las bases de datos informatizadas son las más coste-efectivas, seguidas de la búsqueda manual, y del contacto directo con investigadores y organizaciones. Las bases de datos bibliográficas varían en contenido, actualización, accesibilidad y coste. Solamente decir que las más importantes son MEDLINE, EMBASE, CANCELIT, Current Contents, y HEALTHStar. Las bases de datos solapan información que se puede encontrar en varias de ellas, pero también puede encontrarse información en una y no en las demás, por lo que siempre se deberá buscar en 2 bases de datos como mínimo. Por ejemplo, si en una revisión sistemática utilizamos sólo MEDLINE, no localizaremos el 40% de los ensayos clínicos existentes.

En 1996 la Colaboración Cochrane presentó una biblioteca electrónica: la *Cochrane Library*. Fue diseñada para ofrecer evidencia científica y ayudar a tomar decisiones clínicas. Pero atención, por ahora sólo incluye ensayos clínicos, lo que la hace particularmente eficaz para evaluar procedimientos terapéuticos o preventivos, pero no para los diagnósticos o el pronóstico.

### ¿CÓMO EVALUARON LOS AUTORES LA CALIDAD DE LOS ESTUDIOS INCLUIDOS?

Existe un refrán en inglés que dice: *garbage in, garbage out*. Si metes basura, obtendrás basura. La capacidad de una revisión sistemática para guiar en las tomas de decisión depende de la calidad de los estudios incluidos. Si la calidad de los estudios incluidos es mala, la calidad de revisión

sistemática que los ha incluido será mala, a menos que nos lo avise, comente y discuta. En caso contrario nos confundirá en nuestra práctica clínica y nos hará perder el tiempo.

Existe una gran controversia sobre cómo hacer para evaluar la calidad de los estudios incluidos en una revisión. La mayoría; sin embargo, se decantan por la evaluación de la calidad, como una estrategia importante para identificar y reducir los sesgos.

### ¿Evaluaron los revisores la calidad?

Esto se puede saber leyendo la sección de métodos. Se pueden referir a calidad, validez o rigor. Si no se encuentran estas palabras deberemos hacer un esfuerzo para localizar algún comentario de los revisores que nos sugiera que han utilizado algún método para valorar más algunos de los estudios incluidos que otros. Si seguimos sin encontrarlo podemos empezar a dudar de la utilidad de la RS que estamos leyendo.

### ¿Cómo fue evaluada la calidad?

La calidad de los estudios incluidos se puede resumir en la validez interna: la confianza en que el diseño del estudio, su realización, su análisis y presentación han minimizado y evitado comparaciones sesgadas entre las intervenciones bajo evaluación.

En el caso de estudios, como los ensayos clínicos existe evidencia científica de que los ensayos clínicos que han seguido una secuencia de randomización adecuada, han sido realizados con doble ciego y este es descrito y adecuado, se han descrito apropiadamente los abandonos del tratamiento y se ha ocultado a los investigadores del ensayo la localización de los códigos de localización de los tratamientos y placebos (*allocation concealment*) presentan una mayor validez.

## FIABILIDAD INTEROBSERVADORES

En cualquier momento mientras se trabaja en una revisión sistemática el revisor debe estar atento a evitar sesgos y errores. Estos se pueden producir en varios puntos de la secuencia de realización de la revisión. Uno de los más importantes es el de selección de estudios para inclusión. Estas dificultades se pueden obviar siendo DOS los revisores independientes que llevan a cabo la selección de los

estudios. Es fácil que sus opiniones sean ocasionalmente divergentes (p.e. pueden tener opiniones diferentes sobre si el doble ciego se aseguró adecuadamente: a unos niños con croup se les pone una inyección de dexametasona en el glúteo y al grupo placebo no se le pincha, pero una enfermera no dependiente del grupo investigador les pone a todos un esparadrapo en el glúteo: ¿es esto verdaderamente un doble ciego?).

Para medir estas divergencias existe un estadístico denominado kappa. Utilizar sólo el porcentaje no es suficiente, porque dos personas estarán a veces de acuerdo sólo en base al azar. El kappa corrige esta posibilidad y asegura la medida de un verdadero acuerdo. Si el acuerdo es bajo, entonces los criterios de selección son ambiguos y difíciles de interpretar (p.e. en metformina y diabetes hemos encontrado ensayos clínicos que incluían pacientes diagnosticados de *non-insulin-dependent-diabetes mellitus*, pacientes con diabetes mellitus II, y pacientes con diabetes tipo 2. ¿Serán estos todos estos pacientes un grupo homogéneo?).

Independientemente de que en una RS nos vayan contando que han realizado el kappa y que fue de un 0,61 o de un 0,87, lo básico son dos cosas: la primera es que al final de la revisión los revisores van a tener siempre un kappa de 1, es decir acuerdo total, en caso contrario no habrían llegado a publicar dicha revisión. Y la segunda y más importante es que el leer que han realizado un kappa nos informa que durante dicha revisión se han realizado controles de calidad interna, lo que nos hará apreciar el esfuerzo que han hecho los revisores para ofrecernos conclusiones válidas.

## SÍNTESIS DE DATOS CUALITATIVOS

Es la información general sobre las características de los estudios incluidos que no son generados por técnicas estadísticas. Puede ser sexo, edad, severidad de la enfermedad o dosis. También el año y lenguaje de publicación de los estudios incluidos, la calidad metodológica y fuentes de financiación (al parecer existe una tendencia a que los resultados no demuestren la misma eficacia de un producto si el que financia un estudio es el FIS, la universidad, o el fabricante del producto).

Esto nos facilita suficiente información de los estudios

incluidos, como para poder enjuiciar el uso apropiado que los revisores van a hacer después con las técnicas cuantitativas o estadísticas. La información cualitativa también nos informa de la posible aplicabilidad de las conclusiones a nuestra población (p.e. la tasa de fracturas de en mujeres postmenopáusicas suecas es 3 veces mayor que la de las españolas o 9 veces mayor que la de las africanas. Es de utilidad saberlo al valorar si vamos a instituir una terapia hormonal sustitutiva o con raloxifeno).

Existen revisiones en que la síntesis de datos no es posible y se quedan en revisiones sistemáticas cualitativas. Los revisores pueden comentarnos que les fue imposible combinar datos, porque sólo existía un estudio o porque los estudios que encontraron no ofrecían datos combinables (p.e. en un estudio de claudicación nos cuentan que los pacientes anduvieron 30 metros más con el medicamento y en otro estudio nos dicen que con el medicamento se hicieron menos *by-pass* aortofemorales. Son resultados, pero tan diferentes que en principio no se pueden combinar). Los revisores probablemente nos ofrecerán entonces una revisión cualitativa y no cuantitativa.

## SÍNTESIS DE DATOS CUANTITATIVOS

La síntesis cuantitativa o metaanálisis ocurre cuando al menos una medida de resultados procedente de, al menos, dos estudios diferentes es combinada estadísticamente para obtener un resultado llamado efecto sumatorio global (en literatura inglesa *overall summary effect*). Existen muchas vías para sumarizar efectos en un metaanálisis. En general, se nos ofrecerán datos de la suma combinatoria del efecto (p.e. el tratamiento disminuye la tensión arterial 4,5 mm Hg) con un rango que suele ser el intervalo de confianza del 95% o también la desviación estándar (IC 95% 2,3-6,7), y todo ello se denomina como antes efecto sumatorio global o también tamaño del efecto (*Effect Size ES*) de la intervención (p.e. el ES es - 4,5 [IC95% 2,3-6,7]).

Existen dos grandes modelos para combinar los efectos. Se diferencian no sólo por su estadística, sino por su filosofía. Son el modelo de efectos fijos EF (*fixed effects model*) y el modelo de efectos aleatorios EA (*random effects model*).

El modelo EF asume que todos los estudios incluidos en la revisión estiman el mismo efecto verdadero del trata-

miento (su ES) y si existen diferencias entre estudios se deben al azar. El modelo EA asume que los estudios incluidos son sólo una muestra aleatoria de un teórico universo de estudios existentes (de los que los revisores han podido localizar muchos o pocos), y que el ES se emplaza aleatoriamente alrededor de un valor central.

Por tanto, el modelo de EF asume variaciones intraestudios mientras que el modelo de EA asume variaciones intraestudios e interestudios. El modelo de EF tiende a mostrar antes y más fácilmente la significación estadística de los datos, mientras el de EA tarda más, es más conservador, pero es más seguro. Por fortuna ambos modelos tienden a dar los mismos resultados, pero cuando se demuestra heterogeneidad es mejor no utilizar el modelo EF y más seguro utilizar el modelo EA, con lo que el ES será más válido, más fiable y más creíble. En los casos en los que la heterogeneidad es muy acusada el verdadero camino para conocer el tamaño del efecto sería la regresión logística.

## MÉTODOS PARA CONSEGUIR RESULTADOS DE LA SÍNTESIS DE DATOS

Existen muchas formas de expresar los resultados del MA. Son el efecto medio sopesado (*Weighted Mean Difference, WMD*) y el efecto medio estandarizado (*Standardised Mean Difference, SMD*) para los datos continuos y la *Odds Ratio* OR, riesgo relativo RR y diferencia de riesgos (*Risk Differences RD*) para los dicotómicos.

**WMD:** un reciente MA evaluando la eficacia de la pentoxifilina para la claudicación intermitente mostró un beneficio del medicamento de 28 metros. Es decir, los pacientes que tomaban el medicamento en lugar de placebo anduvieron 28 metros más.

**SMD:** en el mismo metaanálisis existían estudios que utilizaron metros como unidad de medida, otros segundos y otros pasos. Son diferentes unidades de medidas y mientras existe seguridad de que una yarda son 0,91 metros no hay seguridad de que en un segundo se camine un metro. El SMD homogeneiza unidades diferentes de medida.

La estadística óptima en los estudios de casos y controles es la **OR** mientras que en los estudios de cohortes es el **RR**. Pero hasta ahora los estudios incluidos en los MA

son, en general, ensayos clínicos aleatorizados. Como las diferencias entre OR y RR son más epidemiológicas que estadísticas en los MA de ensayos clínicos se pueden utilizar de forma. En casos de eventos poco frecuentes (fractura, hemorragia, muerte...) se utiliza el RR y la RD. La ventaja de la RD es que su inversa es el número necesario para tratar (NNT), que es muy popular entre los médicos porque representa, de una forma muy cercana y comprensible, cuántos recursos en tiempo y hay que dedicar para evitar una fractura de cadera o un accidente cerebrovascular.

Por ejemplo, según un MA el riesgo relativo de presentar una fractura de cadera entre las mujeres que utilizan alendronato y las que toman placebo es de  $-0,5$  (IC 95%  $-0,8-0,2$ ). Esta cifra puede informarnos de que el efecto puede ser ligeramente beneficioso, aunque sin una significación estadística suficiente. Sin embargo, será más útil para hacernos una composición de lugar si nos dicen que tenemos que tratar a 330 (IC 95%: 150-1.000) mujeres postmenopáusicas con alendronato durante 1 año para evitar una fractura de cadera. O que tenemos que tratar a 33 pacientes con Sintrom durante un año para evitar un ACVA.

## EXAMINANDO LA VARIABILIDAD ENTRE ESTUDIOS

Las diferencias entre estudios incluidos en una revisión sistemática pueden ser también evaluados utilizando técnicas estadísticas. El test de Breslow-Day es utilizado en datos dicotómicos y el de la Q de Cochrane en los datos continuos. Como regla general si leemos que los tests han resultado en una  $p < 0,10$  entonces es que existe una heterogeneidad estadísticamente significativa entre los estudios incluidos y, por tanto, intentar encontrar como los revisores han manejado esta dificultad para combinar los estudios.

## ANÁLISIS DE SENSIBILIDAD

La sensibilidad es definida como la evaluación de la robustez de los resultados de la síntesis estadística estimando y comparando el efecto de la intervención utilizando dentro de la misma revisión sistemática todos los estudios y

luego grupos de estudios.

## ANÁLISIS DE SUBGRUPOS

Trata del análisis de subgrupos particulares de pacientes (p.e. mujeres) entre los estudios, no trata de subgrupos de estudios. Por ejemplo si se observan los estudios de hipolipemiantes se observará que existe un déficit importante de mujeres en la muestra. Para tener resultados estadísticamente significativos para las mujeres se necesitará unir subpoblaciones de mujeres a través de muchos estudios de hipolipemiantes para hacer asunciones válidas para las mujeres.

## ESTABLECIENDO SIGNIFICACIÓN CLÍNICA DE LOS RESULTADOS

Finalmente, nos enfrentamos a la gran cuestión: ¿la estimación del efecto estadísticamente significativa es también necesariamente clínicamente significativa?. ¿Qué significado tiene una OR de 0,35 o un ES de 0,42 en términos de beneficio para el paciente?. La adopción de una nueva terapia dependerá del grado de beneficio, ¿cómo de importante sea un resultado en particular?.

Como hemos visto, un camino de convertir las medidas estadísticas dicotómicas en una medida más comprensible es el número necesario para tratar o NNT. El NNT le indica al lector cuantos pacientes se necesitan tratar con el tratamiento objeto de interés para producir un evento de interés (p.e. se evita que un paciente adicional acabe en silla de ruedas por cada 13 nuevos pacientes con esclerosis múltiple que son tratados con interferón durante 3 años, o bien se da un caso al año de hemorragia importante por cada 99 pacientes en fibrilación auricular tratados con Sintrom).

Las medidas continuas miden el grado de beneficio de una forma más familiar para el médico: metros caminados con pentoxifilina, mg de glucemia o % de hemoglobina glicosilada disminuidos con metformina o mm de Hg en tensión arterial.

Finalmente, como resumen de los pasos indispensables a seguir para enjuiciar la calidad y aplicabilidad de una revisión sistemática cuando nos toca ser lectores de una de ellas, lo podemos encontrar resumido en el siguiente esque-

ma<sup>(5)</sup>:

### 1. Formulación de la pregunta

Concreta. Enfocada. Relacionada con las medidas de resultados que se buscan. Objetivo: ¿qué?, ¿cómo?. Intervención, población, resultados, controles y tipos de estudios incluidos.

### 2. Búsqueda de estudios

Cómo mínimo dos bases de datos completas. Lenguaje. Estudios publicados/ no publicados/ duplicaciones. Fechas.

### 3. Estadística aplicada

Estadística óptima (Odds Ratio, Relative Risk, Risk Difference, número necesario para tratar o para tener un efecto adverso NHN, modelo de efectos fijos FE, modelo de efectos aleatorios RE).

Medida del efecto y su dispersión como media y desviación estándar.

Diferencias de escalas (SMD, WMD).

Homogeneidad - heterogeneidad, test Q.

Análisis de sensibilidad (calidad, antigüedad, muestra) y de subgrupos (edad, sexo, pacientes ambulatorios/ingresados, dosis).

### 4. Conclusiones

Consecuentes con lo anterior. Derivadas de los resultados a la pregunta.

### BIBLIOGRAFÍA

1. Berman S. Otitis Media in Children. *N Engl J Med* 1995;**332**:1560-5.
2. Hood SC, Moher D, Barber GG. The management of intermittent claudication with pentoxifylline: a meta-analysis. *Can Med Assoc J*. 1996; **155**:1053-9.
3. Moher D, Pham B, Aulsebrook M, Saenz A, Hood SC, Barber GG. Pharmacologic Management of Intermittent Claudication: A Meta-analysis of Randomised Trials. *Drugs* 2000;**59**:1057-1070.
4. Aulsebrook M, Sáenz A, Pham B, Kellner JD, Johnson DW, Moher D, Klassen T. The effectiveness of glucocorticoids in treating group: meta-analysis. *BMJ* 1999; **319**:595-600.
5. Klassen TP, Jaddad AR, Moher D. Guides for reading and interpreting systematic reviews. I: getting started, II: How did the authors find the studies and assess their quality? III: How did the authors synthesize the data and make their conclusions?. *Arch Pediatr*

*Adolesc Med* 1998; **152**: 700-4/ 812-7/ 915-20.

### 4. EVIDENCIA CIENTÍFICA RELACIONADA CON EL ASMA

#### EVIDENCIAS EN EL ASMA DEL NIÑO Y ADOLESCENTE

Carlos A. Díaz Vázquez, Gervasio Sánchez Iglesias y GAIBE (Grupo de Asma Infantil Basado en Evidencias)

#### ASMA Y ASISTENCIA BASADA EN LA EVIDENCIA

El asma es una patología que interesa y preocupa a los pediatras. Su prevalencia, tendencia en el tiempo y morbimortalidad justifican plenamente esta preocupación, que es compartida con las familias.

El asma es, además, centro de investigación científica. Todos los meses con independencia de la fuente de búsqueda que se utilice (MEDLINE, Amedeo, Biomed, aparecen publicados 80-100 (incluso más) artículos relevantes en torno a esta enfermedad, tanto de investigación básica (genética, biología, patología), como de epidemiología y clínica.

Ello hace que estar al día en el asma, sea una tarea enormemente complicada, y conocer todo lo que se escribe, imposible.

El profesional sanitario reclama información depurada, lista para ser puesta en práctica en el conjunto de sus pacientes asmáticos, y que ésta sea lo más rigurosa posible y se sustente en las mejores evidencias existentes.

La medicina basada en la evidencia busca la toma de las mejores decisiones que afectan a un paciente individual<sup>(1)</sup>. En este método, el clínico se enfrenta a una cuestión que le surge en la atención a un paciente concreto y busca la mejor respuesta en la bibliografía médica, y actúa en consecuencia (aplicación racional) con los hallazgos. Esta práctica de la MBE, si bien tiene el mérito de la rigurosidad metodológica, se enfrenta a numerosas dificultades a la hora de su aplicabilidad real. El número de dudas que surgen a lo largo de una sola jornada de trabajo, son imposibles de solventar día a día y las cuestiones sin resolver se acumulan.

Se hacen precisas estrategias que permitan a los profesionales sanitarios aplicar las evidencias existentes, ya transformadas en recomendaciones y guías de la práctica a con-