

Evaluación de la evidencia científica

Albert J. Jovell^a y María D. Navarro-Rubio^b

^aAgència d'Avaluació de Tecnologia Mèdica. Servei Català de la Salut. Departament de Sanitat i Seguretat Social. ^bDivisió d'Atenció Primària. Hospital Mútua de Terrassa. Barcelona.

tecnologia sanitària, anàlisis coste-beneficio

El proceso de evaluación de tecnologías sanitarias discurre a través de tres fases complementarias. En la primera fase se evalúan los aspectos técnicos, los posibles riesgos y la eficacia de la tecnología sanitaria en su interacción con los sujetos humanos sanos y enfermos. En esta fase se tienen en cuenta las características de la tecnología a evaluar y de la condición clínica a la que se intenta aplicar. En una segunda fase se evalúan la seguridad, la efectividad y la idoneidad de la tecnología en un contexto sanitario concreto y en el grupo específico de pacientes que se pueden beneficiar de la misma. En la tercera fase de evalúa el impacto global de la tecnología o intervención sanitaria, es decir, las consecuencias éticas, sociales y económicas de su introducción, adopción y difusión en un sistema sanitario concreto. Las dos últimas fases constituyen la denominada fase de contextualización de la evaluación de tecnologías médicas¹.

Las tres fases mencionadas permiten distinguir cuáles son los dos objetivos básicos de la evaluación de las tecnologías sanitarias. El primer objetivo consiste en la valoración de la seguridad, la eficacia, la efectividad y la relación coste-consecuencia derivada de la posible adopción de una tecnología sanitaria en la práctica médica habitual. El segundo objetivo es la valoración de las posibles consecuencias éticas, económicas y sociales que se pueden derivar de la introducción, y posterior adopción, de una nueva tecnología sanitaria o procedimiento médico en el sistema sanitario. El cumplimiento de estos objetivos debería constituir un elemento fundamental en el proceso de toma de decisiones en torno a la introducción y la adopción de una innovación sanitaria en la práctica médica habitual o respecto a la nueva indicación clínica de una nueva tecnología ya adoptada con anterioridad. La información dada por el proceso global de evaluación debería ser determinante en la adopción de una nueva (o existente) intervención sanitaria por parte de los profesionales sanitarios, en su aceptación por parte de la sociedad y en la toma de decisiones relativas a su financiación y/o provisión por los sistemas sanitarios públicos.

Las recomendaciones respecto a la decisión de adoptar o no una nueva tecnología o procedimiento médico formuladas por las agencias e instituciones de evaluación de tecnologías sanitarias deberían surgir del análisis y síntesis de la evidencia disponible en la literatura científica, así como de la valoración de otro tipo de información adicional de carácter clínico, técnico y financiero. El análisis y la síntesis de la información disponible comprende la valoración y la clasificación previa de la evidencia según el rigor científico y la calidad de la misma. Cuando el rigor científico es escaso o controvertido, el informe de evaluación debe indicar de forma explícita la necesidad de llevar a cabo estrategias específicas de investigación. Tales estrategias son imprescindibles para poder recomendar la introducción y la adopción

de la tecnología a evaluar en la práctica médica habitual. La misión del proceso de evaluación es la de valorar si los méritos científicos de una tecnología justifican su uso en humanos para indicaciones clínicas específicas y en contextos sanitarios concretos.

Los objetivos del presente artículo son: 1) describir las principales escalas de clasificación de la evidencia científica disponibles; 2) presentar el esquema de gradación utilizado en la Agència d'Avaluació de Tecnologia Mèdica (AATM), y 3) introducir el proceso de formulación de recomendaciones a partir del análisis y de la síntesis de la evidencia científica.

Niveles de evidencia científica

Las escalas de clasificación de la evidencia científica diferencian de forma jerárquica los distintos niveles de la evidencia en función del rigor científico del diseño del estudio. El esquema de gradación se construye a partir del reconocimiento de la capacidad diferencial de los diversos tipos de diseños para determinar la evidencia científica. Diferencias inherentes a las características del diseño determinan, a su vez, la magnitud y el rigor científico de las recomendaciones en torno a la idoneidad de las condiciones de adopción y difusión de una nueva tecnología sanitaria o procedimiento médico en la práctica clínica habitual.

Hay diferentes escalas de gradación de la calidad de la evidencia científica, todas ellas muy similares entre sí. En la **tabla 1** se describe la clasificación o gradación de la evidencia científica formulada inicialmente en 1979 por la Canadian Task Force on the Periodic Health Examination² y adaptada con posterioridad por el mismo grupo³ y por la U.S. Preventive Services Task Force⁴. En la **tabla 2** se describe una clasificación propuesta por The Swedish Council on Technology Assessment in Health Care (SBU) en función del rigor científico de los diferentes tipos de diseños⁵.

Las clasificaciones de la evidencia científica se basan en el mayor rigor científico de determinadas características del diseño en comparación con otras posibles alternativas metodológicas. Los aspectos del diseño que están asociados a un mayor rigor científico son: la asignación aleatoria a los grupos experimental y control de la intervención sanitaria a evaluar, la existencia de un grupo control concurrente en el tiempo, el sentido prospectivo del estudio, el enmascaramiento de los pacientes e investigadores participantes respecto a la tecnología objeto de estudio y la inclusión en el estudio de un número de pacientes suficiente como para detectar diferencias estadísticamente significativas en la determinación del efecto verdadero de la tecnología o procedimiento que se evalúa. Una descripción de las características del diseño de un estudio que influyen en la calidad y en el rigor científico de los resultados puede encontrarse en la literatura^{6,7}.

Análisis de la evidencia científica

Los esquemas de gradación de la evidencia científica constituyen la primera etapa en el proceso de evaluación de una

Correspondencia: Dr. A.J. Jovell.
Agència d'Avaluació de Tecnologia Mèdica.
Travessera de les Corts, 131-159. Pavelló Ave Maria. 08028 Barcelona.

Manuscrito aceptado el 15-4-1995

Med Clin (Barc) 1995; 105: 740-743

tecnología sanitaria. Las clasificaciones descritas en el apartado anterior tienen ciertas limitaciones, al no poder discriminar entre las posibles diferencias de calidad y rigor científico que puedan existir entre dos estudios situados en el mismo nivel de evidencia científica, por ejemplo 2 ensayos aleatorizados y controlados, cuando éstos producen resultados contradictorios. Podría darse el caso de que un estudio de menor gradación científica presentara una mayor calidad y rigor en comparación con otro situado en un nivel superior en el esquema de clasificación. La validez interna de un estudio caso-control bien diseñado podría ser mayor que la de un ensayo controlado no aleatorizado en el que se produjera un sesgo de selección en la asignación de los pacientes.

Con el objeto de mejorar los esquemas de gradación de la evidencia científica antes descritos, se propone un sistema de clasificación que tiene en cuenta además del diseño del estudio una valoración específica de su calidad. El esquema de gradación propuesto por la AATM se describe en la **tabla 3**. Esta clasificación considera en el análisis de la evidencia, además del tipo de diseño, las denominadas condiciones de rigurosidad científica. Estas condiciones permiten seleccionar, cuando se producen resultados dispares, el mejor estudio entre los de diseño similar o de igual nivel de evidencia. El esquema de gradación de la evidencia científica propuesto en la **tabla 3** es similar a los descritos en las **tablas 1 y 2**. La diferencia fundamental que presenta, aparte de la inclusión de las condiciones de rigurosidad, es la consideración del metaanálisis como el diseño de mayor rigor científico⁸. Esto se debe a que el metaanálisis puede presentar la mejor combinación de validez interna u homogeneidad del diseño y validez externa o capacidad de generalizar los resultados a la población general. A pesar de sus ventajas, el metaanálisis no está exento de sesgos ni limitaciones^{9,10}. En el caso del metaanálisis o revisión cuantitativa, las condiciones de rigurosidad señalan un mayor rigor científico si se analizan datos de pacientes individuales –megaanálisis–¹¹, se tienen en cuenta las posibles diferencias entre estudios –metarregresión–¹², los estudios incluidos son homogéneos y consistentes con respecto a los resultados, los ensayos seleccionados han sido valorados mediante un protocolo de calidad¹³ y los resultados obtenidos se han confirmado por más de un método de análisis estadístico¹⁴. En el caso de que los estudios seleccionados para el metaanálisis fueran heterogéneos o de baja calidad, habría que plantear la posibilidad de que el rigor científico y la calidad de un único ensayo aleatorizado y controlado fuera mayor que la del metaanálisis. Hay que tener en cuenta la heterogeneidad en el nivel de calidad de los ensayos controlados y aleatorizados^{15,16}. Las restantes categorías del esquema de gradación propuesto en la **tabla 3** distinguen el mayor rigor científico de los diseños que presentan una mejor combinación de vali-

dez interna y de validez externa. Un poder estadístico que avale la significación estadística de los resultados, la asignación aleatoria de la intervención a evaluar y la presencia de un grupo control concurrente indican un mayor control del sesgo y del error sistemático en el diseño del estudio y, por lo tanto, una mayor validez interna^{17,18}. El carácter multicéntrico del estudio debería suponer la participación de múltiples instituciones sanitarias, de pacientes de diversas procedencias y de diferentes estilos de práctica clínica, lo que garantizaría una mayor validez externa. La valoración del estudio se complementa con el análisis de la calidad del mismo con respecto a diferentes factores relacionados con el tipo de diseño y la metodología del análisis estadístico.

TABLA 2

Clasificación de la evidencia científica según el rigor científico^a

- 1 Ensayo controlado y aleatorizado con una muestra grande
- 2 Ensayo controlado y aleatorizado con una muestra pequeña
- 3 Ensayo no aleatorizado con controles coincidentes en el tiempo
- 4 Ensayo no aleatorizado con controles históricos
- 5 Estudio de cohorte
- 6 Estudio de casos y controles
- 7 Estudios transversales
- 8 Vigilancia epidemiológica (bases de datos o registros)
- 9 Serie consecutiva de casos
- 10 Notificación de un caso aislado (anécdota)

^aTomado de Goodman C⁵. De rigor máximo (1) a rigor mínimo (10).

TABLA 3

Niveles de calidad de la evidencia científica. Esquema propuesto por la AATM

Nivel	Tipo de diseño	Condiciones de rigurosidad científica [*]
I	Metaanálisis de ensayos controlados y aleatorizados	No heterogeneidad Diferentes técnicas de análisis Metarregresión Megaanálisis Calidad de los estudios
II	Ensayo controlado y aleatorizado de muestra grande	Evaluación del poder estadístico Multicéntrico Calidad del estudio
III	Ensayo controlado y aleatorizado de muestra pequeña	Evaluación del poder estadístico Calidad del estudio
IV	Ensayo prospectivo controlado no aleatorizado	Controles coincidentes en el tiempo Multicéntrico Calidad del estudio
V	Ensayo prospectivo controlado no aleatorizado	Controles históricos Calidad del estudio
VI	Estudios de cohorte	Multicéntrico Apareamiento Calidad del estudio
VII	Estudios de casos y controles	Multicéntrico Calidad del estudio
VIII	Series clínicas no controladas Estudios descriptivos: vigilancia epidemiológica, encuestas, registros, bases de datos Comités de expertos	Multicéntrico
IX	Anécdotas o casos únicos	

^{*}Complementado por las características descritas en la tabla de evidencia. De mayor (I) a menor (III) rigor científico.

TABLA 1

Niveles de calidad de la evidencia científica^a

I	Evidencia obtenida a partir de al menos un ensayo aleatorizado y controlado diseñado de forma apropiada
II-1	Evidencia obtenida a partir de ensayos controlados no aleatorizados y bien diseñados
II-2	Evidencia obtenida a partir de estudios de cohorte o caso-control bien diseñados, realizados preferentemente en más de un centro o por un grupo de investigación
II-3	Evidencia obtenida a partir de múltiples series comparadas en el tiempo con o sin intervención ^b
III	Opiniones basadas en experiencias clínicas, estudios descriptivos o informes de comités de expertos

^aDe mayor (I) a menor (III) calidad de diseño y rigor científico. Tomado de U.S. Preventive Task Force³. ^bEste tipo de evidencia también incluye resultados «dramáticos» producidos en experimentos incontrolados, como son los derivados de la introducción de la penicilina en los años cuarenta.

TABLA 4

Idoneidad de las recomendaciones según la calidad de la evidencia científica^a

Grado de las recomendaciones	Niveles de calidad
A: existe ADECUADA evidencia científica para recomendar la adopción de la tecnología	I II-1
B: existe CIERTA evidencia científica para recomendar la adopción de la tecnología	II-1 II-2
C: existe una INSUFICIENTE evidencia científica, por lo que la decisión de adoptar la tecnología debe basarse en otros criterios	II-3 III
D: existe una CIERTA evidencia científica para recomendar la no adopción de la tecnología	II-1 II-2
E: existe una ADECUADA evidencia científica para recomendar la no adopción de la tecnología	I II-1

^aAl no abordar un ejemplo concreto el contenido de esta tabla es sólo orientativo, dado que la relación entre las recomendaciones y los niveles de evidencia científica tienen en cuenta el impacto de la enfermedad en la población, las características de la intervención y, sobre todo, el tipo de diseño. Adaptado de U.S. Preventive Task Force⁴.

TABLA 5

Relación entre los niveles de calidad de la evidencia científica y el grado de recomendaciones^a

Nivel de calidad de la evidencia científica	Grado de recomendación ^b
I: ensayos aleatorizados con una muestra grande y resultados bien definidos (y un riesgo bajo de error estadístico tipo α y β)	A
II: ensayos aleatorizados con una muestra pequeña (y un riesgo moderado a alto de error estadístico tipo α y β)	B
III: estudios no aleatorizados, controles concurrentes en el tiempo	C
IV: estudios no aleatorizados, controles históricos	
V: estudios no controlados, series clínicas	

^aEl esquema de recomendaciones se define en la tabla 4. ^bTomada de Sackett DL²³.

La evaluación de la evidencia científica se completa cuando, además de valorar el tipo de diseño, se tienen en cuenta otras consideraciones relacionadas con la metodología del estudio y las características de la intervención sanitaria que se evalúa. Las tablas de evidencia contribuyen a contextualizar las recomendaciones formuladas en función de otras consideraciones no inherentes al diseño del estudio y relacionadas con la tecnología sanitaria objeto de evaluación. Éstas son: la condición clínica o el problema de salud objeto de intervención, las características de los pacientes y de los proveedores, el tipo de intervención y la forma de provisión, la relevancia clínica de los posibles resultados clínicos y el tipo de institución sanitaria donde se va a adoptar la tecnología. Las tablas antes descritas ponen en evidencia toda la información que se debe considerar en el proceso de evaluación de una tecnología sanitaria antes de recomendar su introducción y adopción en la práctica médica habitual. Ejemplos de esas tablas están en la literatura^{19,20}.

La información descrita en las tablas de evidencia permiten la aplicación de las técnicas de síntesis de la evidencia con el propósito de valorar la utilización de la tecnología en grupos específicos de la población y/o circunstancias clínicas concretas –análisis de decisiones clínicas– o el impacto de la misma cuando se tienen en cuenta otro tipo de factores, como son los costes –análisis coste-efectividad–. En este sentido, el metaanálisis también debería considerarse una técnica de síntesis, debido a que no podría aplicarse sin la existencia de ensayos controlados y aleatorizados o estudios epidemiológicos a combinar. Otras técnicas de síntesis, más complejas y poco comunes, se basan en la elaboración de modelos de simulación o en la aplicación de modelos matemáticos de tipo multivariable que permiten clarificar las in-

certidumbres y ausencias de conocimiento en la interpretación de la evidencia científica. Una última técnica de síntesis, aún en fase de desarrollo, es la denominada síntesis de diseños cruzados que combina los resultados de ensayos aleatorizados y controlados con estudios de tipo observacional²¹. El análisis de la evidencia científica se completa con la valoración de las implicaciones éticas y de la aceptabilidad social de la tecnología objeto de evaluación.

Formulación de recomendaciones

A partir del análisis y de la clasificación de la evidencia científica descrita en la tabla 1, pueden formularse recomendaciones en torno a la idoneidad de las condiciones de adopción de una tecnología sanitaria o procedimiento médico²². El esquema de gradación de las recomendaciones más utilizado en la literatura médica se describe en la tabla 4⁴. La clasificación descrita califica a la evidencia científica como adecuada (grado A) cuando permite formular recomendaciones conclusivas, como cierta (grado B) cuando es factible expresar recomendaciones no concluyentes y como insuficiente (grado C) cuando no se puede manifestar ningún tipo de valoración. Un ejemplo de la aplicabilidad clínica de los esquemas de recomendaciones basados en la evidencia se describe en la tabla 5²³. Este último esquema ha sido modificado por sus autores por otro de mayor complejidad que considera la inclusión del metaanálisis, así como la valoración, en este último tipo de diseño, de la significación clínica del efecto terapéutico combinado y del grado de heterogeneidad de los estudios seleccionados²⁴.

Las futuras escalas de recomendaciones basadas en la calidad de la evidencia científica deberían hacer más énfasis en la valoración de las condiciones de aplicabilidad del diseño y la metodología de análisis estadístico. La valoración de la aplicabilidad del diseño considerará los diferentes aspectos que definen la validez interna y externa del estudio^{17,18}. El análisis estadístico permite valorar la magnitud del beneficio terapéutico estimado de la tecnología, a partir del cálculo del efecto clínico combinado mediante técnicas metaanalíticas y/o a través de la estimación del número de pacientes en los que se necesita aplicar ese procedimiento con objeto de obtener los resultados clínicos deseados²⁵.

Las escalas de recomendaciones revisadas en este apartado sólo califican como buena a la evidencia científica procedente de metaanálisis y de ensayos aleatorizados y controlados. Esto supone un obstáculo en la evaluación de tecnologías o procedimientos en los que la asignación aleatoria no es ética, no existe tratamiento estándar con que comparar o el diseño aleatorizado y controlado precisa de un gran número de pacientes y/o un largo periodo de seguimiento para proporcionar la evidencia necesaria y/o no permite reproducir las condiciones de la práctica clínica habitual. Estas últimas situaciones son frecuentes en el proceso de evaluación de tecnologías sanitarias. Los posibles inconvenientes y ventajas de adoptar o evitar el ensayo controlado y aleatorizado como estándar ideal de diseño en la evaluación de las tecnologías o procedimientos sanitarios, en comparación a otros diseños, pueden encontrarse en la literatura médica²⁶. De todas maneras, la evaluación de una tecnología sanitaria o procedimiento médico debería considerar, siempre que sea posible, los resultados procedentes de diseños controlados²⁷. Estos diseños permiten comparar los beneficios, riesgos y costes de la tecnología a evaluar con los de las posibles alternativas en condiciones de aplicación equivalentes. A pesar del debate existente en torno a cuál debería ser el diseño ideal, el proceso de evaluación de tecnologías sanitarias supone la consideración de toda la información disponi-

bles y su clasificación mediante tablas de evidencia científica que permiten considerar aspectos complementarios al método utilizado. El informe de evaluación debería explicar los criterios en los que se basan las recomendaciones formuladas, incluyendo la procedencia de la información analizada, el esquema utilizado para clasificar la evidencia científica, la descripción de las tablas de evidencia y la interpretación de su contenido, la metodología de síntesis de la evidencia y el análisis de las posibles consecuencias económicas, sociales y éticas que se pueden anticipar como resultado de la adopción y difusión de la tecnología sanitaria evaluada.

Conclusiones

Las recomendaciones sobre la introducción, la adopción y la difusión de una tecnología sanitaria en la práctica médica habitual deberían estar avaladas por la confirmación científica de su seguridad, eficacia, efectividad y relación coste-consecuencia. La aplicación de una tecnología o procedimiento sanitario (nuevo o adoptado) para una indicación clínica cuya idoneidad no está probada científicamente plantea problemas de tipo ético que precisan ser estudiados y abordados mediante propuestas concretas de evaluación. El proceso de evaluación de tecnologías sanitarias supone la formulación de recomendaciones basadas en el análisis, la síntesis y la contextualización de la evidencia científica. Las escalas de clasificación de la evidencia científica contribuyen a determinar la calidad y el rigor científico de la información disponible. La gradación de la evidencia científica permite establecer recomendaciones objetivas respecto al uso apropiado de la tecnología evaluada. Las tablas de evidencia científica permiten la consideración en el proceso evaluador de otro tipo de características necesarias para contextualizar las recomendaciones formuladas. Las técnicas basadas en el análisis de decisiones clínicas y en el análisis económico contribuyen a completar el proceso evaluador al considerar otros factores, diferentes a los resultados procedentes de la evidencia científica, determinantes en la adopción y difusión de la tecnología. El proceso evaluador se amplía y completa con la inclusión de una valoración de las consecuencias sociales, éticas y económicas derivadas de la introducción, adopción y difusión de la tecnología en un sistema sanitario concreto²⁸. Este proceso puede culminar con la elaboración y aplicación de las guías de práctica clínica²⁹ y la determinación del nivel de cobertura sanitaria del sistema.

En el caso de que la gradación de la información disponible indique la carencia, insuficiencia o controversia de la evidencia científica respecto a la utilización de una tecnología concreta no debería recomendarse su adopción universal en la práctica médica habitual. Su utilización tendría que restringirse al contexto de un protocolo de investigación experimental en instituciones sanitarias seleccionadas que permitiera obtener la evidencia que se quiere evaluar. En este sentido, se debería reflexionar en torno a cuáles son los principios éticos que justifican la adopción y difusión de una tecnología sanitaria en la práctica médica habitual cuando no existe evidencia científica sobre los posibles riesgos y beneficios derivados de su aplicación a sujetos humanos. En otras palabras, sin conocer los resultados del proceso de evaluación no se debería recomendar la provisión de una tecnología³⁰. Quizá, si Hipócrates tuviera que rehacer su código ético, al primer principio: «no hacer daño» le añadiría un pie de nota que dijera: «antes de utilizar, primero investigar y evaluar». La producción y la evaluación de la evidencia científica constituyen la primera etapa del proceso.

Agradecimiento

Los autores agradecen los comentarios y las sugerencias realizadas a una versión preliminar del artículo por Ángela Boland, Josep M. Borràs, Alicia Granados y Laura Sampietro-Colom.

REFERENCIAS BIBLIOGRÁFICAS

- Granados A. La evaluación de tecnologías médicas. *Med Clin (Barc)* 1995; 104: 581-585.
- Canadian Task Force on the Periodic Health Examination. The Periodic Health Examination. *Can Med Ass J* 1979; 121: 1.193-1.254.
- Wolff SM, Battista RN, Anderson GM et al. Assessing the clinical effectiveness of preventive maneuvers: analytic principals and systematic methods in reviewing evidence and developing clinical practice recommendations. A report by the Canadian Task Force on the Periodic Health Examination. *J Clin Epidemiol* 1990; 43: 891-905.
- U.S. Preventive Task Force. Guide to clinical preventive services: an assessment of the effectiveness of 169 interventions. Baltimore: Williams and Wilkins, 1989.
- Goodman C. Literature Searching and evidence interpretation for assessing health care practices. Estocolmo: The Swedish Council on Technology Assessment in Health Care, 1993.
- Álvarez-Dardet C, Bolívar F, Porta-Serra M. Tipos de estudios. *Med Clin (Barc)* 1987; 89: 296-301.
- Fowks FGR, Fulton PM. Critical appraisal of published research: introductory guidelines. *Br Med J* 1991; 302: 1.136-1.140.
- Chalmers TC, Lau J. Meta-analytic stimulus for changes in clinical trials. *Stat Meth Med Res* 1993; 2: 161-172.
- Chalmers TC. Problems induced by meta-analyses. *Stat Med* 1991; 10: 971-980.
- Eysenck H. Meta-analysis and its problems. *Br Med J* 1994; 309: 789-792.
- Early Breast Cancer trialists's collaborative Group. Treatment of early breast cancer Volume 1. Worldwide evidence 1985-1990. Oxford: Oxford University Press, 1990.
- Kasiske BL, Kalil RSN, Ma JZ et al. Effect of antihypertensive therapy on the kidney in patients with diabetes: a meta-regression analysis. *Ann Intern Med* 1993; 118: 129-138.
- Chalmers TC, Smith M, Blackburn B et al. A method for assessing the quality of a randomized controlled trial. *Control Clin Trials* 1981; 2: 31-49.
- Dickersin K, Berlin JA. Meta-analysis: state-of-the science. *Epidemiol Rev* 1992; 14: 154-176.
- Gotzsche PC. Methodology of overt and hidden bias in reports of 196 double-blind trials of nonsteroidal antiinflammatory drugs in rheumatoid arthritis. *Control Clin Trials*, 1989; 10: 31-56.
- Sacks HS, Berier J, Reitman D, Pagano D, Chalmers TC. Meta-analysis of randomized control trials: and update of the quality and methodology. En: Bailar JC, Mosteller F, editores. *Medical uses of statistics (2.ª ed.)*. Boston: NEJM Books, 1992: 427-442.
- Porta-Serra M, Álvarez-Dardet C, Bolívar F et al. La calidad de la información clínica (I): validez. *Med Clin (Barc)* 1987; 89: 741-747.
- Plasencia A, Porta-Serra M. La calidad de la información clínica (II): significación estadística. *Med Clin (Barc)* 1988; 90: 122-126.
- Eddy DM. A manual for assessing health practices and designing practice policies: the explicit approach. Filadelfia: American College of Physicians, 1992.
- Morris RD, Lau J, Arena NJ, Nardine FE, Chalmers TC. A clinical trials database as a research tool in health care. *Online J. Current Clin Trials* 1992; 17 de julio de 1992: 14.
- Droitcours J, Silberman G, Chelmsky E. Cross-design synthesis. A new form of meta-analysis for combining results from randomized clinical trials and medical-practice databases. *Int J Technol Assess Health Care* 1993; 9: 440-449.
- U.S. Preventive Services Task Force. Screening for adolescent idiopathic scoliosis. Policy statement. *JAMA* 1993; 269: 2.664-2.666.
- Sackett DL. Rules of evidence and clinical recommendations on the use of antithrombotic agents. *Chest* 1989; 95 (Supl): 2-3.
- Cook DJ, Guyatt GH, Laupacis A, Sackett DL. Rules of evidence and clinical recommendations on the use of antithrombotic agents. *Chest* 1992; 102 Supl 305-311.
- Laupacis A, Sackett DL, Roberts RS. Assessment of clinically useful measures of the consequences of treatment. *N Engl J Med* 1988; 318: 1.728-1.733.
- Warren KS, Mosteller F, editores. *Doing more good than harm: the evaluation of health care interventions*. Nueva York: New York Academy of the Sciences, 1993.
- Bonfill X, Porta-Serra M. ¿Es necesario un grupo control? *Med Clin (Barc)* 1987; 89: 429-435.
- Granados A, Borràs JM. Technology assessment in Catalonia: integrating economic appraisal. *Soc Sci Med* 1994; 38: 1.643-1.646.
- Jovell AJ, Navarro-Rubio MD. Guías de práctica clínica. *FMC* 1995; 2: 152-156.
- Jennet B. Health technology assessment. The rule should be «no evaluation - no technology». *Br Med J* 1992; 305: 67-68.