

## Evaluación crítica de la evidencia científica

A. ARTEAGA DOMÍNGUEZ\* y C. GARCÍA GONZÁLEZ\*\*

\*Servicio de Rehabilitación. Hospital de Cabueñes. Gijón (Asturias). \*\*Médico Rehabilitador. Equipo de Valoración de Incapacidades del Instituto Nacional de la Seguridad Social. Asturias.

**Resumen.**—La evaluación crítica es una forma estructurada de leer los artículos científicos, lo que permite extraer las características clave de los mismos. Constituye uno de los elementos centrales de la Medicina Basada en la Evidencia. Un grupo de expertos ha desarrollado las «guías para usuarios de la literatura médica», en las que se emplean diferentes cuestionarios para los diferentes tipos de publicaciones; todos los cuestionarios están organizados alrededor de tres preguntas básicas: validez, importancia de los resultados y aplicabilidad. Sin embargo, estas guías no contemplan todos los posibles diseños de investigación clínica, entre ellos muchos de los diseños alternativos que existen para conocer la eficacia y efectividad del tratamiento rehabilitador. Además, en los artículos de rehabilitación las medidas de resultados son complejas y diversas, debiéndose comprobar que se han escogido las más relevantes de acuerdo con el objetivo del estudio y, sobre todo, que dichas medidas tengan una aceptable fiabilidad y validez.

**Palabras clave:** *Evaluación crítica. Medicina basada en la evidencia. Rehabilitación.*

### CRITICAL APPRAISAL OF SCIENTIFIC EVIDENCE

**Summary.**—Critical appraisal is a structured form of reading scientific articles, which makes it possible to obtain their key characteristics. It constitutes one of the central elements of Evidence Based Medicine. One group of experts on Evidence Based Medicine has developed the «guidelines for medical literature users», in which different questionnaires are used for different types of publications. All the questionnaires are organized on the basis of three basic questions: validity, importance of the results and applicability. However, these guidelines do not contemplate all the possible designs of clinical investigation, among which many of the alternative designs that exist to know the efficacy and effectivity of the rehabilitation treatment are included. In addition, the measurements of the results are complex and different in the rehabilitation articles, and it must be verified if the most relevant ones have been chosen according to the study objective and, above all, that these measurements have an acceptable reliability and validity.

**Key words:** *Critical appraisal. Evidence based medicine. Rehabilitation.*

### INTRODUCCIÓN

La forma ideal de practicar la Medicina Basada en la Evidencia (MBE) es el empleo de resúmenes producidos por otros, como las revisiones sistemáticas/meta-análisis, y de protocolos basados en la evidencia, como las guías de la práctica clínica. Pero hasta que no se disponga de estas fuentes de una forma generalizada, la capacidad de evaluar críticamente las diferentes publicaciones tiene un gran valor (1, 2).

La evaluación crítica de la evidencia científica es el tercer paso en la práctica de la MBE, tras formular una pregunta en relación al problema del paciente y haber investigado en la literatura los artículos más relevantes. Permite al clínico decidir si un artículo es útil para su práctica clínica, transformándose de esta forma de un espectador pasivo en un «consumidor inteligente» de la bibliografía. Pero para ello es preciso tener unos conocimientos básicos en epidemiología clínica y estadística (1-4).

El presente artículo tiene como *objetivo* servir de introducción en la evaluación crítica de la literatura médica, con especial énfasis en los artículos de Rehabilitación.

### TIPOS DE PUBLICACIONES

Un primer objetivo en toda evaluación crítica es comprobar que se ha escogido el mejor diseño para responder a la pregunta clínica (5). Los diseños básicos de investigación de los estudios primarios se clasifican de acuerdo con el nivel de calidad de la evidencia científica. Las escalas de gradación existentes coinciden en el siguiente orden de mayor a menor rigor científico de los diseños (6): ensayo clínico aleatorizado (ECA), ensayo clínico no aleatorizado, estudio de cohorte, estudio caso-control, estudio transversal, serie de casos y caso clínico. Según el rigor científico el diseño estará más o menos expuesto a la existencia de sesgos.

Aunque se suele considerar que la evidencia científica más sólida es la que surge de los ECA (o de las revisiones de los mismos, las revisiones sistemáticas).

TABLA 1. Diseños de investigación y preguntas clínicas (2, 5).

Pregunta clínica	Diseño principal	Otros diseños
– Biología o factores de riesgo	Estudios de cohorte y caso-control	Estudio transversal
– Pronóstico	Estudio de cohorte	Estudio transversal
– Tests diagnóstico o de <i>screening</i>	Estudio transversal	Estudios longitudinales
– Tratamiento	ECA (*) y revisiones sistemáticas/meta-análisis	Ver tablas 2 y 4

(\*) ECA: Ensayo clínico aleatorizado.

cas/meta-análisis), esto sería cierto si lo que intentamos responder es a una pregunta sobre el tratamiento. Otros diseños de investigación pueden ser los más apropiados para responder a otro tipo de pregunta clínica (tabla 1) (2, 5).

La MBE ha propuesto una jerarquía de la evidencia científica para las decisiones de tratamiento (tabla 2) (7). Este aspecto de la jerarquía es controvertido, y los mismos promotores de la MBE afirman que no todas las preguntas sobre tratamiento requieren un ECA; el ECA sería el diseño más adecuado para conocer la eficacia (en situaciones ideales de un ensayo clínico) pero no la efectividad (en condiciones reales) de un tratamiento, en cuyo caso serían más adecuados otros diseños como los estudios observacionales (estudios de cohorte y caso-control) (4, 8, 9).

## EVALUACIÓN CRÍTICA Y MBE

En el ámbito de la MBE se ha desarrollado un método simple y estructurado de leer la literatura médica, que consiste en el empleo de diferentes cuestionarios para los diferentes tipos de publicaciones (1, 5).

Todos los cuestionarios que emplea la MBE están organizados alrededor de tres preguntas básicas: validez, importancia de los resultados y aplicabilidad. Inicialmente el clínico puede utilizar unos criterios primarios de validez para cribar rápidamente los artículos por el Título y el Resumen, por ejemplo en la pantalla del ordenador, y si se cumplen dichos criterios seguir adelante en la evaluación más profunda del Material-Método y Resultados (3, 10).

La *validez* (validez interna) depende del rigor científico del diseño y de la ejecución del estudio. Respecto a la *importancia de los resultados* la MBE ha contribuido con aportaciones importantes que mejoran la comprensión de la información relativa a los resultados, como los conceptos de «número que es necesario tratar» (NNT) en los estudios sobre tratamiento y de «cociente de probabilidad» en los estudios diagnósticos. Por último, la *aplicabilidad* (validez externa) permite decidir si los resultados del estudio son generalizables a otras situaciones clínicas, en particular nuestro caso clínico, y debe tener en cuenta

los beneficios y riesgos, los costes asociados, y de forma importante los valores y preferencias de los pacientes (5, 7, 11).

La evaluación crítica de la literatura es uno de los elementos centrales de la MBE (11). Para su aprendizaje se han publicado diversos materiales, destacando una serie de artículos en la revista *Journal of the American Medical Association* (JAMA) con el título «guías para usuarios de la literatura médica», que comenzó en 1993 (10) y se sigue publicando (7). Estas guías proporcionan a los clínicos estrategias y herramientas para interpretar e integrar la evidencia de la investigación publicada en el cuidado de los pacientes, e incluye guías sobre el tratamiento (12, 13), el diagnóstico (14, 15), el pronóstico (16), las revisiones sistemáticas (17) y las guías de la práctica clínica (18, 19), entre otras. Están actualmente disponibles en Internet (<http://www.cche.net/>), y recogidas en parte en el libro de Sackett et al (2).

Otros materiales son la serie de cuatro artículos titulada «estadística básica para clínicos» (20-23), que complementa las guías para usuarios de la literatura médica, y la serie de diez artículos sobre aspectos básicos de la MBE con el título «cómo leer un artículo» (24-33). También se dispone en Internet de información sobre un programa de habilidades en lectura crítica, denominado CASP (Critical Appraisal Skills Program), basado en las guías JAMA (<http://www.phru.org/caspinternational/>); en nuestro país, CASPe (CASP España) organiza talleres que enseñan habilidades en la lectura crítica de la literatura científica.

TABLA 2. Jerarquía de la evidencia en estudios de tratamiento propuesta por la MBE (\*) (7).

– ECA N-de-1 (***)
– Revisiones sistemáticas de ECA (**)
– ECA
– Revisiones sistemáticas de estudios observacionales
– Estudios observacionales
– Estudios fisiológicos
– Observaciones clínicas no sistemáticas

(\*) MBE: Medicina basada en la evidencia.

(\*\*) ECA: Ensayo clínico aleatorizado.

(\*\*\*) ECA N-de-1: ECA en un solo individuo.

TABLA 3. Problemas de los ECA (\*) en rehabilitación (43).

- Los pacientes en Rehabilitación frecuentemente tienen *múltiples desórdenes*, lo que dificulta la selección de los sujetos para un ensayo clínico.
- La mayoría de los tratamientos en Rehabilitación tienen *múltiples componentes*, difíciles de cuantificar: terapeutas que aplican el tratamiento, calidad y cantidad de los tratamientos, factores medioambientales, etc.
- La *cumplimentación* del tratamiento rehabilitador es difícil de controlar.
- Existe un riesgo de *contaminación* cuando los mismos terapeutas proporcionan el tratamiento rehabilitador tanto al grupo experimental como al de control.
- El *enmascaramiento* (de los pacientes, de los terapeutas y de los evaluadores) es más difícil de conseguir.
- Las *medidas de resultados* son a menudo diversas, complejas y valorables a largo plazo.

(\*) ECA: Ensayo clínico aleatorizado.

Al margen de la lista de criterios para la evaluación crítica de la literatura producida por expertos en MBE destacamos una «guía introductoria» publicada en 1991 por Fowkes y Fulton (34) que, aunque

TABLA 4. Clasificación de los diseños de investigación de los estudios primarios sobre tratamiento (43-50).

- Estudios de eficacia (diseños experimentales): ECA (\*), ensayo clínico secuencial (\*\*), ECA N-de-1 (\*\*\*).
- Estudios de efectividad o investigación de resultados (\*\*\*\*):
  - Diseños cuasiexperimentales: ensayo controlado no aleatorizado, ensayo antes-después no controlado, diseño de regresión-discontinuidad, diseño de series temporales interrumpidas, diseño de caso único.
  - Diseños no experimentales (observacionales): estudios analíticos (estudios de cohorte y caso-control), estudios descriptivos (estudios de casos).
  - Investigación cualitativa.

(\*) ECA: Ensayo clínico aleatorizado.

(\*\*) Ensayo clínico secuencial: ECA donde la inclusión de participantes y el análisis de resultados se efectúa progresivamente, suspendiéndose el ensayo cuando los datos posibilitan una decisión sobre el tratamiento.

(\*\*\*) ECA N-de-1: ECA en un solo individuo.

(\*\*\*\*) Investigación de resultados: estudia los resultados finales de los cuidados de la salud, es decir, el efecto que el cuidado médico tiene en la salud y el bienestar de los pacientes y las poblaciones. En el diseño de regresión-discontinuidad la asignación a los diferentes grupos no es aleatoria sino basada en criterios estrictos y cuantitativos (como el grado de incapacidad). En el diseño de series temporales interrumpidas lo esencial es la presencia de medidas repetidas y la introducción en este proceso de un cambio experimental. El diseño de caso único es el diseño anterior en un solo sujeto. La investigación cualitativa es el diseño idóneo cuando se intenta conocer los resultados desde la perspectiva de los pacientes.

no cubre todos los diseños de estudios, sirve para recordar las características fundamentales que deben tenerse en cuenta cuando se evalúa un artículo. En los últimos años se han desarrollado unas propuestas o guías específicas para la publicación de los ensayos clínicos y de los metaanálisis, denominadas declaración CONSORT (Consolidated Standards of Reporting Trials) (35) y declaración QUORUM (Quality of Reporting of Meta-analyses) (36), que tienen como objetivo mejorar la calidad de la publicación de estos estudios así como facilitar al lector su evaluación crítica.

Los resultados de cada evaluación crítica se pueden resumir en una hoja de presentación de datos, y en el propio Servicio o Centro de trabajo se podría encontrar un mecanismo eficiente de almacenamiento y recuperación de las mismas (11). La Universidad de McMaster creó un modelo impreso denominado CAT (Critically Appraised Topic) (37), que el Centro de la Medicina Basada en la Evidencia de la Universidad de Oxford ha incorporado a su página web (<http://cebmljr2.ox.ac.uk/>).

Lo distintivo de la MBE es la integración de la mejor evidencia científica con la experiencia clínica, teniendo en cuenta los valores y preferencias de los pacientes (2, 5). Los programas de enseñanza sobre evaluación crítica de la literatura han demostrado que pueden mejorar la competencia de los participantes en este campo (conocimientos en epidemiología y estadística) (38-40), y que también pueden contribuir al objetivo último de influir en el proceso de decisión clínica (41).

## EVALUACIÓN CRÍTICA EN REHABILITACIÓN

Una prioridad en Rehabilitación es establecer la utilidad de los tratamientos que se aplican (42). Aunque el ECA es el diseño más potente para responder a una pregunta sobre tratamiento, su empleo en Rehabilitación es más complejo que en otros campos de la Medicina (tabla 3) (43). Además, los ECA pueden ser éticamente problemáticos, excesivamente caros o no factibles (43, 44). Por otro lado, como se indicó con anterioridad, es el diseño más adecuado para conocer la eficacia pero no la efectividad del tratamiento (4, 8, 9, 44).

Existen otros muchos diseños alternativos, que incluyen otros estudios experimentales y la denominada investigación de resultados o estudios de efectividad, diseños que en su mayoría no vienen contemplados en las guías de la MBE pero que debemos conocer para su correcta evaluación crítica (44-50). En la tabla 4 se expone una posible clasificación de los mismos.



Por la naturaleza de nuestra especialidad las medidas de resultados son complejas y diversas, e incluyen medidas fisiológicas, escalas de valoración para los diferentes conceptos del modelo de discapacidad de la Organización Mundial de la Salud (deficiencia, incapacidad y minusvalía), costes económicos, actividad productiva, satisfacción de los pacientes y sus familiares, destino al alta, etc. (42-44, 46, 51). En la evaluación crítica de los diferentes artículos se debe comprobar que se han escogido las medidas relevantes de acuerdo con el objetivo del estudio y, sobre todo, que dichas medidas tengan una aceptable fiabilidad y validez (46, 51, 52).

Como *conclusión*, la evaluación crítica de la evidencia científica es un paso de un gran valor en la práctica de la MBE, para lo que se necesitan unos conocimientos básicos en epidemiología clínica y estadística. La MBE ha propuesto un método simple y estructurado de evaluación, pero que no incluye muchos de los diseños que existen para conocer la eficacia y efectividad del tratamiento rehabilitador. Por último, en los artículos de rehabilitación es muy importante un uso correcto de las medidas de resultados, como son las escalas de valoración.

## BIBLIOGRAFÍA

- Rosenberg W, Donald A. Evidence based medicine: an approach to clinical problem-solving. *BMJ* 1995;310:1122-6.
- Sackett DL, Richardson WS, Rosenberg W, Haynes RB. Medicina basada en la evidencia. Cómo ejercer y enseñar la MBE. Madrid: Churchill Livingstone; 1997.
- Pozo Rodríguez F. La medicina basada en la evidencia. Una perspectiva desde la clínica. *Med Clin (Barc)* 1999; 112(Supl 1):12-6.
- Vandenbroucke JP. Observational research and evidence-based medicine: what should we teach young physicians? *J Clin Epidemiol* 1998;51:467-72.
- Davies HTO. Introducing critical appraisal. *Hosp Med* 2000;61:432-3.
- Jbvel A, Navarro-Rubio MD. Evaluación de la evidencia científica. *Med Clin (Barc)* 1995;105:740-3.
- Guyatt GH, Haynes RB, Jäschke RZ, Cook DJ, Green L, Naylor CD, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature. XXV. Evidence-based medicine: Principles for applying the users' guides to patient care. *JAMA* 2000;284:1290-6.
- Sackett DL, Wennberg JE. Choosing the best research design for each question. *BMJ* 1997;315:1636.
- Barton S. Which clinical studies provide the best evidence? *BMJ* 2000;321:255-6.
- Oxman AD, Sackett DL, Guyatt GH for the Evidence-Based Medicine Working Group. Users' guide to the medical literature. I. How to get started. *JAMA* 1993; 270:2093-5.
- Gol-Freixa JM. Una perspectiva global sobre la medicina basada en la evidencia. *Enferm Infecc Microbiol Clin* 1999;17(Supl 2):3-8.
- Guyatt GH, Sackett DL, Cook DJ for the Evidence-Based Medicine Working Group. Users' guides to the medical literature. II. How to use an article about therapy or prevention. A. Are the results of the study valid? *JAMA* 1993;270:2598-601.
- Guyatt GH, Sackett DL, Cook DJ for the Evidence-Based Medicine Working Group. Users' guides to the medical literature. II. How to use an article about therapy or prevention. B. What were the results and will they help me in caring for my patients? *JAMA* 1994;271:59-63.
- Jäschke R, Guyatt GH, Sackett DL, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. Are the results of the study valid? *JAMA* 1994;271:389-91.
- Jäschke R, Guyatt GH, Sackett DL, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? *JAMA* 1994;271:703-7.
- Laupacis A, Wells G, Richardson S, Tugwell, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature. V. How to use an article about prognosis. *JAMA* 1994;272:234-7.
- Oxman AD, Cook DJ, Guyatt GH, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature. VI. How to use an overview. *JAMA* 1994;272:1367-71.
- Hayward R, Wilson MC, Tunis SR, Bass EB, Guyatt GH, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature. VIII. How to use clinical practice guidelines. A. Are the recommendations valid? *JAMA* 1995;274:570-4.
- Wilson MC, Hayward RS, Tunis SR, Bass EB, Guyatt GH, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature. VIII. How to use clinical practice guidelines. B. What are the results and will they help me in caring for my patients? *JAMA* 1995;274:1630-2.
- Guyatt G, Jäschke R, Heddle N, Cook D, Shannon H, Walter S. Basic statistics for clinicians: 1. Hypothesis testing. *CMAJ* 1995;152:27-32.
- Guyatt G, Jäschke R, Heddle N, Cook D, Shannon H, Walter S. Basic statistics for clinicians: 2. Interpreting study results: confidence intervals. *CMAJ* 1995;152:169-73.
- Jäschke R, Guyatt G, Shannon H, Walter S, Cook D, Heddle N. Basic statistics for clinicians: 3. Assessing the effects of treatment: measures of association. *CMAJ* 1995;152:351-7.
- Guyatt G, Walter S, Shannon H, Cook D, Jäschke R, Heddle N. Basic statistics for clinicians: 4. Correlation and regression. *CMAJ* 1995;152:497-504.
- Greenhalgh T. How to read a paper. The Medline database. *BMJ* 1997;315:180-3.
- Greenhalgh T. How to read a paper. Getting your bearings (deciding what the paper is about). *BMJ* 1997;315:243-6.
- Greenhalgh T. How to read a paper. Assessing the methodological quality of published papers. *BMJ* 1997;315: 305-8.

27. Greenhalgh T. How to read a paper. Statistics for the non-statistician. I: Different types of data need different statistical tests. *BMJ* 1997;315:364-6.
28. Greenhalgh T. How to read a paper. Statistics for the non-statistician. II: «Significant» relations and their pitfalls. *BMJ* 1997;315:422-5.
29. Greenhalgh T. How to read a paper. Papers that report drug trials. *BMJ* 1997;315:480-3.
30. Greenhalgh T. How to read a paper. Papers that report diagnostic or screening tests. *BMJ* 1997;315:540-3.
31. Greenhalgh T. How to read a paper. Papers that tell you things cost (economic analyses). *BMJ* 1997;315:596-9.
32. Greenhalgh T. How to read a paper. Papers that summarise other papers (systematic reviews and meta-analyses). *BMJ* 1997;315:672-5.
33. Greenhalgh T. How to read a paper. Papers that go beyond numbers (qualitative research). *BMJ* 1997;315:740-3.
34. Fowkes FGR, Fulton PM. Critical appraisal of published research: introductory guidelines. *BMJ* 1991;302:1136-40.
35. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA* 1996;276:637-9.
36. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF, for the QUORUM group. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUORUM statement. *Lancet* 1999;354:1896-900.
37. Sauv   S, Lee HN, Meade MO, Lang JD, Farkouh M, Cook DJ et al. The critically appraised topic: a practical approach to learning critical appraisal. *Ann Roy Coll Phys Surg Canada* 1995;28:396-8.
38. Norman GR, Shannon SI. Effectiveness of instruction in critical appraisal (evidence-based medicine) skills: a critical appraisal. *CMAJ* 1998;158:177-81.
39. Green ML. Graduate medical education training in clinical epidemiology, critical appraisal, and evidence-based medicine: A critical review of curricula. *Acad Med* 1999;74:686-94.
40. Taylor R, Reeves B, Ewings P, Binns S, Keast J, Mears R. A systematic review of the effectiveness of critical appraisal skills training for clinicians. *Med Educ* 2000;34:120-5.
41. Neville AJ, Reiter HI, Eva KW, Norman GR. Critical appraisal turkey shoot: Linking critical appraisal to clinical decision making. *Acad Med* 2000;75(Supl):S87-9.
42. Melvin JL. Outcomes research in rehabilitation. *Am J Phys Med Rehabil* 2001;30:78-82.
43. Tate DG, Findley T, Dijkers M, Nobunaga AI, Karunas RB. Randomized clinical trials in medical rehabilitation research. *Am J Phys Med Rehabil* 1999;78:486-99.
44. Johnston MV, Granger CV. Outcomes research in medical rehabilitation. A primer and introduction to a series. *Am J Phys Med Rehabil* 1994;73:296-303.
45. Johnston MV, Ottenbacher KJ, Reichardt CS. Strong quasi-experimental designs for research on the effectiveness of rehabilitation. *Am J Phys Med Rehabil* 1995;74:383-92.
46. Keith RA. Conceptual basis of outcome measures. *Am J Phys Med Rehabil* 1995;74:73-80.
47. Peters DJ. Qualitative inquiry. Expanding rehabilitation medicine's research repertoire. A commentary. *Am J Phys Med Rehabil* 1996;75:144-8.
48. Tooth L. Use of sequential medical trials in rehabilitation research. *Am J Phys Med Rehabil* 1999;78:87-97.
49. Backman CL, Harris SR. Case studies, single-subject research, and N of 1 randomized trials. *Am J Phys Med Rehabil* 1999;78:170-6.
50. Andresen EM, Lollar DJ, Meyers AR. Disability outcomes research: Why this supplement, on this topic, at this time? *Arch Phys Med Rehabil* 2000;81(Supl 2):S1-4.
51. Hinderer SR, Hinderer KA. Principles and applications of measurement methods. En: DeLisa JA, Gans BM eds. *Rehabilitation Medicine: Principles and Practice*. 3<sup>th</sup> ed. Filadelfia: Lippincott-Raven Publishers; 1998. p. 109-36.
52. Katz RT, Campagnolo DI, Goldberg G, Parker JC, Pine ZM, Whyte J. Critical evaluation of clinical research. *Arch Phys Med Rehabil* 1995;76:82-93.

*Correspondencia:*

Alejandro Arteaga Dom  nguez  
Servicio de Rehabilitaci  n  
Hospital de Cabue  es  
33394 Cabue  es - Gij  n (Asturias)  
E-mail: alecar@teleline.es