

La revista científica en medicina (X de XIV)

4.2 La evaluación de la calidad de un artículo de investigación

DRA. MA. CRISTINA SOSA DE MARTÍNEZ @*, M. EN C. JOSÉ LUIS PABLOS HACH*,
MA. CRISTINA MARTÍNEZ SOSA*

RESUMEN

El décimo artículo de la serie tiene como finalidad describir cómo se evalúa la calidad de un artículo de investigación. Dicha evaluación puede separarse en dos componentes: 1. La calidad de la investigación en sí y 2. La calidad del artículo de investigación. En este artículo nos limitaremos al segundo componente, puesto que la evaluación del primero corresponde solamente a sus «pares», quienes serán quienes juzguen el valor científico y la originalidad de la investigación realizada. A grandes rasgos se presentan algunas de las propuestas teóricas más relevantes para evaluar los artículos sobre investigaciones clínicas, así como algunos estudios empíricos en donde se aplican dichas propuestas.

Palabras clave: Evaluación de la calidad, artículo de investigación, Normas de Vancouver.

ABSTRACT

The purpose of the tenth article of the series is to describe how the assessment of the quality of a research article is performed; the evaluation can be divided in its two components. 1. The quality of the research *per se* and 2. The quality of the research report. In this article we will limit our presentation to the second component, due to the fact that the evaluation of the first one should be performed only by its peers, who will judge the scientific value and originality of the research some of the most relevant theoretical quality assessment tools for clinical research articles are described, as well as some empirical studies performed on their application.

Key words: Quality evaluation, research articles, Vancouver Norms

4.2 Evaluación de la calidad de un artículo de investigación clínica

Si bien la calidad de un artículo científico es algo intangible, existen ciertos factores que permiten valorarla, entre los que se encuentran la cantidad de quienes leen el artículo y lo consideran valioso, la cantidad de cartas recibidas por la revista al respecto, la atención que recibe de parte de los medios de comunicación, la cantidad de veces que es citado, los premios que recibe, etc. y finalmente, la forma en que es capaz de modificar una conducta clínica.¹⁻³ Asimismo, parti-

cipa el prestigio de la revista que lo publica y su selección como “artículo notable”; etc.⁴

La *calidad de un manuscrito* debe separarse conceptualmente en dos componentes: 1. La calidad de la investigación en sí misma y 2. La calidad del informe de investigación. Esta última se presenta cuando “los autores han descrito su investigación con suficiente detalle y con suficiente claridad para que un lector pueda realizar un juicio independiente de la solidez y debilidades de sus datos y sus conclusiones”.⁵

Se han publicado diversos instrumentos para evaluar la calidad de las investigaciones, de sus informes o de ambos, que solamente permiten evaluar un tipo específico de diseño, o bien son tan generales que evalúan no solamente diferentes tipos de estudios, sino áreas fuera de la práctica clínica.⁶

Badgley,⁷ diseña un instrumento con preguntas abiertas que mediante cinco criterios, evalúa estudios clínicos. Selecciona para su trabajo, 103 estudios clínicos, incluyendo ensayos clínicos controlados, estudios

* División de Investigación, Instituto Nacional de Pediatría
** Revista Ingeniería, Investigación y Tecnología, Universidad Nacional Autónoma de México
@ Beca Conacyt No 86532. Maestría en Metodología de la Ciencia. PESTyC. Instituto Politécnico Nacional
Correspondencia: Dra. Ma. Cristina Martínez. Departamento de Metodología de Investigación. Instituto Nacional de Pediatría Col. Insurgentes Cuicuilco 04531 México, D.F. correo electrónico: mcmtz@servidor.unam.mx
Recibido: mayo, 2000. Aceptado: junio, 2000.

epidemiológicos y descripciones de procedimientos terapéuticos, publicados en las revistas *Can Med Assoc J* y *Can J Public Health*. Los criterios utilizados son: 1. presencia de una *definición operacional* de la entidad en estudio, de tal manera que sea posible replicarlo; 2. *muestreo realizado*, para conocer qué criterios se utilizaron para definir a la población en estudio, tomando en cuenta unidad de análisis y procedimiento seguido en la selección de la muestra para determinar su representatividad; 3. uso de grupo *control* cuando se requiere y la idoneidad del mismo; 4. *tipo de análisis estadístico realizado* y 5. *aplicabilidad de las conclusiones*. El 41% de los artículos evaluados por el autor presentan uno o más problemas en los criterios mencionados, lo que impide utilizar los resultados de dichos estudios.

Cabe señalar que el enfoque del citado autor, más que metodológico, es de tipo estadístico. Por otra parte, aunque considera aquellos casos en donde no hay congruencia entre los totales de las columnas y la suma de sus elementos, carece de una categoría para investigaciones que no requieren análisis estadístico. Tampoco describe con suficiente detalle cómo aplica los cinco criterios, lo que impide reproducir el procedimiento, pese a que cuando habla de las conclusiones, señala que como este aspecto es la piedra de toque para modificar la conducta clínica, cualquier problema en esta sección es motivo de que los hallazgos y conclusiones de una investigación deben volverse a examinar para buscar otras posibles interpretaciones alternativas.

Mahon y cols.,⁸ evalúan la validez de los informes de 203 ensayos clínicos controlados sobre fármacos, publicados durante cinco años en *Can Med Assoc J*. Pese a que no definen qué entienden por estudio sobre fármacos, establecen cuatro criterios: 1. La presencia de controles adecuados; 2. La aleatorización del tratamiento; 3. La evaluación objetiva de los efectos del fármaco y 4. El análisis estadístico de los resultados. El autor halló que: en 133 artículos no hay controles; en 49, hay controles válidos, pero el tratamiento no se asigna aleatoriamente; en seis, hay controles válidos y asignación aleatoria del tratamiento, pero la evaluación del mismo no es ciega; en cuatro, se cumplen los tres criterios mencionados, pero no tienen análisis estadístico y solamente en once se cumplen todos los criterios.

Lionel y Herxheimer,⁹ presentan un listado para "la evaluación sistemática" de los ensayos clínicos controlados sobre terapéutica; evalúan 141 informes publicados durante los primeros seis meses entre 1966 y 1968, en cuatro revistas británicas no especializadas, las cuales no identifican. Dicho estudio pretende seguir los lineamientos generales del estudio de Mahon y cols.,⁸ respecto a la descripción de los sujetos en estudio, administración de los fármacos, métodos, diseño y evaluación general del informe. Tiene el inconveniente de que sus preguntas no siguen la progresión lógica que se espera en un artículo de investigación, éstas son ambiguas y no es posible determinar qué respuesta se espera en ellas. Tampoco se especifica cómo se aplica la evaluación final para determinar la validez del estudio publicado en el artículo. Los hallazgos de los autores son: 51% de los informes son aceptables; un 16%, probablemente lo sean y el resto son inaceptables porque no se menciona uno o más de los aspectos evaluados.

Mosteller,¹⁰ señala en forma no cuantitativa las siguientes seis omisiones de información como los hallazgos más relevantes detectados por Gilbert y cols.,¹¹ al revisar 107 ensayos clínicos controlados en cirugía y anestesia publicados entre 1964 y 1972 en revistas que no identifica: 1. Aleatorización; 2. Consentimiento informado; 3. Técnicas de análisis; 4. Estadísticas descriptivas; 5. Potencia de la prueba usada y 6. Selectividad y multiplicidad de técnicas estadísticas empleadas buscando la significancia estadística.

DerSimonian y cols.,¹² parten del supuesto de que un ensayo clínico controlado no puede ser interpretado adecuadamente sin información sobre los métodos utilizados para su diseño y el análisis de sus resultados. Para ello investigan la presencia de once factores en: 13 informes de la revista *N Engl J Med*; 21 en *Lancet*; 19 en *Br Med J*, todos publicados entre julio y diciembre de 1979, así como 14 informes publicados en *JAMA*, entre julio de 1979 y junio de 1980. Los aspectos evaluados y sus hallazgos (entre paréntesis), se presentan a continuación: 1. criterios de selección de los sujetos en estudio (37%); 2. admisión del sujeto al estudio antes de ser asignado al tratamiento (67%); 3. asignación aleatoria del tratamiento (80%); 4. método para calidad aleatorización (19%); 5. el paciente desconoce qué tratamiento recibe (55%); 6. quien evalúa el resultado des-

conoce el tratamiento recibido (30%); 7. complicaciones del tratamiento (64%); 8. pérdidas de sujetos durante el seguimiento, incluyendo las razones de las mismas (79%); 9. los análisis estadísticos van más allá de las técnicas descriptivas (80%); 10. técnicas de estadística inferencial realizadas, incluyendo nombre de las pruebas y programas de cómputo utilizados (67%) y 11. la potencia de la prueba (12%).

Cuando comparan el porcentaje de respuestas positivas en función de la revista en la que se publica el artículo, detectan diferencias significativas: *N Engl J Med* informa 71% de los 11 aspectos; *JAMA* 63%; *Br Med J* 52%, y *Lancet* 46%. Cabe señalar que los autores introducen un sesgo en la interpretación de los resultados al considerar como presente el aspecto investigado, cuando se señala que intencionadamente se omite su realización; por ejemplo, la aleatorización y cuando no se aplica en el estudio. Dicho sesgo hubiera podido evitarse mediante el concepto de "valor faltante", consistente en excluir esos casos del análisis de dicho aspecto.¹³

Tyson y cols.,¹⁴ evalúan la calidad de los estudios terapéuticos en medicina perinatal publicados en *Am J Obstet Gynecol*, *Obstet Gynecol*, *J Pediatr* y *Pediatrics*. El estudio comprende varias etapas. La primera consiste en considerar catorce categorías, entre las que se encuentran: objetivo del estudio, métodos de colección de información, determinación del tamaño de muestra, elegibilidad de los pacientes, uso de controles, enmascaramiento, fuentes de sesgo, pertinencia de la información clínica, evaluación de riesgos, análisis y presentación de la información. En cada categoría se evalúan múltiples criterios, para hacer un total de 72 preguntas que esperan respuestas tales como "sí", "no" "no es claro o se desconoce" y "no aplicable". Acto seguido determinan que 29 criterios son de tipo primario, es decir, esenciales para un ensayo clínico bien diseñado.

A continuación, conforman varias parejas de evaluadores, integradas por un neonatólogo y un experto en bioestadística. Después de que el artículo es "discutido" por cada pareja, cada miembro le asigna dos calificaciones: una objetiva y otra subjetiva; la primera es el porcentaje de criterios primarios cumplidos por el estudio después de descartar los que no son aplicables;

la segunda, es la calificación otorgada a la calidad del estudio con valores entre 0 y 2 (máxima calificación), en cada categoría mencionada. Llama la atención que investiguen aspectos tales como la magnitud especificada del efecto del tratamiento o del manejo investigado y no definan qué entienden por: descripción adecuada de los sujetos, ni por métodos de evaluación apropiados para responder preguntas de investigación, lo que impide reproducir su estudio. En el citado estudio encuentran que el rango de la calificación objetiva es de 11 a 93%, con un promedio de 46% (DE: 18%) de los neonatólogos y de 45% (DE 18%) de los bioestadísticos, mientras que la variación de las calificaciones subjetivas oscila entre 7 y 89% con un promedio de 46% (DE: 18%) de los neonatólogos y de 45% (DE 16%) de los expertos en bioestadística.

Partiendo del supuesto de que la mejor evaluación del proceso de selección de manuscritos en las revistas más importantes se puede realizar después de transcurrido un lapso, Dixon y cols.,¹⁵ seleccionan artículos publicados sobre neumología seis años antes por *Am J Med* (nueve artículos), *Ann Intern Med* (14 artículos) y *N Engl J Med* (11 artículos). A continuación, solicitan por cada artículo dos investigadores neumólogos que han recibido el Premio de los Institutos Nacionales de Salud de Estados Unidos, que califiquen si: 1. el artículo ha tenido impacto, es decir, si ha logrado mejorar el diagnóstico, el tratamiento o se ha traducido en una mejor comprensión de la patogénesis o de la patofisiología de la enfermedad; 2. no ha tenido impacto y 3. se refuta posteriormente. Cabe señalar que se les solicita expresamente no evaluar la calidad del informe de investigación.

Ambos evaluadores concuerdan en la evaluación de 25 artículos: 17 considerados relevantes clínicamente y ocho irrelevantes. Respecto a los 14 artículos en que hay discrepancia, los autores solicitan que sea el comité editorial de la revista donde se publica el artículo el que diga la última palabra: ocho son relevantes y el resto, no. No se detectaron diferencias significativas entre las revistas investigadas con respecto a la relevancia de los artículos que publican. Entre las limitaciones del estudio mencionadas por los autores, se encuentra el hecho de que aunque su objetivo es evaluar el proceso de selección de los manuscritos, no pueden determinar

cuántos artículos de calidad son rechazados durante dicho proceso.

M. Evans y A.V. Pollock,¹⁶ analizan el diseño, métodos de análisis, exactitud y claridad de presentación de 45 ensayos clínicos controlados sobre profilaxis de infección en cirugía abdominal, que localizan mediante *Index Medicus* durante 1980 y 1981. En primer lugar, observan que en seis artículos no se define qué se entiende por infección de herida. Consideran como problemas de diseño los siguientes: en los aspectos éticos, solamente en cinco artículos se menciona el consentimiento informado; en 19 artículos, a los pacientes del grupo control de cirugía de colon y recto por peritonitis bacteriana, no se les administra antibióticos y en otros tres, se les administran inyecciones intramusculares con placebo.

También encuentran que muy pocos autores distinguen entre los pacientes que son excluidos y los que se retiran del estudio. En 22 artículos no se señalan las razones de las exclusiones, mientras que en doce no indican la cantidad de sujetos que se retiran del estudio. Encuentran métodos inadecuados de aleatorización en cuatro. En 22 artículos, no se indica el tiempo de seguimiento. Detectan defectos en el área de estadística en 11 de 25 estudios. Por ejemplo, en nueve de 20 artículos, no se señala qué prueba se utiliza; en dos se viola el postulado de la prueba de Ji-Cuadrada que habla sobre el valor mínimo esperado, etc. También hay defectos en la presentación en 40 artículos. En este último rubro encuentran que siete de los títulos no reflejan el contenido del artículo; en el Resumen se omite información importante; en cuatro artículos, la Introducción es inadecuada; en 28, los Métodos no son reproducibles; en 29 artículos, no se señala la comparabilidad inicial entre los grupos, o bien, no se estratifican las cirugías abdominales en función de los riesgos de la infección de la herida; se presentan conclusiones no justificadas por los datos, en nueve artículos. No se menciona la comparabilidad entre los grupos en 29 y en once artículos las conclusiones no están sustentadas por el estudio.

M. Thorn y cols.,¹⁷ consideran que se puede valorar la calidad de la investigación, cuando la hipótesis, los métodos, resultados y conclusiones siguen una progresión lógica. Proponen cuatro categorías basadas en: 1.

presencia de una progresión lógica, ya que la conclusión es el resultado tanto de la hipótesis, como del análisis estadístico de los resultados, así como que los métodos estadísticos son correctos o aceptables; 2. *ausencia de progresión lógica* entre la hipótesis y los métodos de análisis estadístico de los resultados, o bien el análisis se consideró como incompleto o incorrecto; 3. *carencia de análisis estadístico*; 4. *no evaluable debido a información insuficiente, errores en la documentación, o ambos*.

Para conocer cómo evalúan la calidad científica de un artículo de investigación sobre un problema clínico de actualidad, Williamson y col.,¹⁸ aplican una encuesta, a un grupo médicos generales y de diversas especialidades y a líderes de opinión. Encuentran que en primer lugar, el 87% de los médicos y 90% de los líderes de opinión lo comparan con su propia experiencia; en segundo, que se basan en el aval que creen que confiere la revisión por pares que realiza la revista.

Para evaluar la calidad del informe de investigación, S. Goodman y cols.,⁵ diseñan un instrumento compuesto por 34 reactivos que provienen de listas publicadas en la literatura, de artículos sobre el contenido que deben tener los artículos de revistas, la experiencia editorial de los autores, así como de los comentarios de editores y metodólogos. Cada pregunta se puede calificar entre uno (calificación mínima) y cinco, a fin de que el evaluador pueda graduar la gravedad de la omisión. Organizan el instrumento siguiendo la estructura de un artículo: Título y Resumen (2 preguntas); Introducción (2 preguntas), Métodos (7 preguntas), Resultados (15 preguntas), Discusión y Conclusiones (4 preguntas) y Evaluación General (4 preguntas). Pese al esfuerzo realizado, el instrumento tiene entre sus limitaciones el ser demasiado general y no tomar en cuenta el tipo específico de estudio.

DuRant¹⁹ presenta un cuestionario a manera de listado con el fin de «permitir al clínico realizar decisiones certeras sobre la calidad de un artículo de investigación al señalarle qué preguntas debe realizar». Según el autor, su cuestionario está dirigido a evaluar cuatro tipos de diseños, los cuales no define, sino solamente enuncia: experimentales o cuasi experimentales, encuestas o estudios transversales, revisiones de casos retrospectivos y estudios de casos y controles. El cuestionario

tiene una rama común en donde, sin tomar en cuenta el Resumen, se realizan ocho preguntas sobre la Introducción, once sobre los Resultados y nueve sobre la Discusión. Se presentan instrucciones para dirigir al lector a diferentes grupos de preguntas en función de los diseños señalados y en el caso de los dos primeros, a una sección propia para el análisis estadístico. Debido a que algunos de los estudios mencionados comparten planteamientos en su diseño, los autores se ven en la necesidad de repetir ítems en las diferentes secciones.

Cho y Bero,²⁰ desarrollan dos instrumentos, uno para evaluar la calidad metodológica, la cual definen como "la minimización del sesgo sistemático, así como la consistencia entre las conclusiones y los resultados" y la relevancia clínica de estudios sobre fármacos (que incluye aplicabilidad y ética). El problema es que utilizan preguntas abiertas cuya respuesta puede variar de observador a observador, tales como: ¿Se describe suficientemente la pregunta del estudio? Asimismo espera respuestas de tipo: sí, no y parcialmente.

Con un enfoque diferente, Sonis y Jones,²¹ tratan de evaluar la calidad de 53 ensayos clínicos controlados publicados en la revista *J Fam Pract*, entre 1974 y 1991, para posteriormente tratar de predecirla mediante indicadores externos. Dos evaluadores independientes califican el informe, el diseño y análisis de cada estudio mediante el Índice de Calidad de T. Chalmers.²² En promedio, los artículos obtienen 35% de la puntuación. Al asociar los resultados obtenidos con los indicadores externos para tratar de predecir la calidad del informe mediante análisis de regresión lineal, encuentran que hay factores significativos tales como año de publicación, cantidad de páginas que ocupa la publicación, así como tipo de intervención (estudio sobre fármaco, educación del médico, del paciente, intervención psicosocial, etc.). En cambio, no encuentran significancia respecto al hecho de que esté involucrado un experto en bioestadística, la afiliación a alguna institución, la cantidad de investigadores participantes o la cantidad de ellos con grados de doctorado o el tamaño de la muestra.

Con respecto al formato en que se ofrece el instrumento de evaluación del contenido de una investigación clínica, también se han realizado diversas propuestas. Una es la presentación a manera de listados (checklists) o cuestionario, sea de tipo cuantitativo^{6,22} o cuali-

tativo.¹²⁻¹⁹ El cuestionario cuantitativo parece tener mayor objetividad y permitir una evaluación global del ensayo. Sin embargo, presenta serios problemas en el momento de la evaluación final, ya que no se puede ser rígido con una escala de medición flexible. Asimismo, debido a que no todos los ítems tienen el mismo valor tendrían que ponderarse las calificaciones que se otorgan a cada uno, a fin de permitir discriminar aquellos que evalúan un aspecto de menor importancia, de los que evalúan aspectos primordiales, como por ejemplo, la aleatorización.

La ausencia de cuantificación en los cuestionarios categóricos se tornan en una ventaja, puesto que obligan a que sea el clínico quien dé el valor que considere pertinente a cada cuestión, y determinar así cuáles son los aspectos problemáticos de la investigación que pretende evaluar. Su inconveniente es que generalmente investigan pocos aspectos y no alcanzan suficiente profundidad en sus cuestionamientos.²³

Desafortunadamente, no ha sido posible hallar una fórmula que convierta la evaluación cualitativa de cada uno de los detalles metodológicos de una investigación clínica, en una evaluación cuantitativa, donde se obtenga un valor que represente su calificación.²⁴ Al respecto, Fowkes,²⁵ señala que las bondades y las deficiencias de un artículo deben juzgarse en función de los objetivos y diseño del estudio. Es decir, no es posible asignar arbitrariamente una puntuación debido a que la mención de algún aspecto metodológico puede tener mayor relevancia que la de otro. Más aún, si hubiera un estudio hipotético que satisficiera todos los requerimientos metodológicos, no necesariamente dirimiría para siempre una cuestión clínica. Cuando mucho, posiblemente contribuyera a incrementar el conocimiento médico.

En el siguiente artículo²⁶ se hará una propuesta para un instrumento de evaluación, con la salvedad de que independientemente de todos los aspectos metodológicos que dicho instrumento permita estudiar en una investigación clínica que se publique en una revista médica, sólo será posible determinar el «continente» del artículo. Respecto al «contenido», se reserva para los expertos en el tema, es decir, sus «pares», quienes podrán juzgar el valor científico y la originalidad de una publicación.

REFERENCIAS BIBLIOGRÁFICAS

1. Lundberg GD, Williams E. The quality of a medical article. Thank you to our 1990 peer reviewers. *JAMA* 1991;265(9):1161-2.
2. Garfield E. Which medical journals have the greatest impact? *Ann Intern Med* 1986;105:313-20.
3. Montague T, Taylor L, Martin S, Barnes M, Ackman M et al. The Clinical Quality Improvement Network (CQIN) Investigators. Can practice patterns and outcomes be successfully altered? Examples from cardiovascular medicine. *Can J Cardiol* 1995;11(6):487-92.
4. Lundberg GD. The functions of refereed scientific journals. In: Weeks RA, Kinser DL editores. *Editing the refereed scientific Journal*. Nueva York: IEEE Press. 1994:1-6.
5. Goodman SN, Berlin J, Fletcher SW, Fletcher RH. Manuscript quality before and after peer review and editing at *Annals of Internal Medicine*. *Ann Intern Med* 1994;121:11-21.
6. Langer A, Bobadilla JL. Evaluación hemerográfica de la investigación en salud perinatal realizada en México, 1980-1988. *Bol Med Hosp Inf Mex* 1990; 47:160-7.
7. Badgley RF. An assessment of research methods reported in 103 scientific articles from two Canadian medical journals. *Can Med Assoc J* 1961;85:246-50.
8. Mahon WA, Daniel EE. A method for the assessment of reports of drug trials. *Can Med Ass J* 1964;90:565-9.
9. Lionel ND, Herxheimer A. Assessing reports of therapeutic trials. *Br Med J* 1970;3:637-40.
10. Mosteller F. Problems of omission in communications. *Clin Pharm Ther* 1979;25(2):761-4.
11. Gilbert J, McPeck B, Mosteller F. Progress in surgery and anesthesia: Benefits and risks of innovative therapy. In: Bunker J, Barnes B, Mosteller F. editores. *Cost, risks and benefits of surgery*. Nueva York: Oxford University Press 1977:124-69. En Mosteller F. Problems of omission in communications. *Clin Pharm Ther* 1979;25(2):761-4.
12. DerSimonian R, Charette J, McPeck B, Mosteller F. Reporting on methods in clinical trials. *N Engl J Med* 1982;306(22):1332-7.
13. Sosa-de-Martínez MC, Pablos-Hach JL, Santos-Atherton D. Guía para elaborar el protocolo de investigación. Parte 6. Material y métodos. Hoja de captación de información. *Acta Ped Mex* 1996;17:313-32.
14. Tyson JE, Furzan JA, Reisch JS, Mize SG. An evaluation of the quality of therapeutic studies in perinatal medicine. *J Pediatr* 1983;102(1):10-3.
15. Dixon GF, Schonfeld SA, Altman M, Whitcomb ME. The peer review and editorial process: a limited evaluation. *Am J Med* 1983;74:494-5.
16. Evans M, Pollock AV. Trials on trial. A review of trials of antibiotic prophylaxis. *Arch Surg* 1984;119:109-13.
17. Thorn MD, Pulliam CC, Symons MJ, Eckel FM. Statistical and research quality of the medical and pharmacy literature. *Am J Hosp Pharm* 1985;42:1077-81.
18. Williamson JW, German PS, Weiss R, Skinner EA, Bowes III F. Health science information management and continuing education of physicians. A survey of U. S. primary care practitioners and their opinion leaders. *Ann Int Med* 1989;110:151-60.
19. DuRant RH. Checklist for the evaluation of research articles. *J Adolesc Health* 1994;15:4-8.
20. Cho MK, Bero LA. Instruments for assessing the quality of drug studies published in the medical literature. *JAMA* 1994;272(2):101-4.
21. Sonis J, Joines J. The quality of clinical trials published in *The Journal of Family Practice*, 1974- 1991. *J Fam Pract* 1994;39(3):225-35.
22. Chalmers TC. A method for assessing the quality of a randomized control trial. *Control Clin Trials* 1981;2:31-49. En Sonis J, Joines J. The quality of clinical trials published in *The Journal of Family Practice*, 1974- 1991. *J Fam Pract* 1994;39(3):225-35.
23. Sacristán JA, Soto J, Galende I. Evaluación crítica de ensayos clínicos. *Med Clin (Barc)* 1993;100(20):780-7.
24. Sosa Elízaga de Martínez MC. Metodología descrita en revistas médicas mexicanas indexadas, publicadas durante 1992-1993 [tesis de maestría]. México, DF: Instituto Politécnico Nacional;1999.
25. Fowkes FG. Critical appraisals of published research: introductory guidelines. *Br Med J* 1991;302:1136-40.
26. Sosa-de-Martínez MC, Pablos-Hach JL, Martínez-Sosa MC. La revista científica en medicina (XI de XIV). 4.3. Propuesta de un instrumento para evaluar la calidad de un artículo de investigación clínica. *Acta Pediatr Mex* 2000;22 (1): en prensa.