

Cómo evaluar el trabajo científico

J.A. Caminero Luna

Sección de Neumología. Hospital Universitario Nuestra Señora del Pino. Las Palmas de Gran Canaria.

La evaluación por expertos de trabajos científicos data de fechas muy remotas. Ya en 1665, con el nacimiento de las primeras revistas científicas, los *Philosophical Transactions of the Royal Society* y el *Journal des Scavans*, comenzaron a enviar artículos para evaluar a los miembros de sus sociedades. Esta práctica se formalizó en 1752, cuando la Royal Society creó su Comité de Artículos, hecho que, lentamente, se institucionalizó entre las revistas de mayor prestigio^{1,2}. Aunque en 1937 estas evaluaciones llegaron a convertirse en un requisito legal para obtener una beca de investigación en el US National Cancer Institute², su uso no se generalizó hasta después de la Segunda Guerra Mundial¹. Sorprende, por lo tanto, que a pesar de ser una práctica tan antigua, revistas de notable prestigio introdujeran esta evaluación tan tardíamente y sometieran a esta parte tan importante de la ciencia a un asombroso descuido y abandono¹.

Un trabajo científico puede ser evaluado en tres fases diferentes: cuando se elabora el proyecto de investigación (para obtener ayudas económicas o intentar superar comités de evaluación), en el momento en que se concluye el estudio (resultados y conclusiones) y cuando debe ser analizada su rentabilidad. En los dos primeros casos la valoración sólo se puede realizar a través de revisores expertos en el tema, que es la metodología utilizada (expertos a pares) de forma ordinaria por las agencias de investigación³. Es más discutible la valoración de la rentabilidad del trabajo de investigación, que puede cuantificarse, bien por beneficios económicos (patentes, balanza comercial de la tecnología desarrollada, etc.), o por múltiples indicadores que tratan de analizar el producto de la investigación, fundamentalmente los derivados de las publicaciones (y las citas que ocasionan), premios, conferencias importantes y fomento de la cooperación internacional^{3,4}. La publicación y los indicadores bibliométricos adquieren un papel relevante en esta tercera fase, aunque hay que destacar que estos

indicadores, a pesar de su uso prolífico en la última década, están en fase temprana de adaptación⁵⁻⁸ y hay que darles su justa importancia.

Por lo tanto, la parte más importante de la evaluación de un trabajo científico recae sobre los revisores o árbitros, incluso en la valoración que efectúan del estudio previo a su publicación⁹. Es por ello que el árbitro es, en la actualidad, la pieza clave en torno a la cual gira el método científico¹⁰⁻¹². Ciencia y arbitraje forman una pareja inseparable, ya que mientras la ciencia es el proceso de hacer avanzar el conocimiento, plantear preguntas y contrastar hipótesis, el arbitraje es el proceso de evaluar el grado en que un estudio particular concuerda con el ideal científico^{10,11}. Sin este arbitraje, nos ahogaríamos en un mar de trabajos sin verificar, prematuros e incompletos^{10,13}. Dada la importancia de estos árbitros, sus funciones, características y modo de actuación adquirirían una relevancia fundamental^{1,10} y por ello deberían, a su vez, ser siempre supervisados por un director o un comité^{1,3}.

La misión principal del árbitro es evaluar detenidamente el trabajo para, de forma documentada y razonada, exponer los errores encontrados o las partes susceptibles de mejorar. Es por ello que un buen revisor debe ser un experto en la materia que analiza, además de tener buena capacidad de juicio y facilidad para enseñar y redactar. Ello incluye ser imparcial en el trabajo examinado, riguroso y sistemático en su análisis y positivo a la hora de exponer las conclusiones. No hay que olvidar que todo estudio es el fruto de un trabajo meticuloso e intenso y refleja un esfuerzo colectivo bien intencionado¹⁰. Al final, la labor de un árbitro siempre se debe reflejar en una mejora del trabajo evaluado, a la que tendrán acceso los autores por los razonados comentarios del revisor^{1,10,14}.

Todo trabajo científico debe ser expuesto con claridad, con lenguaje sencillo y conciso, según unas normas básicas¹⁵⁻¹⁷. Éste es el primer punto que debe conocer todo autor que desee realizar un estudio científico y, por supuesto, la premisa inicial que evaluará el revisor. El título tiene que ser claro, concreto, poco extenso y debe definir el trabajo expuesto. A continuación tiene que aportarse un resumen que sintetice adecuadamente el estudio realizado y que incluya suficiente informa-

Correspondencia: Dr. J.A. Caminero Luna.
Sección de Neumología. Hospital Universitario Nuestra Señora del Pino.
Ángel Guimerá, 93. 35005 Las Palmas de Gran Canaria.

Recibido: 5-9-95; aceptado para su publicación 6-9-95.

Arch Bronconeumol 1996; 32: 55-58



ción sobre la hipótesis del trabajo, el diseño de realización, el material y método y, en su caso, los resultados y conclusiones¹⁶⁻¹⁸. El texto ha de comenzar con una introducción que clarifique y actualice el problema estudiado, así como que formule la hipótesis objeto del trabajo, dejando claro su originalidad^{15,17,18}. Después tiene que abordarse la elección de la estrategia a seguir en la investigación¹⁵, incluyendo la metodología de actuación y el diseño, existiendo unos criterios mínimos de validez^{15,17,19} que tienen que ser rigurosamente perseguidos por el árbitro. Se debe exigir una adecuada planificación operativa de la investigación, analizando la población de estudio, los distintos aspectos de la muestra, la definición de las variables y los métodos de recogida y análisis de datos^{15,19}. Tienen que definirse correctamente los criterios de inclusión y exclusión, lo que sucede y cómo se evalúan los casos que abandonan el estudio y se debe ser muy estricto en la valoración de la muestra seleccionada, recurriendo, si es necesario, a la asesoría de un experto en estadística. Por su parte, el diseño tiene que ser expuesto detalladamente, persiguiendo que pueda contestar a todas las preguntas formuladas en la hipótesis. El método, parte sumamente importante en el desarrollo del trabajo, debe mostrarse con claridad, procurando aportar referencias que ayuden a validarlo. También hay que ser muy rigurosos en perseguir que las variables de estudio estén claramente definidas, así como que el análisis realizado sea adecuado al diseño del trabajo y que no existan errores en el mismo. Por último, con el fin de elevar la calidad científica del trabajo, también se deberían reclamar datos como los posibles intereses económicos en el desarrollo del proyecto^{14,20}, los métodos utilizados para la aleatorización, los dispositivos de salvaguardia para garantizar la integridad del estudio, los procedimientos utilizados de control de calidad, etc.^{14,21}.

La evaluación de los aspectos éticos es igualmente importante, tanto los relacionados con la ejecución de la investigación¹⁵, como los referentes a hechos como la certeza de la autoría del trabajo, la posibilidad de publicación múltiple o fraccionada, la omisión de la parte del estudio que se debe al trabajo de otros y la posible falsificación de datos o plagio^{9,10,17,22-25}. Estas faltas se cometen con más asiduidad que la deseada y con frecuencia escapan a la crítica del árbitro¹.

En la valoración final del trabajo también se debe perseguir una conveniente exposición de resultados, haciendo especial énfasis en los que ayudan a contestar las preguntas planteadas en la hipótesis^{17,26}. La discusión, parte importante en la posible publicación del trabajo, ha de debatir y contestar todas las cuestiones planteadas, realizar una interpretación científica de los resultados y contrastar lo obtenido con lo aportado por otros autores^{17,27}. Por último, debe exigirse una adecuación de la bibliografía empleada^{17,28}.

La investigación que se ha efectuado sobre el proceso de evaluación ha sido escasa y ofrece, de modo sorprendente, resultados decepcionantes acerca de su eficacia. No sólo es caro (alrededor de 200 dólares en revistas como *New England Journal of Medicine* o *British Medical Journal*), sino que se estima que se necesitan unos

9 años de trabajo en el tema para acceder a la categoría de experto o revisor¹. Sin embargo, el tiempo que se calcula que emplea un experto en evaluar un trabajo es relativamente escaso, entre 2-2,8 horas^{1,29,30}, lo que conlleva una sorprendente falta de acuerdo en los resultados, objetivada porque sólo un 41-57% de las parejas de revisores que evalúan un trabajo se ponen de acuerdo acerca de algo tan sencillo como la aceptación o rechazo¹. Todo esto resalta los puntos débiles de este sistema y por ello, la evaluación, a pesar de la idoneidad de la elección de los árbitros, está condicionada por la meticulosidad de la realización de su labor y, por supuesto, por su subjetividad personal (sesgos personales o filosóficos)^{1,20}. A esto hay que sumar los mismos sesgos imputables a los encargados de seleccionar los árbitros en los proyectos de investigación y a los editores en los trabajos enviados para su publicación²⁰. Sin embargo, sigue siendo un sistema eficaz y debe continuar practicándose, aunque es necesario mejorarlo¹.

El producto de toda actividad científica sólo se alcanza cuando los autores comunican su aportación al resto de la comunidad científica. Las presentaciones a congresos, aunque muestran la investigación más reciente, no necesariamente pasan controles estrictos de calidad³¹ y, con frecuencia, los criterios de aceptación se ven influidos por motivos económicos. Además, al igual que las discusiones privadas y los resultados no publicados, no generan una amplia difusión³. Por lo tanto, la conclusión deseada de todo trabajo y la obligación ética del investigador es la publicación, llegando a considerarse como el producto final de la actividad científica propiamente dicha^{3,5,6,16,32}. La labor de los árbitros en la evaluación de los trabajos previa a su publicación (para rechazarlos o para mejorar su contenido) convierte de nuevo a éstos en el eje principal de toda esta actividad, aunque en este caso influyen también decisivamente los editores de las diferentes revistas (criba inicial y decisión final)²⁰. Sin embargo, la actual presión por publicar que tienen todos los profesionales para tratar de conseguir logros en su carrera desvirtúa el valor cuantitativo de éstas. Se puede llegar a asistir al hecho de trabajar en el campo en el que es más sencillo publicar²⁰. Tampoco hay que olvidar que existen buenos trabajos cuyos resultados son negativos y éstos tienen menos probabilidad de publicarse, lo que conlleva un sesgo a la hora de analizar los estudios desarrollados en una línea concreta de investigación^{20,33}, sobre todo si en estos trabajos se realizan metaanálisis.

Según todo lo expuesto, los indicadores de actividad científica giran alrededor de las publicaciones que generan los diferentes trabajos (indicadores bibliométricos) y deben ser manejados en estrecha relación con el resto de marcadores económicos y sociales⁵. Inicialmente estos indicadores fueron rechazados por la comunidad científica, argumentándose que comprometían la individualidad, genialidad y filosofía de los científicos^{5,34,35}. Paulatinamente estos prejuicios se fueron superando^{5,36} y se ha asistido a una difusión y aceptación casi generalizada de los mismos. Bien utilizados, los indicadores bibliométricos ocupan un lugar bien definido en el estudio de la actividad científica⁵⁻⁸. Sin embargo, en los últi-



mos años se ha asistido a un uso indiscriminado de algunos indicadores, sobre todo del denominado *factor de impacto* (FI)³⁷, obtenido mediante el Science Citation Index (SCI).

La actividad científica es siempre multidimensional y no puede estudiarse sólo por un indicador aislado^{5,38,39}. Por lo tanto, tiene que descartarse el realizar evaluaciones de trabajos sólo según estudios bibliométricos, debiendo tomarse éstos solamente como un complemento de la valoración que realizan los expertos^{5,39}. Aportan valiosos datos numéricos relativos a la producción, transmisión y consumo de información de la actividad científica, pero no pueden llegar a considerarse como medidores de calidad de los diferentes trabajos. Por ello, su utilidad es limitada en las evaluaciones⁵. No hay que olvidar que entre un 10^{3,40} y un 46,4%⁴¹ de los trabajos médicos publicados jamás son citados posteriormente⁴¹ y que existe un porcentaje importante de citas erróneas, incluso en revistas de prestigio⁴². Además, a la hora de publicar existen importantes limitaciones idiomáticas, del tipo de investigación realizada, de especialidad e, incluso, de distinto interés regional de cada uno de los estudios. No es lo mismo diseñar un trabajo cuyos resultados interesen que se difundan a nivel nacional o extranjero, o que sea de investigación básica (más cara, pero más citada)³¹ o clínico-epidemiológica o, incluso que se desee que sea conocido preferentemente por la especialidad o por la comunidad internacional.

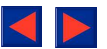
Para la valoración de los artículos y revistas médicas tienen aplicaciones específicas los indicadores bibliométricos de producción, circulación, dispersión, de consumo de información y de repercusión, siendo los más sencillos de este último grupo (número de citas, índice de visibilidad y FI) los de mayor importancia en la evaluación de la actividad médico-científica de autores e instituciones⁸. Según estos últimos parámetros, el Fondo de Investigaciones Sanitarias (FIS) del Ministerio de Sanidad y Consumo evaluó la rentabilidad de los proyectos que becó en 1988 y 1989. En 1988 valoró 270 de los 610 proyectos becados, observando cómo el 35% de éstos no produjeron ninguna publicación (absorbieron el 30% del presupuesto) y otro 21% elaboraron artículos con FI igual o inferior a la unidad. Se obtuvieron 1,74 publicaciones por proyecto becado y un FI medio de 3, lo que suponía, en términos económicos, una cantidad de 1,1 millón gastado por cada artículo publicado y 660.796 pesetas por unidad de FI³. Estos datos, aunque con un coste muy inferior, son cercanos a los que ha obtenido la Sociedad Española de Neumología y Cirugía Torácica (SEPAR) al valorar la rentabilidad de los proyectos de investigación becados en el período 1987-1992. Se han conseguido analizar 57 de los 72 proyectos becados, que han generado 101 publicaciones (49 en revistas extranjeras), lo que supone 1,8 artículos por proyecto y un coste de 502.574 pesetas por publicación. Esta buena rentabilidad obtenida por las becas SEPAR avala el crecer científico de la neumología en nuestro país y debe hacer considerar el incremento de la inversión en este concepto, a pesar de que nuestra Sociedad es una de las que más dinero gasta en ayudas a proyectos de investigación⁴³.

Todas las bases de datos bibliográficos internacionales tienen un sesgo hacia las revistas en inglés^{6,17}, que se hace superlativo en el caso de SCISEARCH, considerada la principal, y de la que se extrae el tan usado FI. Así, el 70% de las revistas que sirven de fuente a esta base de datos pertenecen a Estados Unidos (la mayoría) o Gran Bretaña, cuando la proporción que corresponde al inglés en el periodismo médico mundial no llega al 30%⁶. Este importante sesgo por exceso contrasta con la marcada infravaloración que realiza de las revistas publicadas en ruso, japonés, francés, italiano e, incluso, español. La aportación de nuestras revistas es inferior al 0,5%⁶. Esto da idea de que la validez y fiabilidad del SCI y del FI están limitadas por un problema de cobertura⁴⁴, por lo que el cómputo de citas obtenido por esta base de datos no es el fiel reflejo de la frecuencia con que una publicación es referenciada a nivel internacional^{6,42}.

En nuestro país, cuya producción científica y aportación de publicaciones al resto de las bases de datos son también de segundo nivel (1-2% del total)^{6,25,43-48}, se puede tratar de paliar esta infravaloración usándose el recientemente creado Índice de Citas e Indicadores Bibliométricos de Revistas Españolas de Medicina Interna y sus especialidades^{49,50}, que representa un primer paso hacia la medición de lo que podría denominarse el factor impacto nacional (FIN). Este parámetro se extraería de aquellas revistas y publicaciones que tienen mayor FI en España, para lo que sería necesario utilizar las propias revistas españolas y las extranjeras que más se usan. Sorprende cómo *Medicina Clínica* (3,34) tendría mayor FIN que *American Review Respiratory Disease* (1,57); y cómo ARCHIVOS DE BRONCONEUMOLOGÍA (0,84) superaría a *Chest* (0,73)^{49,51}. Nuestra Revista, que aún no ha podido ser valorada para FI internacional, ocuparía el sexto lugar nacional en cuanto a FIN^{49,50}. Por lo tanto, en España se debería utilizar FIN, al menos en adición al FI internacional^{50,51}.

También hay que tener en cuenta la difusión ideal que tendría que tener el trabajo. Así, se ha llegado a postular razonablemente que, al igual que el FIN, debería existir el FI de especialidad, en el que se mezclasen tanto el FI internacional como el FIN para cada especialidad⁵¹. Es claro que un artículo español sobre cáncer de pulmón tendrá un mayor impacto en nuestro país y en nuestra especialidad si se publica en ARCHIVOS DE BRONCONEUMOLOGÍA que si se hace en *Oncogene* (FI > 5). La realidad es que ni los análisis bibliométricos cuantitativos ni los cualitativos permiten la comparación entre las distintas áreas del conocimiento, por lo que debe darse mayor valor a las comparaciones que se realizan dentro de una misma temática^{3,31}.

Por lo tanto, la evaluación de un trabajo científico es compleja. Es evidente que, a pesar de sus limitaciones, todo debe girar en torno a los revisores. Cuando se produce la publicación, existe la posibilidad de valoración por indicadores bibliométricos, que deben usarse con cautela y conociendo perfectamente lo que pueden aportarnos⁵⁻⁸. Se deben reivindicar los indicadores bibliométricos y, a ser posible, los de especialidad. Sólo compaginando la información aportada por los indica-



dores internacionales, nacionales y de especialidad se podría alcanzar una verdadera valoración del trabajo realizado y de su impacto en la comunidad científica.

BIBLIOGRAFÍA

1. Lock S. La revisión de manuscritos. *Med Clin* 1992; 98: 304-305.
2. Burnham JC. The evolution of editorial peer review. *JAMA* 1990; 263: 1.323-1.329.
3. Ricoy JR, Guasch MF, Jiménez A, Marín E, Medina J, Pozo F. Evaluación del producto de la investigación financiada por el Fondo de Investigación Sanitaria en 1988. *Med Clin* 1992; 99: 690-694.
4. English H, Czerwon JJ. Quantification of performance of research units: a simple mathematical model. *Res Policy* 1990; 19: 477-480.
5. López Piñero JM, Terrada ML. Los indicadores bibliométricos y la evaluación de la actividad médico-científica (I). Usos y abusos de la bibliometría. *Med Clin* 1992; 98: 64-68.
6. López Piñero JM, Terrada ML. Los indicadores bibliométricos y la evaluación de la actividad médico-científica (II). La comunicación científica en las distintas áreas de las ciencias médicas. *Med Clin* 1992; 98: 101-106.
7. López Piñero JM, Terrada ML. Los indicadores bibliométricos y la evaluación de la actividad médico-científica (III). Los indicadores de producción, circulación y dispersión, consumo de la información y repercusión. *Med Clin* 1992; 98: 142-148.
8. López Piñero JM, Terrada ML. Los indicadores bibliométricos y la evaluación de la actividad médico-científica (IV). La aplicación de los indicadores. *Med Clin* 1992; 98: 384-388.
9. Gorstein F. On authors, editors, and publishers. *Human Pathology* 1990; 21: 357-358.
10. Silva A, Campillo Artero C. Cómo se deben evaluar los artículos científicos propuestos para publicación. *Med Clin* 1991; 97: 744-748.
11. Lock S. A difficult balance: editorial peer review in medicine. Filadelfia: ISI Press, 1986.
12. O'Connor M. The scientist as editor: guidelines for editors of books and journals. Nueva York: John Wiley and Sons, 1979.
13. Day RA. Cómo escribir y publicar trabajos científicos. Washington, DC: Organización Panamericana de la Salud, 1990.
14. Spilker B. Cómo elevar la calidad de los ensayos clínicos y su publicación. *Med Clin* 1992; 98: 303-304.
15. Constandriopoulos AP, Champagne F, Podvin L, Denis JL, Boyle P. Preparar un proyecto de investigación. Barcelona: SG Editores, 1991.
16. Day RA. How to write a scientific paper. *Physiotherapy Canada* 1980; 32: 31-36.
17. Lience E. Redacción de un trabajo para una revista biomédica. *Med Clin* 1991; 96: 768-777.
18. Herranz G. La responsabilidad de empezar bien: el resumen y la introducción. *Med Clin* 1986; 86: 205-206.
19. Herranz G. Material y método: cosas básicas dichas en letra pequeña. *Med Clin* 1987; 88: 241-242.
20. Trilla A. Por qué y cómo se aceptan o rechazan artículos para su publicación en las revistas biomédicas. *Med Clin* 1990; 95: 732-734.
21. Chalmers TC. La disminución del sesgo en la publicación de los ensayos clínicos. *Med Clin* 1992; 98: 305-307.
22. Broad W, Wade N. Betrayers of the truth: fraud and deceit in the halls of science. Nueva York: Simon and Schuster, 1982.
23. Barnard H, Overbeke AJ. Duplicate publication of original manuscripts in and from *Nederlands Tijdschrift voor Geneeskunde*. *Ned Tijdschr Geneeskd* 1993; 137: 593-597.
24. Szell K. Major theoretical and practical aspects of publications. *Orv Hetil* 1994; 135: 251-256.
25. Armstrong JD. Plagiarism: what is it, whom does it offend, and how does one deal with it. *AJR Am J Roentgenol* 1993; 161: 479-484.
26. Herranz G. Resultados: el corazón del artículo. *Med Clin* 1988; 90: 500-501.
27. Herranz G. Discusión, o la verdad sin exageraciones. *Med Clin* 1988; 90: 540-541.
28. Herranz G. La bibliografía: más vale poco y bueno que mucho y malo. *Med Clin* 1988; 91: 452-453.
29. Lock S, Smith J. What do peer reviewers do. *JAMA* 1990; 263: 1.341-1.343.
30. Yankauer A. Who are the peer reviewers and how much do they review. *JAMA* 1990; 263: 1.338-1.340.
31. Camí J, Fernández MT, Gómez Caridad I. La producción científica española en biomedicina y salud. Un estudio a través del Science Citation Index (1986-1989). *Med Clin* 1993; 101: 721-731.
32. Price DJS. Toward a model for science indicators. En: Elkana Y, Ledelberg J, Merton RK, editores. *Toward a metric of science. The advent of science indicators*. Nueva York: Wiley, 1978; 69-95.
33. Erill S. La ética de la publicación: el caso de los resultados negativos. *Med Clin* 1992; 98: 308-309.
34. Anónimo. The magic of numbers. *Nature* 1968; 217: 793-794.
35. Anónimo. Can science afford scientist. *Nature* 1970; 226: 10.
36. Price DJS. Smiles at the unobtrusive. *Nature* 1970; 226: 985.
37. Garfield E. Citation analysis as a tool in journal evaluation. *Science* 1972; 178: 471-479.
38. Moravcsik MJ. Life in a multidimensional world. *Scientometrics* 1984; 6: 75-86.
39. Moravcsik MJ. Cómo evaluar la ciencia y a los científicos. *Rev Esp Doc Cient* 1989; 12: 313-325.
40. Peritz BC. The citation impact of letters to the editor: the case of *Lancet*. *Scientometrics* 1991; 20: 121-129.
41. Hamilton DP. Research papers. Who's uncited now. *Science* 1991; 251: 25.
42. George PM, Robbins K. Reference accuracy in the dermatologic literature. *J Am Acad Dermatol* 1994; 31: 61-64.
43. Ruiz Manzano J. Investigación clínica en neumología. *Arch Bronconeumol* 1992; 28: 153-157.
44. Drettner B, Seglen PO, Sivertsen G. The impact factor as an instrument of allocation. Not accepted by the journals in Scandinavia. *Lakartidningen* 1994; 91: 744-745.
45. Terrada ML, López Piñero JM. La producción científica española y su posición en la comunidad internacional. En: López Piñero JM, editor. *Ciencia*. Madrid: Espasa-Calpe 1991; 73-109.
46. Guardiola E. ¿Qué revistas españolas están incluidas en índices biomédicos internacionales? *Med Clin* 1990; 94: 197-198.
47. Excerpta Medica. The Excerpta Medica Abstract Journals. Amsterdam: Elsevier Science Publishers (Esp.), Excerpta Medica, 1988.
48. Perpiñá M. Investigación básica neumológica en España. *Arch Bronconeumol* 1992; 28: 141-144.
49. López Piñero JM, Terrada ML. El consumo de información científica nacional y extranjera en las revistas médicas españolas: un nuevo repertorio destinado a su estudio. *Med Clin* 1994; 102: 104-112.
50. Terrada ML, López Piñero JM, Aleixandre R, Zorrilla V, Mota A, Giménez Sánchez JV. Índice de citas e indicadores bibliométricos de revistas españolas de medicina interna y sus especialidades, 1990. Barcelona: Instituto de Estudios Documentales e Históricos sobre la Ciencia. Barcelona: Ediciones Doyma, 1992.
51. Blasco FJ. Sobre las revistas médicas españolas de especialidad. *Med Clin* 1995; 104: 423-425.