

## Evaluación de la actividad científica a través de indicadores bibliométricos

María Bordons<sup>a</sup> y M.<sup>a</sup> Ángeles Zulueta<sup>b</sup>

<sup>a</sup>Centro de Información y Documentación Científica (CINDOC). Madrid.

<sup>b</sup>Facultad de Documentación. Universidad de Alcalá. Madrid.

*bibliometría / evaluación científica / factor de impacto / indicadores bibliométricos / investigación biomédica / publicaciones biomédicas / referencias bibliográficas*

Los estudios bibliométricos tienen por objeto el tratamiento y análisis cuantitativo de las publicaciones científicas. Forman parte de los «estudios sociales de la ciencia» y entre sus principales aplicaciones se encuentra el área de la política científica. Estos estudios complementan de manera eficaz las opiniones y los juicios emitidos por los expertos de cada área proporcionando herramientas útiles y objetivas en los procesos de evaluación de los resultados de la actividad científica.

Sin embargo, dadas las repercusiones que estas evaluaciones tienen sobre la asignación de fondos para la investigación e, incluso, la promoción profesional de los investigadores, es necesario conocer bien las características de los indicadores bibliométricos y las limitaciones que su uso conlleva.

Una de las bases de datos más utilizada es la del Science Citation Index. En el caso de la investigación biomédica es útil para analizar la producción científica más visible internacionalmente ya que tiene una buena cobertura de revistas biomédicas pero la investigación más clínica y de interés local que se publica en revistas españolas no está recogida en la misma.

Los indicadores bibliométricos más utilizados para medir la actividad científica se basan en los recuentos de las publicaciones y de las citas recibidas por los trabajos publicados, así como en el impacto de las revistas de publicación. El factor de impacto de las revistas es un indicador presente en casi todos los estudios bibliométricos; aunque en ocasiones se equipara al factor de impacto con mayor calidad, este indicador mide específicamente la visibilidad y la difusión de los trabajos publicados en estas revistas más que la calidad científica de los mismos.

**Palabras clave:** *Bibliometría. Biomedicina.*

### EVALUATION OF SCIENTIFIC ACTIVITY THROUGH BIBLIOMETRIC INDICATORS

The scope of bibliometric studies is the treatment and quantitative analysis of scientific publications. They belong to the so-called «social studies of science», and science policy constitutes one of its main applied fields. These studies efficiently complement the opinions and judgements of experts, thus providing objective and useful tools for evaluating the results of scientific activity.

Nevertheless, given the impact that these evaluations have on the assignment of funding for research and even on the professional career of investigators, it becomes essential to know in detail the characteristics of bibliometric indicators and the limitations of their use.

The Science Citation Index database is one of the most employed. In the case of biomedical research it is useful to analyze the most internationally visible scientific production, since it satisfactorily covers biomedical journals; however, clinical research with local interest published in Spanish journals is not included in that database.

Widely employed bibliometric indicators are those measuring scientific activity through the number of publications, those based on the citations received by published studies and, in between them, the impact of journals. The impact factor is an indicator very used in bibliometric studies; though occasionally a high impact factor is assumed to reflect high quality, this indicator specifically measures visibility and diffusion of the works published by these journals rather than their scientific quality.

**Key words:** *Bibliometric. Biomedicine.*

*(Rev Esp Cardiol 1999; 52: 790-800)*

### ¿QUÉ ES LA BIBLIOMETRÍA?

La bibliometría tiene por objeto el tratamiento y estudio de datos cuantitativos procedentes de las publicaciones científicas. Los primeros estudios bibliométricos se remontan a principios de este siglo, y

Correspondencia: M. Bordons.  
Centro de Información y Documentación Científica (CINDOC).  
Consejo Superior de Investigaciones Científicas (CSIC).  
Joaquín Costa, 22. 28002 Madrid.  
Correo electrónico: mbordons@cindoc.csic.es

consistían en recuentos manuales de publicaciones científicas. A partir de los años sesenta, aparece la denominada «ciencia de la ciencia», que nace en la confluencia de la documentación científica, la sociología de la ciencia y la historia social de la ciencia, con el objeto de estudiar la actividad científica como fenómeno social y mediante indicadores y modelos matemáticos. Esta área dará origen a lo que hoy día se conoce como «estudios sociales de la ciencia», campo de carácter claramente interdisciplinario, que se nutre de los recursos técnicos y conceptuales de distintas disciplinas, entre las cuales se encuentra la bibliometría.

Bajo la denominación de «estudios sociales de la ciencia» realizan actualmente su actividad investigadora profesionales de muy diversa formación, algunos de los cuales aplican las técnicas bibliométricas en sus respectivas áreas de actividad. La bibliometría ha experimentado un gran auge y desarrollo facilitado por los avances técnicos y sus variadas áreas de aplicación<sup>1,2</sup>. Los sociólogos de la ciencia profundizan en el estudio de la estructura y dinámica de las áreas científicas mediante distintos indicadores bibliométricos, en especial mediante el uso de los denominados mapas de la ciencia. Por su parte, los historiadores de la ciencia se interesan por realizar un seguimiento de las ideas a lo largo del tiempo. Pero en la actualidad, el campo «estrella» de aplicación de la bibliometría es el área de política científica<sup>3,4</sup>. Su utilidad en esta área explica el auge experimentado por los estudios bibliométricos en las últimas dos décadas, y a su aplicación en el área biomédica nos vamos a referir en este trabajo.

## INDICADORES BIBLIOMÉTRICOS Y EVALUACIÓN DE LA ACTIVIDAD CIENTÍFICA

La importancia de la investigación biomédica para el desarrollo científico de los países y para el bienestar de su población es un hecho incuestionable en la actualidad. Sin embargo, también es una realidad que la investigación conlleva cada vez más costes, por su creciente especialización y complejidad, véase por ejemplo el caso de la investigación en genoma humano, y que los recursos económicos que pueden destinarse a ella son limitados. En este contexto, es fácil entender el creciente interés que muestran los países más desarrollados por los estudios de evaluación de su actividad investigadora tanto en el área biomédica como en las restantes disciplinas.

Los estudios sobre los recursos destinados a la investigación (*estudios de «input»*) permiten conocer el esfuerzo investigador que realizan los países. Estos estudios se basan en datos sobre personal científico, inversiones en investigación y desarrollo (I+D), y otros indicadores definidos en el denominado «Manual de Frascati», punto de referencia metodológico para me-

dir y conceptualizar las actividades científicas y tecnológicas, publicado y revisado periódicamente por la OCDE<sup>5</sup>. Pero cada vez existe un mayor interés por complementar los estudios de «input» con el análisis de los resultados de la investigación (*estudios de «output»*). Aquí es donde tienen cabida los indicadores bibliométricos, que se basan en datos extraídos de las publicaciones científicas asumiendo, que el resultado de la investigación es nuevo conocimiento que se da a conocer a través de publicaciones. Los estudios bibliométricos aportan una interesante visión de la actividad científica del propio país, así como de su situación en el contexto internacional, todo lo cual supone una información básica para facilitar la toma de decisiones de los responsables de la política científica.

Tradicionalmente, la información sobre la situación de la ciencia ha sido siempre proporcionada por los propios científicos. Existe un procedimiento de auto-evaluación dentro de la ciencia por el que la calidad del trabajo de los investigadores es evaluada por el resto de la comunidad científica. Dentro de cada área los expertos revisan la calidad de los manuscritos enviados a las revistas científicas para su publicación (*peer review*). También se forman paneles de expertos que evalúan las solicitudes de ayudas a la investigación de los grupos y la carrera profesional de los investigadores de cara a gratificaciones, promociones o concesión de premios.

El juicio de los expertos ha demostrado ser hasta la fecha el método más apropiado para valorar el grado de desarrollo de un determinado campo de investigación y la calidad de las aportaciones concretas al área de los distintos científicos o grupos. No obstante, el sistema presenta algunas limitaciones, entre las que se puede citar su carácter subjetivo —método muy sensible a la influencia de factores personales—, su elevado coste y su aplicación limitada a pequeñas unidades. Pero, además, las necesidades de la política científica van más allá de las opiniones de los expertos. Los gestores demandan evaluaciones globales de la actividad científica, en todas las áreas, y de la forma más estructurada posible. En este contexto se explica el éxito de los indicadores bibliométricos aplicados a la evaluación científica, ya que se presentan como indicadores objetivos de *output*, que no pueden reemplazar al juicio de expertos, pero que lo complementan con éxito.

En la actualidad, los indicadores bibliométricos o de producción científica se aceptan como un indicador válido de los resultados de la investigación, junto a otros indicadores como patentes o nuevos productos en las áreas más tecnológicas. Como ejemplo del grado de aceptación que han alcanzado se puede mencionar su inclusión en los informes que sobre la situación de la ciencia y la tecnología se emiten periódicamente en los países más desarrollados, donde complementan la información aportada por los indicadores de *input* más tradicionales (recursos humanos, gastos en I+D).

Es el caso de los «Science and Engineering Indicators» de los EE.UU., los «Science & Technologie Indicateurs» publicados por el Observatoire des Sciences et des Techniques (OST) francés, o el «European Report on Science & Technology Indicators» editado por la Unión Europea<sup>6-8</sup>.

## CONSIDERACIONES METODOLÓGICAS

El uso de los indicadores bibliométricos para estudiar la actividad investigadora de un país se basa en la premisa de que las publicaciones científicas son un resultado esencial de dicha actividad. Desde un punto de vista cognitivo, un nuevo conocimiento adquiere valor cuando se da a conocer y difunde dentro de la comunidad científica, porque sólo así podrá contribuir al avance científico. Pero, además, desde un punto de vista social, la publicación de los nuevos descubrimientos es una fase esencial dentro del proceso investigador, ya que permite al científico obtener el reconocimiento por su labor. Por todo ello, la publicación científica se convierte en un resultado importante y tangible de la investigación y los indicadores bibliométricos adquieren validez como medida indirecta de la actividad de la comunidad científica. La fiabilidad de los resultados de los estudios bibliométricos dependerá en gran medida de su correcta aplicación, realizada con conocimiento de sus ventajas, sus limitaciones y sus condiciones óptimas de aplicación, que están ampliamente recogidas en la bibliografía<sup>3,4,9-14</sup>.

### Tipo de área a estudiar

Los indicadores bibliométricos son válidos en aquellas áreas en las que las publicaciones son un resultado esencial de la investigación. Por esta razón, su validez es máxima en el estudio de las áreas básicas, donde predomina la publicación científica, y mucho menor en las áreas más tecnológicas o aplicadas. En las áreas tecnológicas la publicación de los resultados de la investigación es menos relevante, y adquieren especial importancia otros resultados como las patentes o los nuevos productos, por lo que conviene realizar estudios conjuntos de publicaciones y patentes aunque, en ocasiones, ni siquiera las patentes son un buen indicador porque prevalece el secretismo industrial, de cara a proteger el *know-how* y evitar imitaciones.

### Tamaño de la unidad a analizar

El máximo grado de fiabilidad de los estudios bibliométricos se obtiene cuando se analizan grandes unidades, como la producción de un país (macroanálisis) o un área científica (mesoanálisis). Los análisis sobre pequeñas unidades, cuyo caso extremo sería el estudio de la producción de un investigador individual (microanálisis), son menos fiables, ya que muchos de

los indicadores utilizados son de tipo estadístico y su fiabilidad desciende en muestras de pequeño tamaño.

### Las bases de datos como fuentes de información

Las bases de datos bibliográficas son la principal fuente de información que se utiliza en los estudios bibliométricos. En la actualidad existen bases de datos especializadas en todas las áreas científicas, lo que permite analizar cualquier área a través de estas fuentes. Sin embargo, la validez de un estudio dependerá en gran medida de que la base de datos seleccionada cubra de forma adecuada el área objeto de estudio. Las distintas bases de datos difieren en cobertura temática, criterios de selección de revistas y/o documentos, sesgos geográficos y lingüísticos y todas estas características deben estudiarse de forma previa a la realización de un estudio.

## EL SCIENCE CITATION INDEX Y LAS BASES DE DATOS ESPECIALIZADAS

Los análisis bibliométricos de las áreas biomédicas se pueden realizar a través de bases de datos especializadas en medicina (MEDLINE o Excerpta Medica) o en subdisciplinas médicas (p. ej., Cancerlit o AIDS), pero también a través de bases de datos multidisciplinares (p. ej., Science Citation Index [SCI]).

Las principales bases de datos bibliográficas especializadas en medicina son MEDLINE, de productor norteamericano, y Excerpta Medica, que es su homóloga europea. La base de datos MEDLINE, producida por la National Library of Medicine desde 1966 (Bethesda, EE.UU.), recoge cerca de 3.700 revistas biomédicas procedentes de todo el mundo. Entre sus ventajas se puede mencionar su amplia cobertura de revistas y países, y la presencia de descriptores en los documentos que describen su contenido según un tesauro. La utilización de este lenguaje normalizado facilita la delimitación temática de las áreas a estudiar y la identificación de *hot-topics* o temas punteros de investigación. Excerpta Medica, producida por Elsevier Science Publishers (Amsterdam) desde 1974, cubre cerca de 3.600 revistas y tiene una cobertura más amplia de revistas europeas que su homóloga americana.

Los estudios bibliométricos realizados a través de las bases de datos MEDLINE y/o Excerpta Medica permiten estudiar la investigación de mayor difusión internacional, ya que estas bases de datos tienen un amplio uso y difusión. Sin embargo, la actividad de los autores españoles en revistas españolas sólo está parcialmente cubierta en estas bases de datos, ya que recogen sólo una pequeña parte de las revistas biomédicas españolas. Un estudio más exhaustivo de las publicaciones españolas en biomedicina requiere el uso de la base de datos Índice Médico Español (IME), producido desde 1971 por el Instituto de Estudios Docu-

mentales e Históricos sobre la Ciencia de Valencia. Esta base de datos cubre en la actualidad cerca de 120 revistas españolas.

A pesar de la existencia de bases de datos especializadas en medicina, en los estudios bibliométricos se utiliza con frecuencia el SCI, base de datos multidisciplinaria de ciencia y tecnología producida por el Institute for Scientific Information (ISI) de Filadelfia (EE.UU.). El SCI presenta una serie de ventajas que lo hacen especialmente adecuado para su utilización en los estudios bibliométricos:

1. Es multidisciplinario. Cubre cerca de 3.500 revistas de ciencia y tecnología, la mitad de las cuales tienen interés biomédico. La selección de las revistas a cubrir se hace atendiendo a criterios de calidad científica, calidad formal y reconocimiento por parte de la comunidad científica (citas recibidas)<sup>15</sup>.

2. Realiza un «vaciado total» de las revistas seleccionadas, que se recogen en su totalidad. En MEDLINE y Excerpta Medica algunas revistas se revisan y se recogen sólo aquellos artículos considerados de interés («vaciado parcial»).

3. Incluye a todos los autores de los documentos. La base de datos MEDLINE recoge un máximo de 10 autores por documento, de forma que los autores que exceden de esta cifra se incluyen bajo «et al».

4. Para todos los autores recoge el lugar de trabajo, incluyendo el nombre del centro, la ciudad y el país, lo que permite realizar estudios de colaboración. MEDLINE y Excerpta Medica recogen sólo el lugar de trabajo del primer autor, pero esta información no está normalizada, de forma que figura la ciudad pero no siempre el país de procedencia.

5. Contiene las referencias bibliográficas que aparecen en los documentos, dato que no aparece en ninguna otra base de datos y que permite realizar búsqueda de citas.

Todas estas ventajas explican que el SCI sea muy utilizado en los estudios bibliométricos, en ocasiones sin conocer o señalar adecuadamente sus limitaciones. Entre estas últimas se pueden mencionar:

1. Sesgo lingüístico y geográfico, a favor de las revistas en inglés, sobre todo procedentes de los EE.UU. y del Reino Unido. Los distintos países están representados de forma muy desigual en la base de datos, que apenas cubre revistas de los países menos desarrollados. Esto es así porque el ISI pretende ofrecer un panorama representativo de la ciencia internacional, la denominada *main stream science*, por lo que en la selección de una revista se considera su interés internacional y no el local. Pero en realidad, la denominada por el ISI «ciencia internacional» no sólo incluye la investigación básica, por definición de interés internacional, sino también gran parte de la investigación de

interés local que tiene lugar en los países más desarrollados. En lo que se refiere a España, en el año 1997, las únicas revistas españolas recogidas en el SCI eran 4 revistas biomédicas. Aunque la base de datos SciSearch (que corresponde al Current Contents de acceso en línea) tiene mayor cobertura (en 1997 incluía 23 revistas españolas, 15 de ellas biomédicas), su uso en bibliometría queda limitado por los altos costes de su consulta.

2. Sesgo a favor de las áreas básicas, que están mejor representadas que las aplicadas o clínicas. Un estudio sobre la producción española en ciencias de la salud puso de manifiesto la buena cobertura del SCI, comparado con MEDLINE, en las áreas de medicina básica (el SCI recogía el 94% de los documentos previamente identificados en MEDLINE), mientras que su grado de cobertura se situaba en torno al 60% para las áreas de medicina clínica<sup>16</sup>.

## PRINCIPALES INDICADORES BIBLIOMÉTRICOS

Los principales indicadores bibliométricos se pueden agrupar en dos grandes apartados: *a)* indicadores cuantitativos de actividad científica, donde se incluye el número de publicaciones, y *b)* indicadores de impacto, basados en el número de citas que obtienen los trabajos, y que caracterizan la importancia de dicha producción en función del reconocimiento otorgado por otros investigadores. Un estudio completo de la actividad biomédica en su vertiente científica y tecnológica requeriría añadir también un análisis de patentes, ya que las patentes son el principal indicador de actividad tecnológica<sup>17</sup>. Sin embargo, en este trabajo no se aborda el estudio de patentes, ya que son documentos complejos que requieren la consulta de bases de datos específicas y un tratamiento especial de la información.

### Indicadores cuantitativos de actividad científica

El *número de publicaciones* de un centro, área o país es un indicador útil para cuantificar la actividad científica de dichas unidades. Su mayor utilidad se obtiene al efectuar comparaciones con la actividad de otros centros, áreas o países, pues se hace necesario tener un marco de referencia en el que poder ubicar nuestro objeto de estudio. También es interesante realizar seguimientos de la producción científica a lo largo del tiempo. A modo de ejemplo, se puede señalar que durante el período 1986-1989, España ocupaba la séptima posición en la Unión Europea según el número de publicaciones biomédicas recogidas en el SCI. En el siguiente cuatrienio, con un aumento del 70% de documentos, pasó a ocupar el sexto lugar<sup>18,19</sup>. Es claro que el número de publicaciones es un *output* que dependerá en gran medida de los *inputs* introducidos en

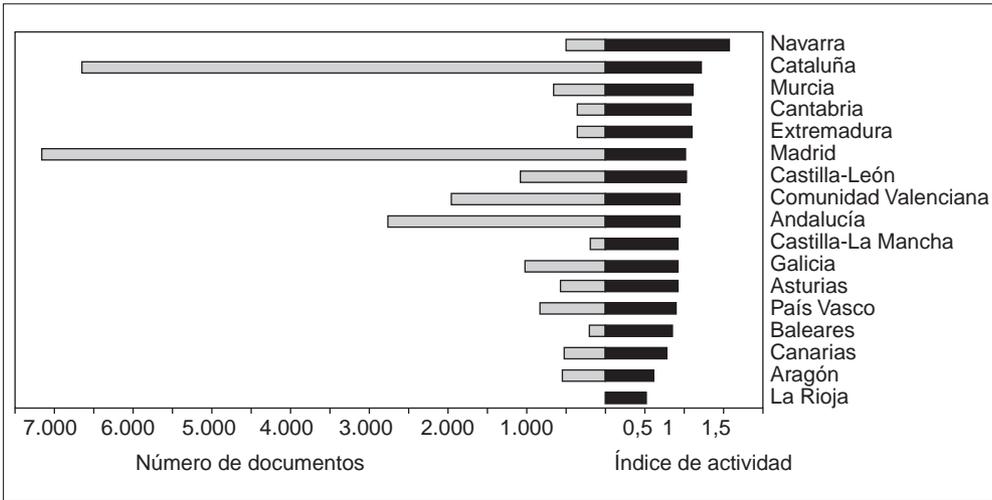


Fig. 1. Distribución de la producción en biomedicina y del índice de actividad por comunidades autónomas. Período 1994-1996. Fuente: Sánchez Nistal et al<sup>20</sup>.

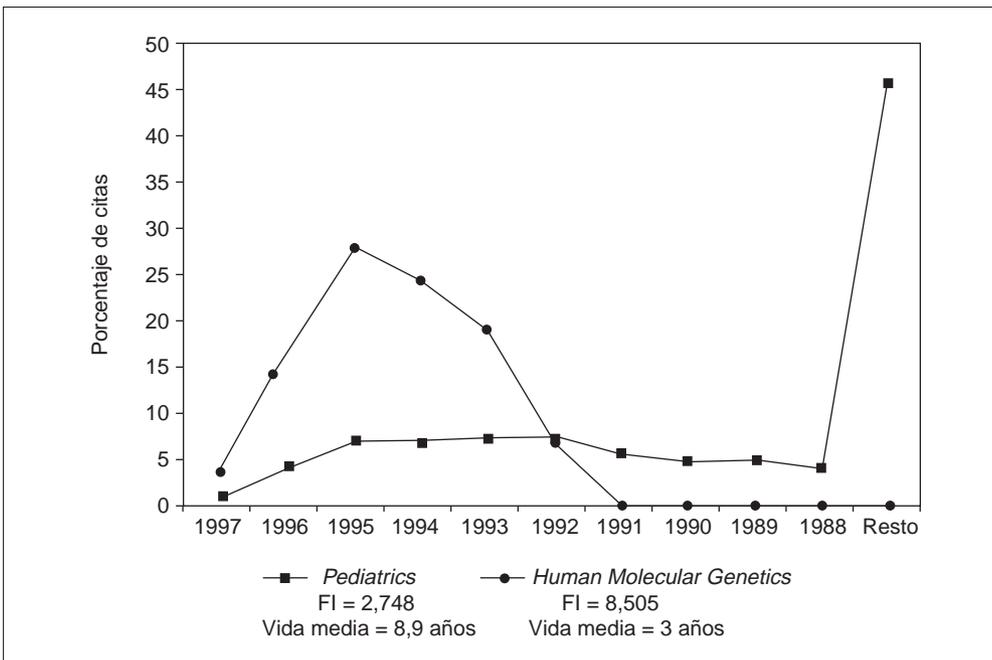


Fig. 2. Citas recibidas en el año 1997 por dos revistas. Distribución por años de los documentos citados; FI: factor de impacto.

el sistema, por lo que es conveniente normalizar este dato en función de los recursos empleados en la investigación. Resulta interesante relativizar la producción en función del personal investigador o de los recursos empleados en I+D pues esto permite identificar centros o regiones que, con un pequeño tamaño desarrollan una importante actividad en el área.

A partir del número de publicaciones se pueden construir otros indicadores, por ejemplo, el *índice de actividad* (IA), que permite comparar el grado de especialización de un centro o una región en un determinado tema. Así, el IA de una comunidad autónoma se calcula como el porcentaje de la producción que dicha comunidad dedica al tema en estudio dividido entre el porcentaje que ese tema representa en la producción nacional, de forma que valores superiores a la unidad

indican mayor actividad en el tema que el promedio nacional y valores inferiores a uno la situación inversa.

En la **figura 1** se expone la especialización de las diferentes regiones españolas en el área biomédica en el período 1994-1996<sup>20</sup>. Destaca Navarra como la región más especializada en el área biomédica que constituye el 73% de su producción, frente al 47% de promedio que dedica el país a la biomedicina. En segundo lugar se encuentra Cataluña que a su alto IA una una alta producción en el área.

### Indicadores de impacto

El número de publicaciones es un indicador meramente cuantitativo, que no valora la calidad o importancia del contenido de los documentos. Es claro que

todas las publicaciones no tienen el mismo interés ni suponen la misma contribución al avance científico. En la búsqueda de indicadores cualitativos, capaces de valorar estos últimos conceptos, se han introducido el *número de citas* que reciben los trabajos y el *factor de impacto* de la revista de publicación. Sin embargo, estos indicadores inicialmente introducidos con el fin de valorar la calidad de los documentos han quedado reducidos después de amplia polémica a indicadores del impacto, influencia o visibilidad de la investigación.

### *Impacto de los trabajos*

El número de citas que recibe un documento es un indicador de la influencia o impacto que produce su contenido sobre la comunidad científica del área. Los avances importantes que abren nuevas líneas de investigación, ya sean teóricos o técnicos, despiertan un gran interés entre los científicos y aparecen con frecuencia referenciados en los trabajos posteriores.

El análisis de citas conlleva el recuento del número de citas que reciben los documentos durante un determinado período de tiempo después de su publicación. Este tipo de estudios nace paralelamente con el SCI, creado en 1963. Todavía en la actualidad, los recuentos de citas se realizan sobre dicha base de datos, que es la única que recoge la bibliografía de los documentos, y que, por tanto, permite estudios de citas.

Es importante destacar que impacto y calidad de la investigación no son términos sinónimos. El término calidad se refiere al contenido científico de la publicación, a lo adecuado de la metodología, a la claridad de exposición y originalidad de planteamientos y conclusiones. El impacto, por su parte, se refiere a la influencia de la publicación sobre la investigación afín en un momento determinado<sup>21</sup>. Aunque nos gustaría contar con algún indicador capaz de valorar la calidad de las publicaciones, de momento tenemos que conformarnos con indicadores indirectos de dicha calidad, como es el impacto que estas publicaciones producen sobre los investigadores del área.

Parece claro que los denominados *hot-papers*, que reúnen las mayores tasas de citación dentro de su área, son documentos relevantes y de alta calidad. Sin embargo, se ha observado que muchas publicaciones que reúnen criterios de calidad atendiendo a la opinión de expertos, apenas reciben citas, es decir, que producen escaso impacto. La razón es que el número de citas recibidas por una publicación depende de su calidad, pero también de otras variables no totalmente identificadas entre las que se incluye el prestigio del autor y/o de su lugar de trabajo, la actualidad del tema, la lengua utilizada y la revista de publicación, que condicionarán la mayor o menor difusión del trabajo<sup>22</sup>.

En definitiva, hoy día se considera que el número de citas que recibe un trabajo es sólo un indicador parcial de su calidad, y lo que muestra es la visibilidad, difu-

sión o impacto del trabajo en la comunidad científica del área.

### *Impacto de las revistas*

La valoración del impacto de los trabajos a través de las citas que reciben no es una medida de tipo inmediato, sino que sólo puede aplicarse varios años después de la publicación de los documentos. Para evitar este problema se introdujo un método alternativo al recuento de las citas que reciben los trabajos consistente en atribuir un peso a las revistas de publicación en función del número medio de citas que reciben sus trabajos. Dado que las revistas que publican más artículos tienen más posibilidades de ser citadas se introdujo el denominado factor de impacto de las revistas, que normaliza el número de citas en función del tamaño de la revista.

El Institute for Scientific Information de Filadelfia calcula anualmente el factor de impacto de todas las revistas incluidas en el Science Citation Index, el Social Science Citation Index y el Arts & Humanities Citation Index, dato que se publica en el denominado Journal Citation Reports (JCR)<sup>23</sup>. El factor de impacto de una revista en un año es el cociente entre el número de citas que han recibido en ese año los documentos publicados en los dos años anteriores y el número de documentos publicados por la revista en esos dos años. Por ejemplo, el factor de impacto de la revista X en el año 1998 se define como el número total de citas recibidas durante el año 1998 por los documentos publicados en los años 1996 y 1997 dividido por el número total de documentos publicados por la revista en cuestión durante los dos años anteriores.

El hecho de publicar en revistas de alto factor de impacto se considera en muchos contextos un criterio de calidad en sí mismo. Dentro de cada disciplina, las revistas de más alto factor de impacto reúnen los mayores criterios de calidad, prestigio y difusión internacional. Dichas revistas suelen contar con un riguroso sistema de selección de originales, por lo que el solo hecho de que un artículo sea seleccionado para su publicación apoya la calidad del mismo.

## **LIMITACIONES DE LOS INDICADORES BIBLIOMÉTRICOS**

Los indicadores bibliométricos constituyen una valiosa herramienta para el estudio de la actividad científica, siempre y cuando se utilicen de forma adecuada, para lo cual es necesario conocer sus limitaciones. Algunas de ellas se exponen a continuación.

### **Limitaciones de los indicadores de actividad científica**

Entre las limitaciones del uso del número de publicaciones como indicador de actividad científica se pueden mencionar:

### *Carácter cuantitativo*

Sólo aportan información sobre la cantidad de publicaciones, pero no sobre su calidad, por lo que tiene mucho interés combinar este indicador con indicadores de impacto y/o juicio de expertos.

### *Diferencias entre áreas*

No pueden efectuarse comparaciones entre áreas temáticas, porque los hábitos de publicación y la productividad de los autores difieren según las áreas. Estas diferencias son especialmente importantes entre las distintas áreas científicas (ciencias sociales, ingeniería y tecnología, ciencias naturales y experimentales), pero también existen diferencias en los hábitos de publicación entre las distintas disciplinas que componen un área. En el caso de la medicina son claras las diferencias atendiendo al carácter básico o clínico de la actividad. Las áreas clínicas con frecuencia presentan una menor productividad, ya que sus profesionales suelen primar la asistencia sanitaria frente a la investigación y publicación de resultados.

### **Limitaciones de las citas**

El uso de las citas en la evaluación de la actividad científica ha sido objeto de amplia polémica<sup>24-26</sup>, y sus principales limitaciones se exponen a continuación.

#### *Problemas conceptuales*

El principal obstáculo del uso de las citas es la ausencia de un modelo aceptado que explique el proceso de citación. Los autores citan aquellos trabajos que han influido en su investigación, pero algunos estudios revelan que no se citan todas las influencias y que sin embargo se incluyen trabajos que no se han utilizado realmente en la investigación. Según Cozzens<sup>27</sup>, en el proceso de citación intervienen varios elementos: *a*) un sistema recompensador por el que se agradece la aportación hecha por el autor citado; *b*) un sistema retórico, por el que se trata de convencer de las tesis mantenidas, y *c*) un sistema de comunicación, por el que se tienden a utilizar revistas de alto prestigio y visibilidad que apoyan la credibilidad del trabajo propio.

#### *Problemas técnicos*

El uso de las citas queda generalmente limitado a la base de datos SCI, única que contiene la bibliografía de los documentos. Algunos problemas observados en el uso de esta base de datos incluyen: errores tipográficos en los nombres de los autores o en la referencia bibliográfica, existencia de homónimos o presencia de autores que firman con distintos nombres a lo largo del tiempo. Por otro lado, hay que señalar que la búsqueda

de citas en el SCI (versión CD-ROM) sólo es posible a través del primer autor de un documento.

### *Diferencias entre tipos documentales*

El tipo de documento puede influir sobre el número de citas que reciben las publicaciones. Se han descrito altas tasas de citación para los documentos metodológicos, que introducen técnicas o métodos que se refieren luego en cada uso. Asimismo, las revisiones y por extensión las revistas especializadas en revisiones obtienen tasas de citación más altas que los otros tipos documentales, debido a que manejan una amplia bibliografía y su consulta es especialmente útil para los científicos.

### *Diferencias entre áreas*

Son inadecuadas las comparaciones entre áreas basadas en el número de citas o en el factor de impacto. La razón es que existen importantes diferencias en el número de citas que pueden recibir los trabajos según las áreas científicas a causa de las diferencias en el tamaño de las comunidades científicas, los hábitos de citación y el ritmo de envejecimiento de la bibliografía.

La frecuencia media de citación en una materia depende del número medio de referencias por artículo en dicha materia, valor determinado en parte por el tamaño del área y sus hábitos de citación. Por otro lado, la probabilidad de que los trabajos sean citados es mayor en áreas grandes y generales, por ejemplo, bioquímica, que en áreas pequeñas o minoritarias, como micología, que difícilmente alcanzarán altas tasas de citación por el menor número de posibles científicos interesados.

Como ejemplo de la diferencia entre áreas se puede señalar que sólo un 5% de los documentos de arte y humanidades reciben alguna cita en los 5 años posteriores a su publicación, frente al 25% de los documentos en ciencias sociales y el 30-40% en ingeniería y tecnología. En el caso de medicina, se ha descrito que un 50-60% de los documentos reciben citas a los 5 años de su publicación<sup>28</sup>. Sin embargo, dentro de la medicina también puede haber diferencias entre disciplinas, atribuibles a su carácter básico o clínico. Los trabajos de investigación básica tienen más posibilidades de ser citados que los trabajos de investigación clínica, ya que esta última utiliza también los resultados de los trabajos de investigación básica.

El ritmo de envejecimiento de la bibliografía en las distintas áreas varía en función de la velocidad con que progresa la investigación en cada área y es un elemento determinante de la rapidez con que son citados los trabajos. En las áreas de rápido crecimiento los investigadores tienden a citar trabajos recientes, mientras que en áreas de lento crecimiento se citan trabajos mucho más antiguos. El ritmo de envejecimiento de

**TABLA 1**  
**Selección de disciplinas del SCI con indicación para cada categoría del número de revistas incluidas, máximo factor de impacto y vida media de la revista. Año 1997**

| Disciplina                      | Número de revistas | FI (primera revista)                  | Vida media (primera revista) |
|---------------------------------|--------------------|---------------------------------------|------------------------------|
| Bioquímica y biología molecular | 249                | 37,297 ( <i>Cell</i> )                | 4,9                          |
| Inmunología                     | 115                | 20,820 ( <i>Immunity</i> )            | 2,3                          |
| Neurociencias                   | 150                | 17,084 ( <i>Trends Neurosci</i> )     | 5,1                          |
| Gastroenterología               | 42                 | 10,250 ( <i>Gastroenterology</i> )    | 6,0                          |
| Sistema cardiovascular          | 61                 | 9,762 ( <i>Circulation</i> )          | 5,5                          |
| Neurología clínica              | 106                | 9,513 ( <i>Ann Neurol</i> )           | 6,5                          |
| Endocrinología y metabolismo    | 81                 | 8,675 ( <i>Diabetes</i> )             | 6,0                          |
| Cirugía                         | 114                | 5,594 ( <i>Ann Surg</i> )             | > 10                         |
| Oftalmología                    | 39                 | 5,250 ( <i>Invest Ophth Vis Sci</i> ) | 5,6                          |
| Nutrición y dietética           | 50                 | 4,536 ( <i>Prog Lipid Res</i> )       | 7,9                          |
| Dermatología                    | 32                 | 4,584 ( <i>J Invest Dermatol</i> )    | 6,2                          |
| Sistema respiratorio            | 22                 | 4,705 ( <i>Am J Resp Crit Care</i> )  | 2,5                          |

FI: factor de impacto.

**TABLA 2**  
**Citas y factor de impacto**

| Citas de documentos  | Factor de impacto (FI) de revista   |
|--|---|
| Las citas son un indicador de la visibilidad, difusión o impacto de la investigación publicada en un documento                   | El FI de una revista es un indicador de su visibilidad y difusión internacional                           |
| Gran parte de las publicaciones nunca son citadas. El 15% de los artículos publicados en una revista reciben el 50% de las citas | El FI de una revista no es una buena estimación del número de citas que va a recibir un documento aislado |
| Las revisiones y los artículos metodológicos reúnen altas tasas de citación  | Las revistas de revisiones tienen altos factores de impacto dentro de su área                             |
| La probabilidad de que un trabajo sea citado varía según las áreas   | Existen variaciones en el FI según las áreas  |
| Las publicaciones tienen más posibilidades de ser citadas en las áreas generales o con gran número de investigadores             | Las revistas de áreas generales tienen mayor FI   |
| Las publicaciones básicas tienen más posibilidades de ser citadas que las clínicas   | Las revistas de áreas básicas tienen mayor FI que las clínicas  |
| La ventana de citación debe variar según las áreas: más amplia para áreas de lento envejecimiento                                | El FI calculado con una ventana de citación de 2 años favorece a las áreas de rápido envejecimiento       |

los documentos de una revista o de un área puede estudiarse a través de la *vida media* de las citas que recibe dicha revista o área<sup>23</sup>. Si una revista tiene una vida media de 5,5 años en 1998 significa que la mitad de las citas que esa revista recibe en 1998 son citas a artículos publicados durante los 5,5 años previos. El resto de las citas se hallará disperso entre todos los restantes años desde que nació la revista. Cuanto menor sea la vida media de una revista, más rápido es el envejecimiento de sus publicaciones (ver por ejemplo la [fig. 2](#)). El recuento del número de citas que recibe un documento suele realizarse a los 2 o 3 años después de su publicación, ya que se ha observado que en promedio, los artículos reciben el máximo número de citas a los 2 años de su publicación. Sin embargo, la situación de este pico máximo de citación puede variar según las áreas. El recuento del número de citas que reciben los documentos a los dos años de su publicación resultará especialmente favorable para las áreas de rápido

crecimiento, en las que se citan documentos muy recientes, y desfavorable para las áreas de lento crecimiento y larga vida media.

Por todo ello, diferencias en los recuentos de citas de documentos publicados por investigadores de áreas distintas no implican necesariamente una diferencia en el impacto de los documentos ni en la calidad de la investigación realizada.

### Limitaciones del factor de impacto

Las limitaciones expuestas para el análisis de citas son también válidas para el uso del factor de impacto, ya que este indicador se calcula en función de las citas que reciben las revistas.

Al igual que ocurría con las citas, el factor de impacto de las revistas también muestra valores muy distintos según las áreas. A modo de ejemplo, la [tabla 1](#)

recoge una selección de disciplinas científicas cubiertas por el SCI. Para cada una de las disciplinas seleccionadas se indica el factor de impacto máximo obtenido por una revista en el año 1997, excluidas las revistas de revisiones. Se observa que el factor de impacto máximo de bioquímica y biología molecular (37,297) es bastante más elevado que los correspondientes a otras áreas como neurociencias (17,084) y sistema cardiovascular (9,762). Ni que decir tiene que este dato no indica la mayor calidad de la investigación publicada en la revista de biología molecular frente a la publicada en las revistas de las otras áreas. Tan sólo pone de manifiesto las diferencias en el número de citas que se reciben en las distintas áreas, hecho en el que influyen, como ya se señaló anteriormente, el tamaño de la comunidad científica, el ritmo de envejecimiento de la bibliografía y los hábitos de citación.

El análisis de las diferencias en factor de impacto según las áreas pone de manifiesto algunos hechos ya comentados con relación a las citas y que se recogen esquemáticamente en la **tabla 2**:

1. Las áreas más *clínicas* presentan los factores de impacto más bajos, mientras que los factores de impacto más elevados corresponden a áreas más básicas, como la bioquímica y la biología molecular o la inmunología.

2. El tamaño pequeño del área, reflejado en el número de revistas que la componen, suele asociarse a bajos factores de impacto, mientras que las áreas más grandes presentan valores más elevados.

3. Las áreas con un rápido envejecimiento de la bibliografía presentan valores altos de factor de impacto. Esto es así debido a que se calcula el factor de impacto en función de las citas recibidas durante los dos años siguientes a la publicación de los documentos (impacto a corto plazo). En las áreas de rápido envejecimiento (p. ej., la biología molecular o la genética) se citan sobre todo documentos muy recientes, y todas estas citas se considerarán en el cálculo del factor de impacto. Por el contrario, un alto porcentaje de las citas que reciben las revistas de áreas de lento crecimiento (como cirugía o pediatría) tendrán una antigüedad mayor de 2 años y no se considerarán en el cálculo del factor de impacto. Para analizar estas últimas revistas sería más adecuado utilizar una ventana de citación más amplia, por ejemplo, de 4 o 6 años (impacto a medio plazo).

4. Distribución asimétrica de las citas. La estimación del impacto de un documento en función del factor de impacto de la revista de publicación es más adecuada para el análisis de un conjunto amplio de documentos, y lo es menos al descender a pequeñas unidades como la producción de un autor en el que se deben contabilizar las citas reales recibidas por el mismo. La razón es que la distribución de citas entre artículos de una revista es muy asimétrica<sup>29</sup>, de forma

que un pequeño núcleo de artículos concentra gran cantidad de citas y otros artículos nunca son citados, incluso las revistas más prestigiosas pueden publicar artículos que nunca se citan. Se ha descrito que el 15% de los artículos publicados en una revista concentran el 50% de todas las citas.

### ALGUNAS REPERCUSIONES DEL USO DE LOS INDICADORES BIBLIOMÉTRICOS EN LA EVALUACIÓN CIENTÍFICA

El uso de los indicadores bibliométricos en los procesos de evaluación de la actividad científica, tanto para la concesión de ayudas como para la promoción profesional, ha contribuido en cierto modo al deterioro de las prácticas de publicación.

La consideración del número de publicaciones como indicador de actividad científica ha desencadenado el denominado «síndrome de publicar o perecer» (síndrome POP o *publish or perish*). Bajo este nombre se designa la situación actual en la que los científicos se ven presionados a publicar, no sólo para dar a conocer los resultados de su investigación, sino también como la única vía de justificar su actividad y obtener un reconocimiento. En esta misma línea se han descrito otros comportamientos como son una tendencia a aumentar el número de autores por documento, hecho que no siempre se asocia a un aumento real de la colaboración entre autores, sino a la denominada «autoría gratuita», y la fragmentación de los trabajos en varias publicaciones que podrían haberse publicado en un solo artículo más completo y más coherente (síndrome LPU o *least publishable unit*)<sup>30</sup>. Entre las medidas tomadas para evitar estas conductas se pueden mencionar todas aquellas orientadas a primar la calidad frente a la cantidad de publicaciones. En este sentido, es frecuente que en algunos procesos de evaluación de personal investigador se solicite al científico que adjunte sólo sus 3 o 5 publicaciones más relevantes con el fin de centrar la evaluación en la calidad de sus contribuciones.

Por otro lado, el uso del número de citas como indicador del impacto o visibilidad de las publicaciones conlleva el riesgo de que se produzca un proceso de «citación dirigida». Conscientes del valor otorgado a las citas, algunos investigadores intentan aumentar su tasa de citación a través de la autocitación o de acuerdos de citación mutua con otros investigadores o grupos. En cualquier caso, la autocitación por sí misma no puede alterar de forma significativa el número de citas que reciben los documentos y cierto grado de autocitación es lógico, ya que la investigación es un proceso acumulativo y los resultados actuales de la investigación suelen ser el punto de partida para nuevos avances. Es interesante señalar cómo los propios editores de las revistas, conocedores del interés de las citas para aumentar la visibilidad de su revista, intentan

de forma dirigida fomentar la autocitación de la revista en los documentos que publican.

Más importante que los cambios en los hábitos de publicación son las repercusiones que puede tener el uso de los indicadores bibliométricos sobre la propia investigación y sobre la carrera investigadora de los profesionales. Aunque parece claro el interés de fomentar la calidad frente a la cantidad de publicaciones, la evaluación de la calidad no es una tarea fácil. No siempre se puede recurrir al juicio de expertos, que, dicho sea de paso, tampoco es un sistema perfecto, y aunque los indicadores de impacto pueden facilitar la tarea, su uso correcto requiere conocer las limitaciones ya comentadas anteriormente<sup>30-32</sup>.

Un primer elemento de distorsión se refiere a los sesgos y limitaciones de las bases de datos del ISI. Los productos del ISI han conseguido ocupar un importante papel en los sistemas de evaluación de la actividad investigadora de distintos países, debido a que estas bases de datos son las únicas que permiten la búsqueda de citas. El ISI es, además, productor del JCR, exhaustivo estudio —único en su especie— que recoge interesantes datos e indicadores bibliométricos sobre las revistas que analiza. Lamentablemente, se ha creado una excesiva dependencia de estos indicadores en muchos países. Es frecuente ver que se premia la publicación en revistas cubiertas por el SCI porque se considera que han pasado un filtro de calidad que ha hecho posible su inclusión en dicha base de datos. También se utiliza la clasificación de revistas en función de su factor de impacto, alentando a la publicación en las de mayor factor de impacto. Estos criterios, que en principio pueden ser válidos para las áreas básicas, de interés internacional, pueden no serlo para las áreas clínicas o aplicadas o de interés más local, peor representadas en la base de datos cuanto más periférico sea el país. El uso de estos criterios en los procesos de evaluación de la investigación conduce a que los investigadores de los países no dominantes también tiendan a publicar en revistas del SCI, en detrimento de las revistas nacionales, escasamente recogidas en dicha base de datos.

Las revistas nacionales constituyen un importante vehículo para la difusión de los resultados de la investigación en las áreas más clínicas y cumplen una valiosa función en la vertebración de la comunidad científica nacional. La calidad y continuidad de estas revistas puede verse amenazada por la creciente tendencia a publicar en revistas del SCI, ya que disminuye el número de originales que éstas reciben y la calidad de los mismos. Este riesgo ha sido puesto de manifiesto por distintos autores y entre las soluciones sugeridas se incluye la elaboración en cada país de listas de publicaciones nacionales de calidad, y su posterior consideración positiva en las evaluaciones, con el fin de permitir la supervivencia y fortalecimiento de las revistas nacionales y de la investigación que, por sus ca-

racterísticas, es de difícil publicación en revistas del SCI. Por otro lado, el uso inadecuado e indiscriminado de los indicadores bibliométricos como son las citas o el factor de impacto puede provocar la migración de investigadores desde pequeñas áreas mal recompensadas en término de citas hacia áreas con mayor visibilidad, cuyo desarrollo puede no ajustarse a las necesidades del país.

Finalmente, la comparación de la productividad de grupos o centros activos en distintas áreas es un error muy extendido que puede resultar especialmente perjudicial. Diferencias en el número de publicaciones o de citas en estos casos no indican investigación de distinta calidad o impacto sino distintos hábitos de publicación y citación.

## CONSIDERACIONES FINALES

La instauración de sistemas de evaluación de la actividad científica es hoy día una necesidad ineludible, ya que es la única forma de optimizar los recursos, siempre limitados, que se destinan a la investigación. Junto a los tradicionales indicadores de *input* como el número de investigadores, los fondos en I+D y/o los equipos disponibles, los indicadores bibliométricos basados en publicaciones se han constituido en los últimos años como instrumentos útiles para estudiar los resultados de la actividad científica. Estos indicadores aportan una valiosa información sobre la situación de la investigación y complementan con éxito al juicio de expertos.

La fiabilidad de los estudios bibliométricos dependerá en gran medida del uso adecuado de los indicadores, que debe hacerse con conocimiento de sus limitaciones y de sus condiciones óptimas de aplicación. En la interpretación de los resultados de estos estudios es importante tener en cuenta las diferencias en los hábitos de publicación según las disciplinas y el tipo de investigación, que limitan las comparaciones entre áreas. Asimismo, se recomienda el uso de diversos indicadores complementarios, cuya convergencia aumenta la fiabilidad de los análisis, y cuyo uso conjunto minimiza el riesgo de que los científicos puedan manipular los indicadores a su favor, con la consiguiente distorsión de la realidad y riesgo de inhabilitación de los indicadores como instrumento de medida.

Sin embargo, la ciencia es una actividad multidimensional, tanto en su naturaleza como en sus resultados, que tienen una dimensión científica, pero también económica, tecnológica y social. Esta realidad multidimensional complica enormemente los procesos de evaluación de la actividad científica, para los cuales se requiere el uso combinado de diversos indicadores<sup>22,32</sup>. Asimismo, no se puede olvidar que la actividad investigadora en biomedicina coexiste en muchos casos con la actividad docente o asistencial que también es necesario evaluar.

No cabe duda de que a pesar de las objeciones que se pueden y se deben hacer a los indicadores bibliométricos, éstos constituyen en la actualidad una herramienta útil y objetiva que facilita una mejor comprensión de la actividad investigadora.

## BIBLIOGRAFÍA

- Narin F, Moll JK. Bibliometrics. *Annual Review of Information Science and Technology* 1977; 12: 35-58.
- White HD, McCain KW. Bibliometrics. *Annual Review of Information Science and Technology* 1989; 24: 119-186.
- Bellavista J, Guardiola E, Méndez A, Bordons M. Evaluación de la investigación. Cuadernos Metodológicos del CIS número 23. Madrid: Centro de Investigaciones Sociológicas, 1997.
- Van Raan AFJ. Advanced bibliometric methods to assess research performance and scientific development: basic principles and recent practical applications. *Research Evaluation* 1993; 3: 151-166.
- OCDE. La mesure des activités scientifiques et techniques, méthode type proposée pour les enquêtes sur le recherche et de développement experimental. París: Manual de Frascati, 1976.
- National Science Board. *Science & Engineering Indicators*, 1996. Washington, DC: US Government, Printing Office, 1996 (NSB 96-211).
- Science & Technologie Indicateurs 1998. Observatoire des Sciences et des Techniques. París: Ed. Economica & OST, 1998.
- Second European Report on S&T Indicators, 1997. EUR 17639. Bruselas: European Commission, 1997.
- Gómez I, Bordons M. Limitaciones en el uso de los indicadores bibliométricos para la evaluación científica. *Política Científica* 1996; 46: 21-26.
- Moed HF. The use of bibliometric indicators for the assessment of research performance in the natural and life sciences. Leiden: DSWO Press, 1989.
- Sancho R. Indicadores bibliométricos utilizados en la evaluación de la ciencia y la tecnología. Revisión bibliográfica. *Revista Española de Documentación Científica* 1990; 13: 842-865.
- 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 (Barc)* 1991; 98: 64-68.
- López Piñero JM, Terrada ML. Los indicadores bibliométricos y la evaluación de la actividad médico-científica (IV). La aplicación de los indicadores. *Med Clin (Barc)* 1992; 98: 384-388.
- King J. A review of bibliometric and other science indicators and their role in research evaluation. *J Inform Sci* 1978; 13: 261-276.
- Garfield E. How ISI selects journals for coverage: quantitative and qualitative considerations. *Current Contents* 1990; 22: 5-13.
- Pestaña A. MEDLINE como fuente de información bibliométrica de la producción española en biomedicina y ciencias médicas. Comparación con el Science Citation Index. *Med Clin (Barc)* 1997; 109: 506-511.
- García Ecudero P, López P. Análisis bibliométrico y literatura de patentes. *Revista General de Información y Documentación* 1997; 7: 182-199.
- Camí J, Fernández MT, Gómez I. La producción española en biomedicina y salud. Un estudio a través del Science Citation Index (1986-89). *Med Clin (Barc)* 1993; 101: 721-731.
- Camí J, Zulueta MA, Fernández MT, Bordons M, Gómez I. Producción científica española en biomedicina y ciencias de la salud durante el período 1990-1993 (Science Citation Index y Social Science Citation Index) y comparación con el período 1986-1989. *Med Clin (Barc)* 1997; 109: 481-496.
- Sánchez Nistal JM. La producción científica de la Comunidad de Madrid en el trienio 1994-1996. En: *Investigación y desarrollo en la Comunidad de Madrid. Tres estudios sobre los recursos, producción y distribución de la actividad científica madrileña*. Madrid: Dirección General de Investigación de la CAM, 1998.
- Martin BR, Irvine J. Assessing basic research. Some partial indicators of scientific progress in radio astronomy. *Research Policy* 1983; 12: 61-90.
- Martin BR. The use of multiple indicators in the assessment of basic research. *Scientometrics* 1996; 36: 343-362.
- Journal Citation Reports. Filadelfia: Institute for Scientific Information.
- Garfield E. Is citation analysis a legitimate evaluation tool? *Scientometrics* 1979; 1: 359-375.
- Garfield E, Welljams-Dorof A. Citation data: their use as quantitative indicators for science and technology evaluation and policy-making. *Science and Public Policy* 1992; 19: 321-327.
- MacRoberts MH, MacRoberts BR. Problems of citation analysis. *Scientometrics* 1996; 36: 435-444.
- Cozzens SE. What do citations counts? The rethoric-first model. *Scientometrics* 1989; 15: 437-447.
- Hamilton DP. Research papers: who's uncited now? *Science* 1991; 251: 25.
- Seglen PO. From bad to worse: evaluation by Journal Impact. *Trends in biochemical Sciences* 1989; 14: 326-327.
- Broad WJ. The publishing game: getting more to less. *Science* 1981; 211: 1.137-1.139.
- Camí J. Impactología: diagnóstico y tratamiento. *Med Clin (Barc)* 1997; 109: 515-524.
- Moravcsik MJ. ¿Cómo evaluar a la ciencia y a los científicos? *Revista Española de Documentación Científica* 1989; 12: 313-325.