



REVIEW

Google Scholar and the h-index in biomedicine: The popularization of bibliometric assessment[☆]

A. Cabezas-Clavijo, E. Delgado-López-Cózar*

EC3: Evaluación de la Ciencia y de la Comunicación Científica, Departamento de Información y Comunicación, Facultad de Comunicación y Documentación, Universidad de Granada, Granada, Spain

Received 5 November 2012; accepted 17 January 2013

Available online 1 September 2013

KEYWORDS

Research evaluation;
Bibliometrics;
h-Index;
Citation analysis;
Periodicals;
Databases,
Bibliographic;
Google Scholar;
Google Scholar
Metrics;
Google Scholar
Citations;
Biomedicine;
Health sciences

Abstract The aim of this study is to review the features, benefits and limitations of the new scientific evaluation products derived from Google Scholar, such as Google Scholar Metrics and Google Scholar Citations, as well as the h-index, which is the standard bibliometric indicator adopted by these services. The study also outlines the potential of this new database as a source for studies in Biomedicine, and compares the h-index obtained by the most relevant journals and researchers in the field of intensive care medicine, based on data extracted from the Web of Science, Scopus and Google Scholar. Results show that although the average h-index values in Google Scholar are almost 30% higher than those obtained in Web of Science, and about 15% higher than those collected by Scopus, there are no substantial changes in the rankings generated from one data source or the other. Despite some technical problems, it is concluded that Google Scholar is a valid tool for researchers in Health Sciences, both for purposes of information retrieval and for the computation of bibliometric indicators.

© 2012 Elsevier España, S.L. and SEMICYUC. All rights reserved.

PALABRAS CLAVE

Evaluación de la
investigación;
Bibliometría;
Índice h;
Análisis de citas;
Revistas científicas;
Bases de datos
bibliográficas;

Google Scholar e índice h en biomedicina: la popularización de la evaluación bibliométrica

Resumen El objetivo de este artículo es hacer una revisión de las características, prestaciones y limitaciones de los nuevos productos de evaluación científica derivados de Google Scholar, como son Google Scholar Metrics y Google Scholar Citations, y del índice h, el indicador bibliométrico adoptado como estándar por estos servicios. Asimismo se reseña la potencialidad de esta nueva base de datos como fuente para estudios en biomedicina y se realiza una comparación del índice h obtenido por las revistas e investigadores más relevantes en el ámbito de la medicina intensiva, a partir de los datos extraídos de Web of Science, Scopus y Google Scholar.

[☆] Please cite this article as: Cabezas-Clavijo A, Delgado-López-Cózar E. Google Scholar e índice h en biomedicina: la popularización de la evaluación bibliométrica. Med Intensiva. 2013;37:343–354.

* Corresponding author.

E-mail address: edelgado@ugr.es (E. Delgado-López-Cózar).

Google Scholar;
 Google Scholar
 Metrics;
 Google Scholar
 Citations;
 Biomedicina;
 Ciencias de la salud

Los resultados muestran que, pese a que los valores medios de índice h en Google Scholar son casi un 30% más elevados que los obtenidos en Web of Science y en torno a un 15% más altos que los recogidos por Scopus, no hay variaciones sustantivas en las clasificaciones generadas a partir de una u otra fuente de datos. Aunque existen algunos problemas técnicos, se concluye que Google Scholar es una herramienta válida para los investigadores en ciencias de la salud, tanto a efectos de recuperación de información como de cara a la extracción de indicadores bibliométricos.

© 2012 Elsevier España, S.L. y SEMICYUC. Todos los derechos reservados.

Introduction

From the perspective of the evaluation of science and of scientific communication, the opening years of the XXI century have witnessed the birth of new tools destined to supplement, if not to replace, the traditional databases offering access to scientific information, such as PubMed, Web of Science (WOS) or Scopus, and the established bibliometric indicators headed by the impact factor. In this context, the new scientific evaluation and access tools examined in the present study (h-index and the Google Scholar family) have popularized access to scientific information and the evaluation of research. The mentioned tools are simple and easy to use and understand, and can be freely accessed at no charge; these features have caused them to gain acceptance in scientific circles at a tremendous pace.

The aim of this article is to review the characteristics of the novel scientific evaluation systems auspiced by Google on the basis of Google Scholar (Scholar Metrics, Scholar Citations) and the h-index (the standard bibliometric indicator adopted by these services). In addition, a comparison is made of the indicators offered by these products with respect to those afforded by the already established Web of Science and Scopus tools – taking as study case the most relevant journals and investigators in the field of Intensive Care Medicine.

The h-index: a new bibliometric indicator for measuring scientific performance

The impact factor was originally introduced in the 1970s by Eugene Garfield with the purpose of selecting journals for library collections under the assumption that the most cited journals would also be the most useful for investigators. Developed on the basis of the famed Citation Indexes (Science Citation Index, Social Sciences Citation Index, now forming part of the ISI Web of Science), acceptance of the impact factor far exceeded the expectations of its creator, and it soon became the gold standard for evaluation, most particularly in the field of biomedicine.

Abuses in the application of this indicator have been denounced for years,^{1,2} and recently it has received criticism both due to the way in which it is calculated and because of its indiscriminate application in evaluative processes,³⁻⁵ particularly referred to the rating of individual investigator performance—a practice which, as was already pointed out by Eugene Garfield years ago, is clearly inappropriate.⁶

The h-index was created in the context of the mentioned search for valid indicators for measuring individual performance. Developed by Jorge Hirsch in August 2005, the novel proposal of the Argentinean physicist condenses the quantitative (production) and qualitative dimensions (citations) of research in a single indicator—this having been a classical aspiration in bibliometrics. A given scientist has an index of “h” when “h” of his or her articles has received at least “h” citations each.⁷ In other words, an investigator with an h-index of 30 has published 30 articles with at least 30 citations each.

The new indicator generated extraordinary interest among specialists in bibliometrics, who soon started to discuss its advantages and inconveniences, and to offer alternatives. This gave rise to an enormous number of publications—to the point whereby in the year 2011 more than 800 articles on the h-index had been published.[†] Apart from the many variants proposed to overcome its limitations (g, e, hc, hg, y, v, f, ch, hi, hm, AWCR, AWCRpA, AW, etc.),⁸ the h-index has expanded from the evaluation of investigators (its original purpose) to the assessment of institutions,⁹ countries¹⁰ and scientific journals.^{11,12}

Support of the new metrics came from the core databases in bibliometrics, Web of Science and Scopus, which immediately incorporated it to its set of indicators – facilitating consultation at all levels of aggregation. Definitive consolidation in turn was facilitated by Google upon incorporating the tool to the two scientific evaluation products derived from its academic browser: Google Scholar Metrics, for journals, and Google Scholar Citations, for investigators.

The h-index has left its mark among many scientists as the definitive indicator for measuring the impact of their investigations. In general terms, it has been very well received by the scientific community, since it is easy to calculate and is able to evaluate investigators using a single number.

The h-index is characterized by progressiveness and robustness, as well as tolerance of statistical deviations and error. It is a progressive marker that can only increase in the course of the career of a scientist, and increasing current values are associated to exponential growth of the difficulty of reaching even higher values. This confers robustness, since the h-index of an investigator either increases in a sustained manner over time or remains stable—thus allowing us

† <http://sci2s.ugr.es/hindex/biblio.php>.

to evaluate the trajectory of the author beyond only widely cited point publications. This likewise protects the indicator from the possible citation or literature search errors commonly found in commercial databases.

Nevertheless, this measure, in the same way as with all the bibliometric indicators, is not without some limitations:

- For adequate calculation of the h-index, we should have an exhaustive information system or database containing all the publications and citations associated with a given investigator. While this is currently elusive, considering the existing overabundance of scientific information, it is essential to take into account that the indicator obtained is directly proportional to the coverage and quality of the database used. In this sense, and as will be seen further below, the h-index varies according to the database used (Google Scholar, Web of Science, Scopus).
- The h-index is affected by the scientific production of the unit studied (organization, journal, investigator). Units with greater scientific production may have higher h-indexes, thereby penalizing procedures based on selective publication strategies.¹³
- The h-index is not valid for comparing investigators from different areas, due to the different production and citation patterns of the different disciplines. This is a limitation inherent to all bibliometric indicators, and can only be overcome by adopting the following principle: compare what is effectively comparable; never compare the h-indexes or impact factors of different disciplines or specialties, since the sizes of the corresponding scientific communities and their behavior in terms of production and citation differ considerably. As an example, in the year 2011, the journal with the greatest impact in Intensive Care Medicine had an impact factor of 11.080, while in Neurology the leading journal had an impact factor of 30.445—this in no way implying that the former publication is less relevant than the latter. It is also known that clinical journals reach more modest values than journals in basic areas, and that reviews draw a larger number of citations than original papers.² These are only some examples of the limitations inherent to the impact factor.
- It is not appropriate to compare junior investigators, with scant production, versus veteran authors with a long scientific trajectory. In order to avoid this problem, it is advisable to compare indicators within similar time windows. Google Scholar has adopted this perspective, offering the indicator for individual investigators referred to both the overall career and the most recent time period (the last 5 years, in the case of journals). On the other hand, in both the Web of Science and in Scopus we can easily delimit the production time periods of the studied units.
- Self-citation (an unavoidable limitation of bibliometric indicators) is not excluded. However, it should be noted that the impact of self-citation practices upon this indicator is smaller than in the case of other metrics such as the citation average, since it only affects documents at the counting limit for the h-index.¹⁴ In any case, the important issue is that we can quickly calculate the indicator with or without self-citations. This can be done with both Web of Science and Scopus, but not with Google Scholar.
- The h-index is not notoriously discriminating, as it takes discrete values and is usually located in a very small range for most investigators. It is therefore advisable to accompany the index with other metrics capable of supplementing the information afforded by the h-index.
- It does not take into account differences in author responsibility and contribution to the publications. This aspect is particularly important in biomedicine, where research constitutes team effort—this in turn being reflected by the large average number of signing authors. In this context, and as occurs with the rest of the bibliometric indicators, the h-index does not ponder the position of the authors in the chain of signatures—this commonly being used to rate the relative contributions of the different authors to a given article.
- The h-index does not take highly cited articles into account. Once the limit established for the h-index has been exceeded, the further citations received by a given article are regarded as irrelevant for increasing the h-index of an investigator. On the other hand, and as has been mentioned above, this fact makes the h-index very stable over time, avoiding the deviations which such highly cited articles induce in other impact indicators.

Thus, rather than measuring point successes in the course of a scientific career, it can be concluded that the h-index essentially measures investigator regularity in terms of production and scientific impact. Although the h-index is very easy to calculate and use, it must be employed with great caution, since it is never a good idea to assess the performance of an investigator or institution on the basis of a single indicator.

The Google family: Google Scholar, Google Scholar Metrics and Google Scholar Citations

Google Scholar

Google Scholar, created in November 2004, is an academic document (articles in journals, books, preprints, doctoral theses, communications at congresses, reports) search tool that provides information on the number of times which a study has been cited, as well as links to the citing documents. Access to full text documents is possible, provided the publications are freely available online, or if a subscription has been obtained, when accessing from a research institution.

Thanks to its excellent information retrieving capacity, easy use and design along the same lines as the general Google browser, Google Scholar has become an obligate resource for an important percentage of scientists seeking information.¹⁵ Some specific studies in biomedicine have shown that Google Scholar is only surpassed in terms of use by the Web of Science, and presents practically the same results as PubMed.¹⁶ This same publication indicates that Google Scholar is used as a data source complementary to Web of Science or PubMed, and that the users positively rate its easy use, fast response and cost-free or open-access nature versus precision and quality of the results, which are regarded as determinant factors in the other two mentioned data sources. In the case of biomedicine, different studies

underscore the importance of Google Scholar in affording access to scientific information, as evidenced both by editorials in the main biomedical journals¹⁷ and by comparative studies on the exhaustiveness and precision or data retrieval with Google Scholar versus PubMed.^{18–21} In the case of Spain, and as refers to the primary care setting, González de Dios et al.²² reported that 70% of the primary care physicians claimed to access scientific literature through general browsers such as Google or Yahoo, while only 29% use specialized databases.

Google Scholar allows access to the full text version of academic works not only from the “official” sources or sites, but also from repositories, personal websites and alternative locations—thus affording enormous visibility of open-access material. In this sense, it must be mentioned that the open-access diffusion of academic material is of great help in facilitating accessibility for research and in increasing diffusion among investigators in the field, as evidenced by the data corresponding to the specialty of Intensive Care Medicine.²³

Furthermore, ordering of the results according to relevance (a criterion based mainly on the number of citations received by the articles) ensures not only that the findings are pertinent to our search but that they are also scientifically relevant—at least as regards the first registries produced by the browser. Such results usually consist of highly cited articles published in journals of great reputation and well known to the investigator. Consequently, and in addition to the ease and speed with which information is retrieved, the investigator feels “comfortable” with the results obtained. As shown by a recent study on the 20 first results of different clinical searches, the articles retrieved by Google Scholar receive a larger number of citations and are published in journals of higher impact factor than the documents retrieved by PubMed.²¹

However, Google Scholar has a series of limitations that must be taken into account when using it as a source of information.^{24,25} The main limitation is the opaque source inclusion policy of the browser. Although it is known that Google Scholar scans the information contained in the websites of universities and research institutions, repositories or websites of scientific publishers, we do not exactly know what these sources are or what the size or scope of the documentary base is in each case. On the other hand, along with research publications we can also find teaching material and even data of an administrative nature, which are scantily relevant to research work. Likewise, the data generated by Google Scholar are not normalized in any way—this being a consequence of the broad coverage involved, the variety of information sources used, and automatic processing of the information.²⁵

Comparative studies of Google Scholar versus PubMed, the genuine *sancta sanctorum* of the biomedical databases, show very similar results in terms of the exhaustiveness of information retrieval, though performance is considered more favorable to the product of the United States National Library of Medicine when it comes to defining searches of certain complexity.²⁰ For this reason it may be useful to employ a hybrid strategy for document retrieval, using PubMed to conduct the literature search prior to any study,¹⁸ and Google Scholar for identification and quick access to concrete documents, or in the context of very defined

searches such as articles on a topic in a given journal or by a concrete author.

The fact that Google Scholar generates citations of indexed articles has made it a serious competitor for the traditional databases used in scientific evaluations. In this context, an extensive body of literature compares results on data from the Web of Science, Scopus and Google Scholar.^{26–28} Authors such as Harzing and Der Wal²⁹ even claim that Google Scholar represents a democratization of the analysis of literature citations.

Google Scholar citations: bibliometric profiles of investigators

Knowing the rapid expansion of Google Scholar among scientists, and aware of the needs of this sector in relation to scientific evaluation, Google in July 2011 launched a new product: Google Scholar Citations (“Mis citas”, in Spanish). This service allows an investigator to automatically establish a webpage with his or her scientific profile, specifying the published documents registered in Google Scholar, along with the number of citations received by each of them—thereby generating a series of bibliometric indicators headed by the ever present h-index.³⁰ These statistics in turn are automatically updated with the new information located by Google Scholar (new authored publications or new citations of previous work); as a result, the scientific profile is continuously kept up to date without the investigator having to intervene in any way. Scholar Citations offers three bibliometric indicators. In addition to the mentioned h-index, it specifies the number of citations of the works published by the investigator, and the i10-index, i.e., the number of studies published by the investigator which have received more than 10 citations (Fig. 1). The mentioned indicators are calculated for both the global academic career and for the most recent period (the last 6 years).

The registry process for investigators is very simple: after entering an institutional mail address and the signing names (and variants) we usually use, Google Scholar displays the articles of which we are signing authors, for confirmation or rejection purposes. These articles conform the body of information from which our bibliometric data are drawn. In addition, the investigator can decide whether or not to make his or her profile public, and can moreover edit the registries—thereby contributing to correct or normalize the information of the browser, merge duplicate registries, and even manually enter other works not contemplated in the academic search engine. In this latter case, however, we do not have the number of citations received by these documents. Likewise, the system allows us to export the registries in the most common bibliographical formats.

Another advantage of establishing a profile in Scholar Citations is that we can receive updates on documents of relevance to our academic interests. By means of algorithms that take into account the words we use in our works and their co-authors, as well as the flow of citations among articles and journals, Google Scholar offers us documents suited to our scientific needs.

On the other hand, we can establish a link to co-authors who are registered in Scholar Citations with a public profile and add the key words or descriptors that best define our

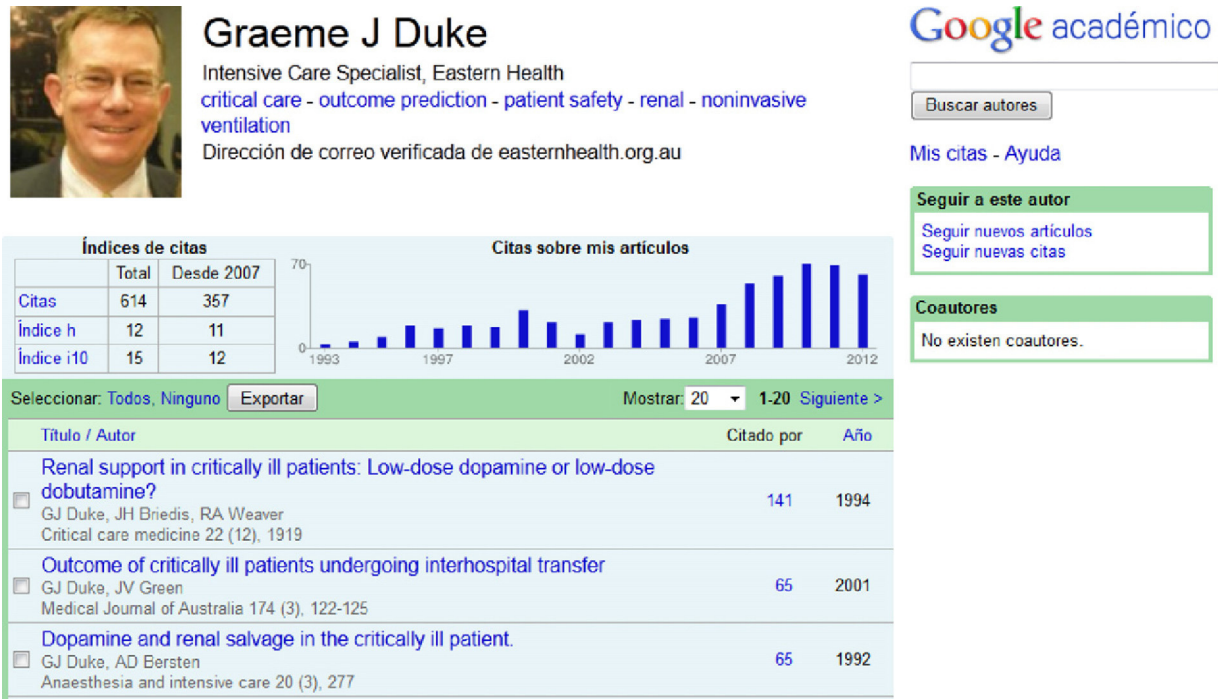


Figure 1 Profile of an investigator in Google Scholar Citations.

work. This service is worth mentioning, since the investigator can personally indicate the disciplines or sub-disciplines in which he or she is expert. This allows the creation of a list of investigators according to the total number of citations received (Fig. 2). In this context, mention must be made of the unfortunate lack of normalization in the

assignment of descriptors, which allows one same concept to be represented in different ways. As an example, in the field of Intensive Care Medicine we can find the descriptors Critical Care, Critical Care Medicine, Intensive Care or Intensive Care Medicine, among others. Likewise, it has been found that some investigators use uncommon terms with the

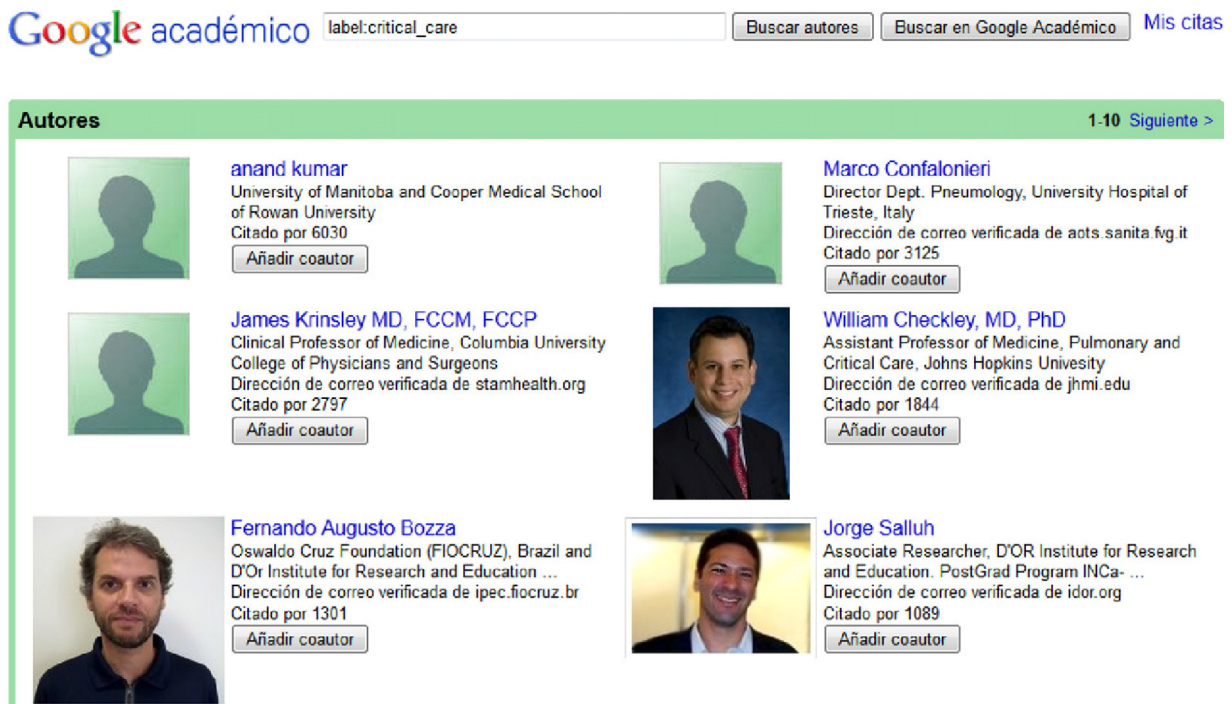


Figure 2 Investigators with the largest number of citations in Google Scholar Citations, linked to the descriptor Intensive Care Medicine (Critical Care).

purpose of appearing as leading specialists in a given sub-specialty. Another phenomenon refers to investigators with activity in more than one field of knowledge. As an example, it is common in bibliometrics to find scientists in fields such as physics, chemistry or computer science that allow them to appear in leading positions in disciplines that do not correspond to the main body of their activity—giving rise to lists that are scantily realistic or even directly false. A notorious situation in the disciplines in which the authors of the present study are specialists (bibliometrics, scientometrics) is the presence alongside leading investigators in the field of scientists who are completely unknown to the specialists in the discipline. In this sense, more seriousness could be expected from Google by joining or merging descriptors that may be regarded as almost synonymous. We even could suggest the use of standardized or normalized vocabularies such as the MESH headings in biomedicine. In turn, in the case of the investigators, greater rigor in establishing links to the different disciplines would be desirable.

From a bibliometric perspective, this service has some limitations, particularly the absence of the total number of publications of a given investigator—this being the most evident indicator for assessing the scientific activity of a professional. Another shortcoming is the lack of indicators linked to the capacity to publish in reputed journals or diffusion media—though the introduction of Scholar Metrics possibly may overcome this defect in the short term, if Google decides to integrate both products.

Another aspect to be taken into account is possible deceitful or fraudulent behavior on the part of some authors. In the same way as has been detected in relation to the disciplines and sub-disciplines to which investigators ascribe themselves, it is also feasible for people to assign as their own articles which in fact have been published by others, as has been well pointed out by Anne W. Harzing³¹—though such practices are easily detectable. Another possibility is manipulation of the database, entering false authors that inflate our registry of citations. The experiments of Labbe³² and Delgado López-Cozar et al.³³ show that this is feasible and simple to do—in contrast to the situation in closed and controlled environments such as the large databases of the Web of Science or Scopus, where such manipulation, while possible, is much more complicated.

Another uncertainty is whether the directories generated are representative of the research done in a given field. Investigators with little academic impact have few incentives for creating a public profile of their own, since they could face embarrassing comparisons with other colleagues who have a stronger curriculum. Therefore, one of the risks is the creation of a directory limited to high-impact scientists that is scantily representative of the research done in a given field.³⁰

Google Scholar metrics: rankings of scientific journals

In April 2012, Google launched Google Scholar Metrics (GSM), called “Estadísticas” in Spanish. This is a free and open-access bibliometric product offering the impact of scientific journals and other documental sources, and which uses the h-index as ordering criterion.

This product can be consulted in two ways. A first option is to access the rankings of the 100 journals with the highest h-index according to the language in which they are published (Fig. 3). At present, 10 languages are available: English, Chinese, Portuguese, German, Spanish, French, Korean, Japanese, Dutch and Italian—though in this aspect it must be pointed out that Google Scholar Metrics does not appear to be good at differentiating journals with bilingual contents. Since November 2012, which is when the latest upgrade on the product was launched, we can also consult the first 20 publications according to their h-index in 8 areas of knowledge and corresponding to 313 disciplines—though only for journals in English.³⁴

The second option is to use the search box to directly insert words included in the titles of the journals. In this case, the search is made in all the sources indexed by Google Scholar Metrics (without necessarily having to rank at the top of each language or discipline), with the return of a maximum of 20 results ordered according to the h-index. The sources indexed in Google Scholar Metrics are those that have published at least 100 articles in the period 2007–2011 and which have received some citation (i.e., those with an h-index of 0 are excluded). This indicator covers the works published between 2007 and 2011, as well as the citations received up until 15 November 2012. It is therefore a static information system with regular updates, though Google has not provided details on how these will be made.

For each of the journals we can access the information on the articles that contribute to the h-index. Thus, for a journal with an h-index of 14, as is the case of *Medicina Intensiva*, Google Scholar Metrics shows us the 14 articles that have received 14 citations or more (Fig. 4). In the same way, by clicking on them, we can access the citing documents; it is therefore easy to analyze the sources that largely contribute to the h-index of a given journal.

The fact that relatively large time windows are used (5 years) implies that the indicator will remain quite stable over time. Since the h-index is an indicator of weak discriminatory capacity (it takes discrete values, and it is very difficult to further raise the index beyond a certain citation threshold), Google has included another indicator—the median of the citations received by the articles that contribute to the h-index—as a second criterion for ordering the journals.

However, Google Scholar Metrics has some very important limitations that make it scantily reliable when used for scientific evaluation purposes^{34,35}:

- Coverage: indiscriminate mixing in the rankings of documental sources as diverse as journals, series and collections included in repositories, databases or congress proceedings.
- Errors in citation counts, due to a lack of normalization of journal titles, particularly referred to journals with versions in different languages.
- Main presentation of results by languages, not by scientific disciplines, as is usual in products of this kind. Thus, comparisons are made of journals corresponding to different scientific disciplines, with different production and citation patterns.
- The impossibility of searching by areas or disciplines in the journals published in languages other than English.

	Título	Índice h5	Mediana h5
inglés			
chino	1. Revista Española de Cardiología	24	32
portugués	2. Psicothema	22	33
alemán	3. Salud Pública de México	21	28
español	4. International Journal of Clinical and Health Psychology	20	25
francés	5. Emergencias: Revista de la Sociedad Española de Medicina de Urgencias y Emergencias	17	22
italiano	6. Nutrición hospitalaria	16	22
japonés	7. Medicina Clínica	15	25
neerlandés	8. Revista de educación	15	25
coreano	9. Gaceta Sanitaria	15	24
	10. Revista Española de Salud Pública	15	21
	11. Archivos de Bronconeumología	14	21
	12. Revista médica de Chile	14	19
	13. Medicina intensiva	14	18
	14. Nefrología: publicación oficial de la Sociedad Española de Nefrología	14	18

Figure 3 Lists of publications with the highest h-indexes in Spanish (2007–2011), according to Google Scholar Metrics.

Limitation of results in the key word-based search to the 20 titles with the highest h-index.

- Insufficient citation window for evaluating publications of national impact in disciplines referred to the human and social sciences.

Having pointed out the above limitations, a positive evaluation must be made of the step taken by Google, and which will facilitate consultation of the impact of journals by investigators who do not have access to the traditional citation databases, which entail important economical costs. In

	Título / Autor	Citado por	Año
inglés			
chino	Neumonía y síndrome de distrés respiratorio agudo producido por el virus influenza A (H1N1)		
portugués	C de Haro López, RF Roca, JV Daunis Med Intensiva 33 (9), 455-458	124	2009
alemán			
español	Estudio Nacional de Vigilancia de Infección Nosocomial en Unidades de Cuidados Intensivos. Informe evolutivo de los años 2003-2005	60	2007
francés	F Álvarez-Lerma, M Palomar, P Olaechea, JJ Otaol, J Insausti, E Cerdá Medicina Intensiva 31 (1), 6-17		
italiano			
japonés	Current insights in intra-abdominal hypertension and abdominal compartment syndrome	24	2007
neerlandés			
coreano	Recomendaciones de tratamiento al final de la vida del paciente crítico	22	2008
	JL Monzón Marín, I Saralegui, R Abizanda i Campos, L Cabré Pericas, S ... Med Intensiva 32 (3), 121-133		
	Documento de Consenso (SEMES-SEMICYUC). Recomendaciones del manejo diagnóstico-terapéutico inicial y multidisciplinario de la sepsis grave en los Servicios de Urgencias hospitalarios	20	2007
	G DE TRABAJO DE SEPSIS, C LEON GIL, L GARCIA-CASTRILLO RIESGO, M MOYA MIR, A ... Medicina intensiva 31 (7), 375-387		
	Guía de práctica clínica basada en la evidencia para el manejo de la sedo-analgésia en el paciente adulto críticamente enfermo	19	2007
	E CELIS-RODRIGUEZ, J BESSO, C BIRCHENALL, MA DE LA CAL, R CARRILLO, G ... Medicina intensiva 31 (8), 428-471		
	Prevalencia y tratamiento de la anemia en el paciente crítico	19	2007
	M Muñoz, SR Leal-Noval, JA García-Erce, E Naveira Medicina Intensiva 31 (7), 388-398		

Figure 4 Most cited articles in *Medicina Intensiva* (2007–2011), according to Google Scholar Metrics.

addition to offering bibliometric indicators to all investigators, this situation may serve to stimulate competition among the different products.

Intensive care medicine through the h-index: comparing Google Scholar, Web of Science and Scopus

In order to evaluate the behavior of Google Scholar and Google Scholar Metrics, and to clearly visualize the differences in results compared with the traditional databases, we have contrasted the h-indexes obtained in a sample of journals and of significant authors in the field of Intensive Care Medicine. As measure of comparison of the positions of the journals and authors, we have used the Spearman correlation coefficient (ρ)—a statistic commonly used in bibliometric studies to measure the association between two variables based on their position in different rankings.³⁶

The comparison has been made with the 26 journals registered in the Journal Citations Reports (JCR) in 2011, within the category Intensive Care Medicine (Critical Care Medicine), based on three data sources: Web of Science,

Scopus and Google Scholar Metrics. Considering the articles published between 2007 and 2011, the journals *American Journal of Respiratory and Critical Care Medicine*, and *Chest and Critical Care Medicine* show the highest h-indexes, independently of the data source used. The mean h-index for the global 26 journals is 28 in the case of the Web of Science, 32 for the journals in Scopus, and 36 in the case of Google Scholar Metrics (Table 1).

This means that the data from Google Scholar Metrics are 28% higher than those obtained from the Web of Science, and 13% higher than those contained in Scopus. It should be mentioned that the data from Google Scholar Metrics count the citations received until April 2012, while the calculations of the other two databases were made 6 months later. As a result, the percentage difference between Google Scholar Metrics and the other two databases may be somewhat underestimated.

On the other hand, the mean h-index of the sample of journals in Intensive Care Medicine analyzed is 13.5% greater in Scopus than in the Web of Science (Table 2). However, comparison of the order of the journals according to their h-index based on the Spearman correlation coefficient (ρ) shows the differences among databases to be negligible. The correlation coefficient, 0.99 in all three cases, at a

Table 1 h-Indexes of journals in intensive care medicine (n = 26) according to the Web of Science, Scopus and Google Scholar Metrics (2007–2011).

Journal	H-index			Ranking		
	WOS	Scopus	GSM	WOS	Scopus	GSM
<i>American Journal of Respiratory and Critical Care Medicine</i>	81	93	96	1	1	1
<i>Chest</i>	68	79	89	2	2	2
<i>Critical Care Medicine</i>	64	74	83	3	3	3
<i>Intensive Care Medicine</i>	47	55	63	4	4	4
<i>Journal of Trauma-Injury Infection and Critical Care</i>	47	40	53	4	6	5
<i>Critical Care</i>	40	45	53	6	5	5
<i>Journal of Neurotrauma</i>	37	40	40	7	6	8
<i>Resuscitation</i>	36	40	45	8	6	7
<i>Shock</i>	31	34	38	9	9	9
<i>Injury</i>	29	34	38	10	9	9
<i>Current Opinion in Critical Care</i>	24	27	32	11	11	11
<i>Pediatric Critical Care Medicine</i>	23	27	30	12	11	12
<i>Burns</i>	23	25	29	12	13	13
<i>Neurocritical Care</i>	21	24	25	14	14	15
<i>Journal of Critical Care</i>	20	23	26	15	15	14
<i>Seminars in Respiratory and Critical Care Medicine</i>	19	23	sd	16	15	sd
<i>Respiratory Care</i>	19	22	23	16	17	17
<i>American Journal of Critical Care</i>	17	21	25	18	18	15
<i>Critical Care Clinics</i>	17	19	22	18	19	18
<i>Minerva Anesthesiologica</i>	17	19	19	19	19	19
<i>Anesthesia and Intensive Care</i>	14	17	18	21	21	20
<i>Critical Care Nurse</i>	9	13	18	22	22	20
<i>Critical Care and Resuscitation</i>	8	12	12	23	23	23
<i>Anesthesiologie und Intensivmedizin</i>	8	8	4	23	25	25
<i>Medicina Intensiva</i>	7	10	14	24	24	22
<i>Anästhesiologie, Intensivmedizin, Notfallmedizin, Schmerztherapie</i>	6	7	7	25	26	24
Promedio	28	32	36			

Source: Google Scholar Metrics data, taken from http://scholar.google.com/citations?view_op=top_venues (number of citations updated to 1 April 2012). Data from Web of Science and Scopus compiled on 26 September 2012.

Table 2 Comparison of the h-indexes of journals in intensive care medicine according to databases.

Source	No. ^a	Rho ^b	Difference in h-index
WOS-SCOPUS	26	0.994	+13.5% (SCOPUS)
WOS-GSM	25	0.989	+28.2% (GSM)
SCOPUS-GSM	25	0.994	+12.9% (GSM)

Difference in h-index: percentage differences in the mean h-indexes between databases. The database with the largest mean h-index appears in parentheses. No.: journals; rho: Spearman correlation coefficient.

^a Since Google Scholar Metrics contains a journal for which no data are available, the statistical analysis is performed on 25 journals in the case of this database, instead of on the 26 journals of the analysis in the WOS-Scopus comparison.

^b The correlation is significant at 1%.

significance level of 1%, implies that despite the quantitative differences in h-index among the databases, ordering of the journals is practically the same, regardless of the data source employed.

In the same way as for the journals, we have calibrated the h-index of reputed scientists in the field of Intensive Care Medicine. We have taken the 20 most productive investigators in 2007–2011 within the category Intensive Care Medicine (Critical Care Medicine) in Web of Science. To this effect, we designed a search equation through the advanced search option (WC = Critical Care Medicine AND PY = 2007–2011). In this way, and making use of the option Refine Results of the mentioned database, we considered the most productive investigators and calculated their h-indexes

Table 4 Comparison of the h-indexes of investigators in Intensive Care Medicine according to databases.

Source	No.	Rho ^a	Difference in h-index
WOS-SCOPUS	20	0.918	+10.9% (SCOPUS)
WOS-GS	20	0.934	+28.9% (GS)
SCOPUS-GS	20	0.968	+16.2% (GS)

Difference in h-index: percentage differences in the mean h-indexes between databases. The database with the largest mean h-index appears in parentheses. No.: journals; rho: Spearman correlation coefficient.

^a The correlation is significant at 1%.

in Web of Science, Scopus and Google Scholar. Some investigators with common names that could complicate the search process were excluded from the sample. The data referred to the h-indexes of these investigators are shown in Table 3. The mean h-index is seen to be 23 in the case of Web of Science, 25 for Scopus, and 29 when using Google Scholar as source for the analysis.

The differences in mean h-index among the databases are very similar to those observed in the case of the journals, with values in Google Scholar almost 30% higher than those in Web of Science, and 16% higher than in Scopus. On the other hand, the difference between Web of Science and Scopus, favorable to the latter, is a mere 10.9% (Table 4).

Few differences are observed in the rankings according to one variable or other, though the most relevant investigator in Web of Science (J.B. Holcomb) is not the leading investigator in either Scopus or Google Scholar (R. Bellomo). The

Table 3 h-Indexes of leading investigators in intensive care medicine (n=20) according to the Web of Science, Scopus and Google Scholar (2007–2011).

Investigator	h-Index			Ranking		
	WOS	Scopus	GS	WOS	Scopus	GS
Holcomb JB	37	36	43	1	2	2
Bellomo R	36	40	48	2	1	1
Matthay MA	27	29	35	3	3	3
Rello J	25	27	33	4	5	4
Angus DC	25	26	31	4	8	6
Reinhart K	25	26	31	4	8	6
Moore EE	24	27	31	7	5	6
Kollef MH	24	27	31	7	5	6
Azoulay E	23	26	28	9	8	11
Wade CE	23	25	29	9	11	10
Kellum JA	22	28	33	11	4	4
Ely EW	21	24	25	12	12	14
Scalea TM	21	22	25	12	14	14
Herndon DN	20	22	28	14	14	11
Giannoudis PV	19	24	27	15	12	13
Annane D	19	21	25	15	16	14
Gajic O	18	20	22	17	17	18
Demetriades D	18	19	23	17	18	17
Brochard L	15	19	21	19	18	19
Sprung CL	15	19	20	19	18	20
Promedio	23	25	29			

Data compiled on 28 September 2012.

Spearman correlation coefficient (ρ) is very high among the three data sources, particularly between Scopus and Google Scholar (0.968), at a significance level of 1%.

Discussion

At present we have many ways to evaluate research on the basis of bibliometric parameters, and different data sources are available for this purpose. As a result, the impact factor and ISI Web of Science no longer monopolize these evaluating functions. Indicators such as the h-index or databases and products such as those offered by Google are only some of the most recognizable examples of the existing alternatives.

The h-index above all is an indicator used to assess the scientific trajectory of an investigator, measuring his or her regularity in terms of scientific productivity and impact. It can also be used to measure the performance of journals and institutions, though in this case it is advisable to confine calculation of the index to limited periods of time. Despite its very easy calculation and use, and its solidness and tolerance of statistical deviations, the h-index must be used with great caution, and in any case measures and controls must be adopted to neutralize or minimize its sources of bias. One of these controls, in the case of biomedicine, would be to always perform comparisons within one same discipline, and in the case of investigators, to consider scientific careers of similar duration or apply some normalization criterion according to the years of research trajectory or the academic position held. Likewise, it must be stressed that it is never a good idea to measure the performance of an investigator or institution with a single indicator. In this sense, it is advisable to use a battery of indicators in evaluative processes based on bibliometric parameters—if possible, in conjunction with expert counseling of specialists in the corresponding scientific field.

It must be made clear that use of the h-index does not mean that the impact factor is no longer applicable as a scientific quality indicator. It is a good marker for measuring difficulties and barriers in accessing a given scientific journal, and for evaluating its impact. In the case of Medicine, it is essential for investigators to publish their work in journals with an impact factor, since the latter remains the gold standard for assessing the ranking of scientific journals.³⁷ In this context, journals with higher impact factors receive larger numbers of manuscript submissions, and are obliged to reject many of them. Authors therefore have to compete among each other for publication in these journals. For this reason the impact factor can be regarded as a competitiveness indicator.³⁸ A good example is provided by this same journal (*Medicina Intensiva*), which saw a 200% increase in the number of original manuscripts received after obtaining its first impact factor.³⁹ This phenomenon shows that the scientific community tends to regulate itself, since the leading journals in turn incorporate more exhaustive and demanding controls—allowing publication only by the more prominent members of the scientific community. Logically, this does not prevent such journals from also publishing works of limited value or usefulness for investigators, but in this case such studies represent a significantly lesser percentage than in

the case of journals that have no impact factor. For this reason it is also advisable to remember that the impact factor of a journal is not representative of the impact of the articles that are published in it. Believing that a given work, the authors signing it, and the institution behind it “inherit” the impact factor of the journal in which it has been published is a widespread and mistaken practice—particularly among Spanish investigators.⁴⁰

In sum, it is important to take into account that all bibliometric indicators have limitations, and that we should avoid using a single criterion for the evaluation of research; rather, a series of indicators should be used in order to ensure rigorous and fair assessments.

Regarding the use of Google Scholar as a source of scientific information, we can conclude that it is a serious alternative to the traditional databases. The browser developed by Google offers results similar to those of other data sources, and may be preferable for establishing a first documental approach to a given topic, or for locating certain literature references. However, in biomedicine it is still preferable to make use of the controlled environment afforded by specialized services such as PubMed, with journals and contents validated by the scientific community.

In turn, although Google Scholar Citations and Google Scholar Metrics (the products derived from Google Scholar) pose problems associated with the lack of normalization of the browser, they offer fast and cost-free bibliometric information of great value, referred to the performance of both individual investigators and scientific journals. This undoubtedly is of help in familiarizing investigators with the bibliometric indicators and with their use in the evaluation of scientific activity.

A surprising observation on comparing the bibliometric data offered by Google Scholar with the traditional bibliometric databases Web of Science and Scopus, considering both investigators and journals, is that the results are practically identical, at least in Intensive Care Medicine. There is no disparity in the rankings obtained: all three systems coincide in identifying the most relevant and influencing authors and journals. The only differences refer to scale and size: Google Scholar retrieves more citations than Scopus, and the latter in turn retrieves more citations than the Web of Science.

Beyond the technical aspects, it is true that there are two important differences. The first is of a conceptual nature. Scientific ethics demands prior controls to certify and validate the generation and diffusion of knowledge, i.e. it demands controlled settings. In this context, Google Scholar works in an uncontrolled environment, since it processes all the scientific information it is able to capture on the Internet, regardless of how it is produced and communicated. In this sense, the new information sources are in stark contrast to the traditional databases that exclusively rely on filtered and screened scientific sources (journals and congresses) with the most rigorous scientific controls. However, to the effects of scientific evaluation, control or the lack of control makes no difference. The other mentioned difference refers to economical considerations, since the price to pay for one source or other is not the same. In this context, while Google Scholar is a free and open-access service, the multidisciplinary databases represent great costs for the

research institutions, and the licenses are moreover very restrictive.

In conclusion, while the impact factor introduced the evaluation of scientific performance in the academic world, and Google popularized the access to information scientific, the h-index and bibliometric tools derived from Google Scholar have further extended scientific evaluation, making the retrieval of academic information and the calculation of bibliometric indicators more accessible. In sum, the great novelty of these last few years has been the fact that these products are no longer restricted to specialists and are now open to any investigator, thanks to the ease with which these metrics are calculated and the free and open access allowed by the new tools—giving rise to what may be regarded as a kind of popularization of scientific evaluation.

Financial support

This study was supported by project HAR2011-30383-C02-02 of the General Research and Management Board of the National R&D and Innovation Plan, Spanish Ministry of Economy.

The funding sources have not participated in the design of the study, in data collection, analysis or interpretation, in drafting of the manuscript, or in the decision to submit the study for publication.

Conflicts of interest

The authors declare that they have no conflicts of interest.

References

- Seglen P. Why the impact factor of journals should not be used for evaluating research. *Br Med J.* 1997;314:498–502.
- Bordons M, Fernandez M, Gómez I. Advantages and limitations in the use of impact factor measures for the assessment of research performance in a peripheral country. *Scientometrics.* 2002;53:195–206.
- Rossner M, van Epps H, Hill E. Show me the data. *J Cell Biol.* 2007;179:1091–2.
- Simons K. The misused impact factor. *Science.* 2008;322:165.
- Vanclay JK. Impact factor: outdated artefact or stepping-stone to journal certification? *Scientometrics.* 2012;92:211–38.
- Garfield E. The impact factor and using it correctly. *Unfallchirurg.* 1998;48:413–4.
- Hirsch J. An index to quantify an individual's scientific research output. *Proc Natl Acad Sci USA.* 2005;102:16569–72.
- Alonso S, Cabrerizo FJ, Herrera-Viedma E, Herrera F. h-Index: a review focused in its variants, computation and standardization for different scientific fields. *J Informetr.* 2009;3:273–89.
- Prathap G. Hirsch-type indices for ranking institutions' scientific research output. *Curr Sci.* 2006;91:1439.
- Csajbók E, Berhidi A, Vasas L, Schubert A. Hirsch-index for countries based on essential science indicators data. *Scientometrics.* 2007;73:91–117.
- Braun T, Glänzel W, Schubert A. A Hirsch-type index for journals. *Scientometrics.* 2006;69:169–73.
- Delgado López-Cózar E, Orduña Malea E, Marcos Cartagena D, Jiménez Contreras E, Ruiz-Pérez R. Journal Scholar: una alternativa internacional, gratuita y de libre acceso para medir el impacto de las revistas de Arte, Humanidades y Ciencias Sociales. *EC3 Working Papers.* 2012;5:1–25.
- Costas R, Bordons M. Una visión crítica del índice h: algunas consideraciones derivadas de su aplicación práctica. *Prof Inf.* 2007;16:427–32.
- Costas R, Bordons M. The h-index: advantages, limitations and its relation with other bibliometric indicators at the micro level. *J Informetr.* 2007;1:193–203.
- Nicholas D, Williams P, Rowlands I, Jamali HR. Researchers' e-journal use and information seeking behaviour. *J Inf Sci.* 2010;36:494–516.
- Hightower C, Caldwell C. Shifting sands: science researchers on Google Scholar, Web of Science, and PubMed, with implications for library collections budgets. *Issues Sci Tech Libr.* 2010;63. <http://dx.doi.org/10.5062/F4V4054J>.
- Giustini D. How Google is changing medicine. *Br Med J.* 2005;331:1487–8.
- Shultz M. Comparing test searches in PubMed and Google Scholar. *J Med Libr Assoc.* 2007;95:442–5.
- Falagas ME, Pitsouni EI, Malietzis GA, Pappas G. Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses. *FASEB J.* 2008;22:338–42.
- Anders ME, Evans DP. Comparison of PubMed and Google Scholar literature searches. *Respir Care.* 2010;55:578–83.
- Nourbakhsh E, Nugent R, Wang h Cevik C, Nugent K. Medical literature searches: a comparison of PubMed and Google Scholar. *Health Info Libr J.* 2012;29:214–22.
- González de Dios J, Flores Canoura A, Jiménez Villa J, Gutiérrez Fuentes JA. Qué revistas médicas españolas leen y cómo se informan los médicos de atención primaria. *Aten Primaria.* 2011;43:629–36.
- Riera M, Aibar E. ¿Favorece la publicación en abierto el impacto de los artículos científicos? Un estudio empírico en el ámbito de la medicina intensiva. *Med Intens.* 2013;37:232–40.
- Jacsó P. Google Scholar revisited. *Online Inf Rev.* 2008;32:102–14.
- Torres-Salinas D, Ruiz-Pérez R, Delgado-López-Cózar E. Google Scholar como herramienta para la evaluación científica. *Prof Inf.* 2009;18:501–10.
- Meho L, Yang K. Impact of data sources on citation counts and rankings of LIS faculty: Web of Science versus Scopus and Google Scholar. *J Am Soc Inf Sci Technol.* 2007;58:2105–25.
- Kulkarni A, Aziz B, Shams I, Busse JW. Comparisons of citations in Web of Science, Scopus, and Google Scholar for articles published in general medical journals. *JAMA.* 2009;302:1092–6.
- Bar-Ilan J. Citations to the Introduction to informetrics indexed by WOS, Scopus and Google Scholar. *Scientometrics.* 2010;82:495–506.
- Harzing A-W, van der Wal R. Google Scholar as a new source for citation analysis. *Ethics Sci Environ Polit.* 2008;8:61–73.
- Cabezas-Clavijo Á, Torres-Salinas D. Google scholar citations y la emergencia de nuevos actores en la evaluación de la investigación. *An Think EPI.* 2012;6:147–53.
- Hertther NK. Scholar citations—Google moves into the domain of Web of Science and Scopus. *Inf today.* 2011.
- Labbé C. Ike Antkare one of the great stars in the scientific firmament. *ISSI Newslett.* 2010;6:48–52.
- Delgado-López-Cózar E, Robinson-García N, Torres-Salinas D. Manipulating Google Scholar Citations and Google Scholar Metrics: simple, easy and tempting. *EC3 Working Papers.* 2012:6.
- Delgado-López-Cózar E, Cabezas-Clavijo Á. Google Scholar Metrics revisado: Ahora empieza a ir en serio. *EC3 Working Papers.* 2012:8.
- Delgado-López-Cózar E, Cabezas-Clavijo Á. Google Scholar Metrics: an unreliable tool for assessing scientific journals. *Prof Inf.* 2012;21:419–27.
- Torres-Salinas D, Delgado-López-Cózar E, Jiménez-Contreras E. Rankings of departments and researchers within a university

- using two different databases: web of Science versus Scopus. *Scientometrics*. 2009;80:761–74.
37. Fernandez-Mondéjar E, Hugo Ceraso D, Gordo Vidal F, Guerrero López F, Colmenero Ruiz M. Factor de impacto, una herramienta imperfecta pero imprescindible. *Med Intens*. 2010;34: 493–4.
 38. Delgado López-Cózar E. Herramientas e indicadores bibliométricos para la evaluación de la investigación en Ciencias de la Salud. In: *Procedimientos y herramientas en la traslación de la investigación biomédica en cooperación*. Universidad Internacional Menéndez y Pelayo. Santander, 26 July. 2010.
 39. Fernández Mondéjar E, Guerrero López F, Colmenero Ruiz M, Gordo Vidal F. Actividad editorial y otros acontecimientos durante 2011 en Medicina Intensiva. *Med Intens*. 2012;36:1–2.
 40. Jiménez-Contreras E, Torres-Salinas D, Ruiz-Pérez R, Delgado-López-Cózar E. Investigación de excelencia en España: ¿actores protagonistas o papeles secundarios? *Med Clin (Barc)*. 2010;134:76–81.