

A Bibliometria e a Avaliação da Produção Científica: indicadores e ferramentas

Teresa Costa

Fundação para a Computação
Científica Nacional
Inst. Investigação e Formação
Avançada, Universidade de Évora,
FCCN, Av. do Brasil, n.º 101
1700-066 Lisboa
Tel: 300005184
E-mail: teresa.costa@fccn.pt

*Silvia Lopes, Fernando
Fernández-Llimós*

Dep^{to}. Sócio-Farmácia, Faculdade de
Farmácia, Universidade de Lisboa
Av. Prof. Gama Pinto
1640-003 Lisboa
Tel: 217946484
E-mail: {slopes, f-llimos}@ff.ul.pt

*Maria João Amante,
Pedro Faria Lopes*

ISCTE - Instituto Universitário de
Lisboa
Av. Forças Armadas
1649-026 Lisboa
Tel: 217903082/
Email: {maria.amante, pedro.lopes}
@iscte.pt

RESUMO

A bibliometria é uma técnica quantitativa e estatística para medir índices de produção e disseminação do conhecimento, bem como acompanhar o desenvolvimento de diversas áreas científicas e os padrões de autoria, publicação e uso dos resultados de investigação. A avaliação da produção científica, importante para o reconhecimento dos investigadores junto da comunidade científica, é feita através da aplicação de diversos indicadores bibliométricos, que se dividem em indicadores de qualidade, importância e impacto científicos. As limitações e controvérsias apontadas por vários autores vêm colocar dúvidas acerca da validade da aplicação dos mesmos. Como resultado das críticas levantadas sobre os principais indicadores bibliométricos, nos últimos anos têm surgido novos indicadores. Assim, numa primeira fase deste estudo, pretende-se caracterizar e comparar os principais indicadores bibliométricos tendo em atenção as suas vantagens e limitações bem como campo de aplicação. Atualmente são várias as bases de dados que utilizam indicadores bibliométricos e disponibilizam resultados de análise bibliométrica, sendo a mais conhecida a Web of Science (WoS) da Thomson Reuters. Até recentemente, a WoS foi a principal ferramenta utilizada para a realização de análise de citações. Esta situação, no entanto, já não se verifica, uma vez se existem atualmente outras ferramentas que também fornecem dados com base nas citações. No presente estudo destacamos duas delas: a Scopus da Elsevier e o Google Scholar Metrics (GSM) sendo feita uma análise e comparação ao nível das características, vantagens e desvantagens destas três bases de dados com vista a perceber se são concorrentes ou complementares.

PALAVRAS-CHAVE: Bibliometria, Avaliação da produção científica, Indicadores e ferramentas bibliométricos

ABSTRACT

The bibliometry is a technique for measuring quantitative and statistical indices of production and dissemination of knowledge, and monitor the

development of several scientific areas and patterns of authorship, publication and use of research results. An evaluation of scientific production, important for the recognition of researchers from the scientific community, is made by applying several bibliometric indicators, which are divided into quality indicators, scientific importance and impact. The limitations and controversies mentioned by several authors have put doubts about the validity of their application. As a result of the criticisms raised about the main bibliometric indicators in recent years new indicators have emerged. Thus, at the first stage of this study, we intend to characterize and compare the main bibliometric indicators taking into account their advantages and limitations as well as scope. Currently there are various databases that use bibliometric indicators and provide results of bibliometric analysis, being the most well known the Web of Science (WoS) of Thomson Reuters. Until recently, WoS was the main tool used to conduct citation analysis. This situation, however, no longer applies, since there are currently other tools that also provide data based on citations. In this study we emphasize two of them: the Elsevier's Scopus and Google Scholar Metrics (GSM) with an analysis and comparison in terms of the characteristics, advantages and disadvantages of these three databases with the aim of understand whether they are competing or complementary.

KEYWORDS: Bibliometrics, Scientific evaluation, bibliometric indicators and tools

INTRODUÇÃO

A produção científica reveste-se da maior importância no conjunto das actividades académicas e de investigação sendo um instrumento pelo qual a comunidade científica mostra os resultados, a pertinência e a relevância da investigação. Nesse sentido, é o espelho do desempenho da instituição e dos docentes e investigadores, no conjunto das suas actividades de ensino e de investigação.

Como conclusão de um estudo, HALLIDAY (2001) considera que o resultado das actividades de investigação deve cumprir com os seguintes critérios:

ser credível, ser divulgado e estar acessível. Segundo este autor para que a produção científica tenha credibilidade é altamente desejável que uma vez publicada, não seja modificada, que diferentes versões estejam claramente identificadas e que a credibilidade se baseie em critérios de avaliação reconhecidos, como o *peer-review* (arbitragem científica). Também é conveniente que cada publicação tenha pelo menos um autor identificável.

Relativamente à divulgação, é aconselhável que as publicações tenham metadados e é altamente desejável que o público-alvo tenha conhecimento da sua existência. Para cumprir os critérios de acessibilidade, é essencial que a publicação tenha um formato de registo durável e seguro, que seja recuperável ao longo do tempo, que exista o compromisso de não retirar a publicação, que esta esteja acessível publicamente, independentemente do acesso ser gratuito ou de implicar algum pagamento. Finalmente, será altamente desejável que a publicação tenha identificadores persistentes (HALLIDAY, 2001).

Importa, portanto, compreender o que é considerado “produção científica”. Assim, os tipos de produção mais relevantes são (CORTEZ, 2011, 3):

- Livros
- Teses
- Capítulos de livros
- Artigos publicados em revistas científicas
- Comunicações em atas de conferências
- Relatórios técnicos, materiais pedagógicos, white papers e páginas web

E podem:

- Ter ou não arbitragem científica (*peer-review*)
- Ser de âmbito nacional ou internacional

Uma vez identificado o que se considera produção científica, torna-se importante compreender de que forma pode ser avaliada.

BIBLIOMETRIA

Surge pela primeira vez em 1917 quando Cole e Eales publicam uma análise estatística da história da disciplina de anatomia comparada (SANCHO, 2002, 80; OKUBO, 1997, 10; VANTI, 2002, 153), o conceito de bibliografia estatística, actualmente designada por bibliometria. Trata-se de uma técnica quantitativa e estatística que permite medir índices de produção e disseminação do conhecimento, acompanhar o desenvolvimento de diversas áreas científicas e os padrões de autoria, publicação e uso dos resultados de investigação (OKUBO, 1997, 8; ARAÚJO, 2006, 12). Inicialmente utilizada na medição estatística de monografias (exemplares, edições, etc.) com especial utilidade nas bibliotecas, rapidamente a sua aplicação passou a contemplar também a análise de outra documentação como periódicos e artigos científicos. Actualmente, é utilizada para avaliar a produtividade dos autores e realizar estudos de citações (ARAÚJO, 2006, 13). Desta forma, facilmente se percebe a sua importância e benefícios práticos (SANCHO, 2002; ARAÚJO, 2006, 13):

- nas bibliotecas – no desenvolvimento das colecções e na gestão da informação
- na comunidade académica e científica - na análise e avaliação da sua produção.

A avaliação da produção científica, factor primordial para o reconhecimento dos investigadores junto da comunidade científica, nacional e internacional, e das agências financiadoras, faz-se através da aplicação de diversos indicadores bibliométricos (SANCHO, 2002).

INDICADORES BIBLIOMÉTRICOS

Os indicadores bibliométricos são ferramentas de avaliação e podem ser divididos em:

- Indicadores de qualidade científica
- Indicadores de actividade científica
- Indicadores de impacto científico
- Indicadores de associações temáticas

Os indicadores de qualidade científica baseiam-se na percepção ou opinião dos pares que avaliam as publicações pelos seus conteúdos. Indicadores de actividade científica permitem contabilizar a actividade científica desenvolvida, nomeadamente o nº e distribuição dos trabalhos publicados, a produtividade dos autores, a colaboração na autoria dos trabalhos, o nº e distribuição das referências entre trabalhos e autores, entre outros. Os indicadores de impacto científico podem subdividir-se em dois tipos: indicadores de impacto dos trabalhos e indicadores de impacto das fontes. Como exemplo de indicador de impacto dos trabalhos temos o nº de citações recebidas. Por outro lado, como indicadores de impacto das fontes temos o factor de impacto das revistas, o índice de citação imediata e a influência das revistas. Associações temáticas também podem ser importantes. A análise de citações e a análise de referências comuns são exemplos de indicadores de associações temáticas (SANCHO, 2002, 83).

É importante realçar que todos os indicadores deverão ser aplicados de forma cuidada dado o conjunto de características e limitações que acarretam. Por outro lado, é fundamental considerar que os principais resultados de investigação se transmitem através da comunicação formal, que os trabalhos publicados representam o volume de investigação produzida e que estes trabalhos (fontes primárias) são indexados em bases de dados (fontes secundárias), ferramentas que permitem recuperar informação publicada em qualquer área científica.

Factor de Impacto

Trata-se de um indicador de avaliação do impacto de revistas, que determina a frequência com que um artigo é citado. Foi criado por Eugene Garfield em 1955 para classificar e avaliar as revistas incluídas na Science Citation Index, do Institute for Scientific Information (ISI).

Na prática, o factor de impacto (FI) contabiliza as citações efectuadas, num dado ano, a documentos publicados nos 2 anos anteriores.

Para uma dada revista, a fórmula de cálculo do FI é:

$$FI_n = \frac{\text{citações}_{n-2} + \text{citações}_{n-1}}{\text{artigos}_{n-2} + \text{artigos}_{n-1}}$$

Assim, o FI dessa revista no ano 2011 é:

$$FI_{2011} = \frac{\text{citações}_{2009} + \text{citações}_{2010}}{\text{artigos}_{2009} + \text{artigos}_{2010}}$$

Ou seja, o FI da dada revista em 2011, contabiliza o nº de citações ocorridas em 2011, a documentos publicados em 2009 e 2010.

No entanto, existem diversas e importantes limitações apontadas ao FI que condicionam a credibilidade do mesmo. A primeira grande limitação relaciona-se com a fórmula de cálculo. Assim, vejamos as diferenças entre o que é contabilizado no numerador e o que é tido em conta no denominador. No numerador entram todos os documentos citáveis, isto é, tudo o que é publicado: artigos originais, artigos de revisão, relatórios, editoriais, cartas, resumos, etc. enquanto no denominador apenas são contabilizados os “*meaty items*”: artigos originais, artigos de revisão e relatórios (GARFIELD, 1986, 445; FERNANDEZ-LLIMÓS, 2003, 43; McVEIGH et al., 2009, 1107).

Como resultado desta limitação na fórmula de cálculo, podemos verificar que revistas com uma maior

correspondem necessariamente ao FI da revista onde o artigo é publicado.

H-index

Criado por J. Hirsch em 2005, o h-index (h_i) é uma proposta para quantificar a produtividade e o impacto dos investigadores, baseando-se nos seus artigos mais citados. Corresponde ao número de artigos de um determinado autor com, pelo menos, o mesmo número de citações (COSTAS et al., 2007, 193). Ou seja, se o h_i de um investigador for 16, quer dizer que, da totalidade de artigos publicados, esse investigador, tem 16 artigos com, cada um deles, pelo menos 16 citações.

O cálculo do h_i é efectuado a partir da lista de artigos do investigador. Esta lista deverá ser apresentada por ordem decrescente do nº de citações recebidas. Quando o nº de ordem na lista iguala o nº de citações recebidas, obtém-se o h-index (Fig.1).

| | | | | | | | | |
|--------------------------|--|---|---|---|---|---|----|------|
| <input type="checkbox"/> | 26. Title: GENETIC STUDIES OF MACUSHI AND WAPISHANA INDIANS. 2. DATA ON 12 GENETIC POLYMORPHISMS OF RED-CELL AND SERUM-PROTEINS - GENE FLOW BETWEEN TRIBES Author(s): NEEL, J.V, GERSHOWITZ H, SPIELMAN RS, et al. Source: HUMAN GENETICS Volume: 37 Issue: 2 Pages: 207-219 Published: 1977 | 0 | 0 | 1 | 0 | 0 | 32 | 1.00 |
| <input type="checkbox"/> | 27. Title: EFFECT OF SEX, AGE AND CULTIVATION TIME ON NUMBER OF SATELLITES AND ACROCENTRIC ASSOCIATIONS IN MAN Author(s): MATTEVI MS, SALZANO FM Source: HUMANGENETIK Volume: 29 Issue: 4 Pages: 265-270 Published: 1975 | 0 | 0 | 0 | 0 | 0 | 32 | 0.94 |
| <input type="checkbox"/> | 28. Title: ABNORMAL HEMOGLOBINS IN A BRAZILIAN NEGRO POPULATION Author(s): TONDO CV, SALZANO FM Source: AMERICAN JOURNAL OF HUMAN GENETICS Volume: 14 Issue: 4 Pages: 401-& Published: 1962 | 0 | 0 | 0 | 0 | 0 | 32 | 0.68 |
| <input type="checkbox"/> | 29. Title: Further evidence for the existence of a major founder Y-chromosome haplotype in Amerindians Author(s): Santos FR, Hutz MH, Coimbra CEA, et al. Source: BRAZILIAN JOURNAL OF GENETICS Volume: 18 Issue: 4 Pages: 669-672 Published: DEC 1995 | 1 | 1 | 1 | 1 | 0 | 30 | 2.31 |
| <input type="checkbox"/> | 30. Title: SKIN COLOUR OF CAINGANG AND GUARANI INDIANS OF BRAZIL Author(s): HARRISON GA, SALZANO FM Source: HUMAN BIOLOGY Volume: 38 Issue: 2 Pages: 104-& Published: 1966 | 0 | 0 | 1 | 0 | 0 | 30 | 0.70 |

Figura 1: Exemplo do cálculo do h-index

percentagem de editoriais, cartas e resumos publicados conseguem obter valores de FI superiores.

Por outro lado, existem diferenças no nº de citações por artigo entre diferentes áreas temáticas. Por exemplo, a média de citações por artigo na bioquímica é de 1,12 enquanto na botânica é de 0,43. Há ainda a considerar que existem diferenças na obsolescência dos trabalhos entre áreas do conhecimento distintas e, por essa razão, a vida média de citação (VMC) (mediana da antiguidade das referências) também varia entre áreas científicas. Assim, é incorrecto efectuar comparações entre diferentes áreas já que diferentes padrões de citação não têm a ver com qualidade mas sim com o enquadramento/dinâmica de cada área temática.

O FI tende, portanto, a beneficiar periódicos das áreas científicas com ritmo mais acelerado de renovação do conhecimento, ao contabilizar apenas as citações dos 2 anos seguintes à sua publicação.

O facto de o FI ser calculado apenas a partir dos conteúdos da *Science Citation Index* é outra importante desvantagem. Existem limitações quer idiomáticas quer geográficas na inclusão de periódicos nesta fonte secundária (FERNANDEZ-LLIMÓS, 2003, 49), com predomínio de revistas de língua inglesa (de origem americana) (MARZIALE et al., 2002, 466) e a ausência de revistas de carácter regional, nomeadamente em áreas como Saúde Pública, Botânica, Geologia ou Astronomia.

Finalmente é de registar a utilização incorrecta do FI como indicador de impacto dos trabalhos, uma vez que este indicador pretende medir o impacto das revistas. Os trabalhos mais citados nem sempre são publicados em revistas com FI e as citações dos artigos não

Existem diversas vantagens associadas ao h-index, nomeadamente o facto de combinar medidas de quantidade (publicações) e de impacto (citações), e de poder ser aplicado a grupos de investigação, departamentos e/ou países. Permite ainda criar “rankings” de Universidades, analisando e comparando universidades produtivas e de baixo impacto *versus* universidades menos produtivas e maior impacto.

No entanto, também é alvo de críticas na avaliação dos investigadores, uma vez que não entra em linha de conta com alguns factores importantes como a longevidade da carreira bem como a estratégia de publicação do investigador e as diferenças existentes entre áreas científicas (COSTAS et al., 2007, 194). Utilizando o h_i para avaliar investigadores em início de carreira, com baixa produção e/ou com uma estratégia de publicação mais selectiva podemos estar a incorrer em alguns erros (Fig.2).

| Investigador | Nº Artigos | Nº Citações | H-index |
|--------------|------------|-------------|---------|
| Inv. 1 | 1 | 500 | 1 |
| Inv. 2 | 5 | 5 | 5 |
| Inv. 3 | 30 | 1 | 1 |

Figura 2: Exemplo de comparação entre 3 investigadores através do h-index

De modo a colmatar as falhas associadas ao h_i , Hirsch apresentou o “Parâmetro m” para que se possam comparar investigadores de diferentes áreas científicas e com distintos tempos de carreira:

$$h = \frac{h_i}{m}$$

em que: h_i = h-index; m = nº de anos desde a 1ª publicação do autor.

Apesar das melhorias introduzidas, mantêm-se algumas críticas. Documentos altamente citados têm importância relativa: entram para a contagem, mas o nº de citações é irrelevante. As auto-citações dos próprios podem enviesar os resultados, introduzindo um risco de aumento desmesurado do valor do índice em benefício dos próprios. À semelhança de outros indicadores, existem limitações técnicas como a dificuldade de identificar todas as publicações dos autores (devido às variantes dos nomes e ao facto de ser calculado apenas a partir do que está indexado nas bases de dados).

No entanto, o h-index é considerado um bom indicador bibliométrico, preferível a usar apenas o nº total de trabalhos produzidos, o nº total de citações, o nº citações por artigo ou o nº de artigos mais citados (COSTAS et al., 2007, 194).

Eigenfactor™ Metrics

Eigenfactor™ Metrics são medidas de prestígio e credibilidade das revistas que utilizam o Eigenfactor™ score (EF) e o Article Influence™ score (AI) (Eigenfactor, 2012).

O EF mede o grau de importância de revistas científicas calculando o número de citações recebidas por essas revistas a partir de outras publicações. Baseia-se no número de vezes que artigos publicados numa dada revista, nos cinco anos anteriores, são citados no ano mais recente e exclui as auto-citações das revistas.

O AI determina a influência média de um artigo publicado em determinada revista, nos primeiros 5 anos após ter sido publicado. O cálculo de AI é baseado no EF:

$$AI = \frac{\text{Eigenfactor } r}{T_a}$$

em que T_a – Total de artigos publicados.

O EI considera “a qualidade das revistas citadas ponderando as suas citações através do seu impacto na comunidade científica (...) assumindo que a literatura científica constitui uma vasta rede de artigos ligados uns aos outros através das citações (...) e usa a estrutura da rede para medir o impacto relativo das revistas” (DURIEUX et al., 2010, 347). Franceschet (2010) identifica um conjunto de vantagens no uso deste indicador, nomeadamente, o facto de ponderar citações tendo em conta a importância das revistas citadas, considerar a quantidade de referências das revistas citadas, ter em consideração uma janela temporal de 5 anos, explorar a totalidade da rede de citações, ignorar as auto-citações das revistas, estar fortemente relacionada com outras métricas, ser um indicador que pode ser aplicado em diversos contextos e estar disponível gratuitamente num conjunto significativo de fontes secundárias (FRANCESCHET, 2010, 556). No entanto, Durieux e Gevenois (2010) consideram que embora o EI se baseie “numa ideia muito promissora, o seu papel potencial na avaliação das revistas científicas não ter sido ainda suficientemente estudado” (DURIEUX et al., 2010, 347). Serrano (2004) salienta

que têm sido identificadas algumas críticas no que se refere à robustez do algoritmo utilizado (idêntico ao utilizado pelo Google na ordenação das páginas web) quando em causa estão valores mais baixos de nº de citações (VILLAR, 2011, 94).

Relativamente ao AI, VILLAR (2011) considera que este é um excelente indicador a ter em conta uma vez que evita comparar revistas que publicam dezenas de artigos por ano com as revistas que publicam centenas. Assim, alguns autores consideram que este é o método alternativo à utilização do FI (VILLAR, 2011, 90).

Para além de todas as características já identificadas, pode ser ainda referido que o EF (e por conseguinte o AI) gera uma classificação temática endógena e unívoca em 87 categorias, permitindo que cada periódico pertença a apenas uma categoria, usa a mesma fonte do FI – o Journal Citation Reports (JCR), e está acessível gratuitamente na Web (VILLAR, 2011, 93).

SCImago Journal Rank (SJR) indicator

O indicador SJR é um recurso em acesso aberto e que lista um conjunto considerável de títulos de revistas de uma grande variedade de países e idiomas já que se baseia na base de dados Scopus (FALAGAS et al., 2008, 2623). Mede o prestígio de uma revista, através das referências que um periódico faz a outros ou a si próprio. No entanto, de forma a prevenir o uso excessivo de auto-citações da revista, destas apenas são consideradas no máximo de 33% do total de citações (GONZÁLEZ-PEREIRA et al., 2009, 5). Utiliza um algoritmo semelhante ao Google PageRank e expressa o número médio de citações ponderadas recebidas num determinado ano, através de documentos publicados na revista nos três anos anteriores (SCImago, 2007; FALAGAS et al., 2008, 2623; GONZÁLEZ-PEREIRA et al., 2009, 5). Ao contrário do que acontece com o cálculo do FI, o indicador SJR considera no denominador da sua fórmula de cálculo, o total de documentos publicados no periódico. Portanto, o indicador SJR apresenta-se como uma boa alternativa ao FI sendo que, para além das características já mencionadas, também se destaca pela avaliação da qualidade das citações (FALAGAS et al., 2008, 2623).

FERRAMENTAS BIBLIOMETRICAS

São pois várias as métricas e os indicadores utilizados, estando associados a diversas bases de dados que são ferramentas muito úteis para o desenvolvimento de estudos bibliométricos. Podem fornecer pistas sobre a investigação de um tema, pesquisando um tópico, sobre as mudanças de orientação que o mesmo possa ter sofrido ao longo do tempo, a evolução de um autor, etc. Assim, e sendo usadas como meios, não como fins, podem fornecer aos investigadores uma preciosa ajuda aquando dos seus trabalhos de investigação.

Atualmente são três as ferramentas bibliométricas mais usadas: a Web of Science, a Scopus e o recente Google Scholar Metrics. Estas cobrem uma gama de diferentes dados e usam diferentes métricas, pelo que tal deve ser sempre mantido em mente aquando da sua utilização.

Web of Science

A Web of Science (WoS) da Thomson Reuters é uma das bases de dados mais importante ao nível das revistas científicas e foi durante mais de quarenta anos a maior

base de dados bibliométrica (ARCHAMBAULT, 2009, 1320). É multidisciplinar e contém o Science Citation Index, o Social Sciences Citation Index e o Arts & Humanities Citation Index. Indexa mais de 12.700 periódicos, nas diferentes áreas científicas, contendo informação desde início do século XX, sendo atualizada semanalmente.

É a partir destas bases de dados que é calculado o fator de impacto das publicações periódicas, bem como outros indicadores bibliométricos presentes no JCR. Este é a ferramenta de classificação de periódicos, foi desenvolvido na década de 1950, e é a ferramenta líder de mercado dos rankings de periódicos.

É um recurso essencial para a avaliação e comparação de periódicos, recolhendo dados referentes a mais de 9.000 periódicos científicos a nível mundial.

O JCR permite conhecer dados bibliométricos dos diferentes periódicos e compará-los dentro de uma mesma área científica.

O FI é a medida mais comumente utilizada, porém convém relembrar que este é um indicador de qualidade dos periódicos que permite, ainda que com alguns cuidados, comparar/avaliar revistas dentro de uma mesma área, não devendo ser aplicado aos investigadores.

A WoS foi projetada desde início como uma base de dados de citação (índice de citações). Permite, além da pesquisa habitual por ocorrência de palavras no registo, também a pesquisa de artigos relacionados e o estabelecimento de ligações entre artigos que citam outros ou são citados por outros.

Como principais vantagens desta base de dados podemos indicar:

- excelente cobertura temporal (a partir de 1900-presente para algumas revistas)
- um grande número de registos é enriquecido com as referências citadas
- melhorou a cobertura regional e muito recentemente adicionou 700 periódicos regionais
- primeira base de dados a incorporar o h-índice
- pode ver-se h-index negativo (apenas do primeiro autor)
- inclui conference proceedings
- inclui monografias
- pode ver registos órfãos usando a opção de pesquisa “cited references”

Como principais desvantagens tem:

- melhor cobertura na área das ciências do que das artes e humanidades
- c. 80% dos conteúdos são da áreas das Ciências
- pouca cobertura de conteúdos em Acesso Aberto
- as opções/alternativas para encontrar e distinguir autores não são grandes
- falta de standardização de autores e instituições/afiliação
- privilegia os conteúdos anglo-saxónicos

A WoS manteve o seu domínio enquanto base de indicadores bibliométricos até 2004, ano em que a editora Elsevier lançou a Scopus.

Scopus

Contrariamente ao que acontece com a WoS, a Scopus não foi projetada como um índice de citações, mas inclui citações de artigos desde 1996. A sua principal intenção é a pesquisa por autor e assunto. Procura ter uma cobertura detalhada desde 1996 até à atualidade. Conteúdos de outras bases de dados da Elsevier, desde 1966, têm sido seletivamente incluídos na Scopus, de modo a aumentar e melhorar a cobertura.

A equipa da Scopus tem melhorado o seu Journal Analyzer tendo acordado com a SCImago que o cálculo do SJR esteja incluído na base de dados.

Como principais vantagens podemos indicar:

- indexa mais de 18.000 títulos de periódicos
- inclui títulos em Acesso Aberto, conferências, páginas web, patentes e livros
- a funcionalidade “more” permite visualizar rapidamente os registos órfãos
- cobertura muito forte ao nível das revistas de ciência e tecnologia
- contém ferramentas úteis para identificação dos autores
- gera automaticamente o h-índice
- tem mais conteúdos europeus que a WoS, e inclui mais idiomas para além do Inglês - 60% de cobertura é de fora dos EUA

Ao nível das desvantagens podem indicar-se:

- a cobertura temporal não é muito significativa, uma vez que no caso de muitas revistas só indexa os últimos 5 anos
- cobertura deficiente das áreas das artes e humanidades, mas maior que a da WoS
- a maior parte das citações remonta a 1996. Tal resulta num enviesamento do h-index para investigados com carreiras mais longas
- citações de artigos pré-1996 feitas em artigos publicados depois de 1996 não estão incluídas no cálculo no h-index, não sendo contabilizadas, o que limita de duas formas – citação e citado.

Relativamente à análise e ranking de periódicos fornece duas métricas principais: o SJR e o Source Normalised Impact Per Paper (SNIP) que podem dar resultados bastante variados. Ao contrário do JCR, não fornece listas ordenadas de revistas por área temática. No entanto, o site SCImago, de utilização gratuita, baseia-se nos dados Scopus para fornecer este tipo de listagem (SCImago, 2007).

Google Scholar Metrics

Em abril passado, o Google lançou o Google Scholar Metrics (GSM), uma nova ferramenta que fornece métricas de impacto de revistas científicas, obtidas a partir da contagem de citações e que, por ser de acesso gratuito, se apresenta como uma alternativa muito apetecível.

O GSM fornece acesso ao h-index de uma ampla gama de revistas científicas e outras fontes de informação mas não aplica a investigadores, contrariamente às duas bases de dados já referidas.

Vantagens:

- a principal vantagem desta ferramenta é para já o facto de ser gratuita, uma vez que vai estimular a competição entre os vários fornecedores
- como também adotou como indicador o h-index, pode encorajar a adoção deste indicador na e para a avaliação de periódicos

São muitas as desvantagens apontadas a este novo produto do Google, que não inclui todos os periódicos indexados no Google Scholar.

Segundo Delgado-López-Cózar et al. (2012) “It is disappointing to see how Google has delivered a product that is so unambitious and full of mistakes. Google should be aware that producing professional bibliometric tools requires effort and infrastructure beyond algorithms and robots that automatically produce results.” (DELGADO-LÓPEZ-CÓZAR, et al., 2012, 425).

Assim, como desvantagens temos:

- cobertura temporal limitada a artigos publicados nos últimos cinco anos (2007-2011), inclui as citações recebidas até Abril de 2012.
- junta periódicos científicos com outras fontes como repositórios, bases de dados, proceedings e working papers
- não fornece uma lista de periódicos cobertos (peer-reviewed ou não)
- falta de standardização dos títulos dos periódicos
- apresenta erros na identificação de autores, periódicos e outros dados bibliográficos.~
- desconhece-se a periodicidade com que faz atualização dos dados

O uso de distintas bases de dados pode conduzir a diferentes resultados de uma mesma publicação e/ou investigador. São vários os aspetos a que deve estar atento.

Veja-se, por exemplo, o caso da revista *Scientometrics*, no que se refere aos diversos indicadores (Fig. 3).

| Indicador | Valor |
|--------------|-------|
| JCR IF | 1.966 |
| SCOPUS SRJ | 0.066 |
| SCOPUS SNIP | 1.418 |
| GSM H5-INDEX | 38 |

Figura 3: Cálculo de indicadores para a revista *Scientometrics*

Também no caso do estudo de citações de autores os valores podem divergir. Veja-se o exemplo apresentado na Fig.4, para o autor Andras Schubert.

| | h-index |
|--------|---------|
| WoS | 28 |
| SCOPUS | 22 |
| GSM | n.a |

Figura 4: Cálculo do h-index para autor Andras Schubert

CONSIDERAÇÕES FINAIS

A bibliometria e os indicadores bibliométricos, no seu conjunto, são utilizados para avaliar a produtividade dos autores, realizar estudos de citações e avaliar a qualidade dos periódicos científicos.

Apesar da quantidade de indicadores existentes e das inúmeras críticas que surgem relativamente à sua utilização estes evidenciam benefícios práticos na análise e avaliação da produção da comunidade académica e científica.

Nos últimos anos têm surgido novas ferramentas e novos indicadores o que ilustra que, apesar das críticas apresentadas, estes acabam por estar enraizados na avaliação da comunicação científica. Estas ferramentas e indicadores têm surgido com vista a avaliar não só as publicações mas também os investigadores, pelo que é importante ter alguns cuidados a quando da sua utilização.

Como exemplo, temos o h-index que é utilizado e disponibilizado nas diversas bases de dados. Porém, e apesar de ser comum, pode divergir quanto aos resultados ou quanto ao âmbito de aplicação já que, no caso do GSM, é aplicado às publicações.

A avaliação destas conta também com diversos indicadores distintos nos parâmetros utilizados o que tem consequências no resultado final.

A utilização das bases de dados bibliométricas é hoje uma realidade. No entanto há que ter em consideração que as mesmas não só divergem nos conteúdos que indexam, ao nível de tipologia de documento (artigos, periódicos, patentes, livros, etc.), como também ao nível das coberturas temática, temporal, geográfica e idiomática. Há um predomínio da área das ciências e tecnologias com um considerável défice das artes e humanidades, nas bases de dados mais recentes a cobertura temporal é significativamente reduzida e abundam os conteúdos anglo-saxónicos com predomínio de conteúdos norte americanos. Podemos então concluir que mais do que concorrentes as diferentes bases de dados são complementares. É de salientar que, no caso das três bases de dados analisadas, duas são de subscrição. Porém, de referir que associado às mesmas existe um conjunto de ferramentas e indicadores disponíveis gratuitamente, como o SJR e o Eigenfactor. No caso do GSM, apesar de gratuito, sendo ainda muito recente, a sua qualidade/aplicação é reduzida. Todavia é de salientar a iniciativa do Google no desenvolvimento e disponibilização desta ferramenta em acesso aberto.

REFERÊNCIAS

ARAÚJO, Carlos Alberto – Bibliometria: evolução histórica e questões atuais. Em *Questão*. Porto Alegre. [Em linha]. Vol.12, nº 1 (2006), p.11-32. [Consult. 21 de Abril 2012]. Disponível em [www: http://revistas.univerciencia.org/index.php/revistaemquestao/article/viewFile/3707/3495](http://revistas.univerciencia.org/index.php/revistaemquestao/article/viewFile/3707/3495)

ARCHAMBAULT, Éric et al. – Comparing Bibliometric Statistics Obtained From the Web of Science and Scopus. *Journal of the American Society for Information Science and Technology*. [Em linha]. Vol. 60, nº 7 (2009), p.1320-1326. [Consult. 8 Setembro 2012]. Disponível em [www: http://onlinelibrary.wiley.com/doi/10.1002/asi.21062/pdf](http://onlinelibrary.wiley.com/doi/10.1002/asi.21062/pdf)

CORTEZ, Paulo – Some scholarly communication guidelines: teaching report. Guimarães: Department of

- Information Systems of University of Minho, 2011. [Consult. 24 de Abril 2012]. Disponível em [www: http://hdl.handle.net/1822/11599](http://hdl.handle.net/1822/11599)
- COSTAS, Rodrigo; BORDONS, María – The h-index: advantages, limitations and its relation with other bibliometric indicators at the micro level. *Journal of Informetrics* [Em linha]. Vol. 1 (2007), p. 193-203. [Consult. 21 de Abril 2012]. Disponível em [www: http://www.liquidpub.org/mediawiki/upload/1/11/Journal_of_Informetrics_1\(3\).pdf](http://www.liquidpub.org/mediawiki/upload/1/11/Journal_of_Informetrics_1(3).pdf)
- DELGADO-LÓPEZ-CÓZAR, Emilio; CABEZAS-CLAVIJO, Álvaro – Google Scholar Metrics: na unreliable tool for assessing scientific journals. *El Profesional de la información*. [Em linha]. Vol.21, n.º4 (2012), p. 419-425. [Consult. 27 Agosto 2012]. Disponível em [www: http://elprofesionaldeinformacion.metapress.com/media/2gurm834tjdxqneejmfk/contributions/q/0/3/g/q03gk6v6u5875872.pdf](http://elprofesionaldeinformacion.metapress.com/media/2gurm834tjdxqneejmfk/contributions/q/0/3/g/q03gk6v6u5875872.pdf)
- DURIEUX, Valérie; GEVENOIS, Pierre Alain – Bibliometric indicators: quality measurements of scientific publication. *Radiology*. [Em linha]. Vol. 255, n.º 2 (2010), p. 342-351. [Consult. 21 de Abril 2012]. Disponível em [www: http://radiology.rsna.org/content/255/2/342.full.pdf+html](http://radiology.rsna.org/content/255/2/342.full.pdf+html)
- Eigenfactor.org: ranking and mapping scientific knowledge [Em linha]. Washington: University of Washington, 2012. [Consult. 21 de Abril 2012]. Disponível em [www: http://www.eigenfactor.org](http://www.eigenfactor.org)
- FALAGAS, Matthew E. et al. – Comparison of SCImago journal rank indicator with journal impact factor. *The FASEB Journal* [Em linha]. Vol. 22, n.º 8 (2008), p. 2623-2628. [Consult. 21 de Abril 2012]. Disponível em [www: http://www.fasebj.org/content/22/8/2623.full.pdf+html](http://www.fasebj.org/content/22/8/2623.full.pdf+html)
- FERNANDÉZ-LLIMÓS, Fernando – Análisis de la cobertura del concepto de Pharmaceutical Care en fuentes primarias y secundarias de información. Granada: Universidad de Granada, 2003. Tesis de doutorado apresentado à Universidad de Granada.
- FRANCESCHET, Massimo – Ten good reasons to use the Eigenfactor™ metrics. *Information Processing and Management*. [Em linha]. Vol. 46, n.º 5 (2010), p.555-558. [Consult. 21 de Abril 2012]. Disponível em [www: http://www.sciencedirect.com/science/article/pii/S0306457310000075](http://www.sciencedirect.com/science/article/pii/S0306457310000075)
- GARFIELD, Eugene – Journal Impact vs Influence: A need for five-year impact factors. *Information Processing & Management*. Vol. 22, n.º5 (1986), p. 445.
- GONZÁLEZ-PEREIRA, Borja; GUERRERO-BOTE, Vicente P.; MOYA-ANEGÓN, Félix – The SJR indicator: A new indicator of journals' scientific prestige. [Em linha] [Consult. 21 de Abril 2012]. Disponível em [www: http://arxiv.org/ftp/arxiv/papers/0912/0912.4141.pdf](http://arxiv.org/ftp/arxiv/papers/0912/0912.4141.pdf)
- HALLIDAY, Leah L. – Scholarly communication, scholarly publication and status of emerging formats. *Information Research* [Em linha]. Vol. 6, n.º 4 (2001). [Consult. 24 de Abril 2012]. Disponível em [www: http://InformationR.net/paper111.html](http://InformationR.net/paper111.html)
- MARZIALE, Maria Helena P.; MENDES, Isabel Amélia C. – O fator de impacto das publicações científicas. *Revista Latino-Americana de Enfermagem*. Ribeirão Preto. [Em linha]. Vol. 10, n.º 4 (2002), p.466-471. [Consult. 21 de Abril 2012]. Disponível em [www: http://www.scielo.br/pdf/rlae/v10n4/13356.pdf](http://www.scielo.br/pdf/rlae/v10n4/13356.pdf)
- McVEIGH, Marie E; MANN, Stephen J. – The Journal Impact Factor Denominator: Defining Citable (Counted) Items. *Journal of the American Medical Association*. Vol. 302, n.º 10 (2009), p.1107-1109.
- OKUBO, Yoshiko – Bibliometric indicators and analysis of research systems: methods and examples. *OECD Science, Technology and Industry Working Papers*. [Em linha]. N.º 1 (1997). [Consult. 21 de Abril 2012]. Disponível em [www: http://dx.doi.org/10.1787/208277770603](http://dx.doi.org/10.1787/208277770603)
- SANCHO, Rosa – Indicadores bibliométricos utilizados en la evaluación de la ciencia y la tecnología: revisión bibliográfica. In *Inteligencia competitiva: documentos de lecture*. [Em linha]. Barcelona: Fundació per a la Universitat Oberta de Catalunya, 2002, p.77-106. [Consult. 21 de Abril 2012]. Disponível em [www: http://www.tramasoft.com/documentos/I+D+i/UND2/Lecturas%20complementarias/79059.Inteligencia%20Competitiva.Lecturas.pdf#page=77](http://www.tramasoft.com/documentos/I+D+i/UND2/Lecturas%20complementarias/79059.Inteligencia%20Competitiva.Lecturas.pdf#page=77)
- VANTI, Nadia Aurora Peres - Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento. *Ciência da Informação*. [Em linha]. Vol. 31, n.º 2 (2002), p.152-162. [Consult. 21 de Abril 2012]. Disponível em [www: http://www.scielo.br/pdf/ci/v31n2/12918.pdf](http://www.scielo.br/pdf/ci/v31n2/12918.pdf)
- VILLAR, Antonio - El «eigenfactor»: un nuevo y potente instrumento bibliométrico para evaluar la investigación. *Aula Abierta* [Em linha]. Vol. 39, n.º 3 (2011), pp. 85-96. [Consult. 20 de Setembro 2012]. Disponível em [www: http://www.uniovi.net/ICE/publicaciones/Aula_Abierta/numerosanteriores/i17/10_AA_vol39_n3_sept_2011](http://www.uniovi.net/ICE/publicaciones/Aula_Abierta/numerosanteriores/i17/10_AA_vol39_n3_sept_2011)
- SCImago. SJR — SCImago Journal & Country Rank [Em linha]. Spain: SCImago, 2007. [Consult. 24 de Setembro 2012]. Disponível em [www: http://www.scimago.com](http://www.scimago.com)