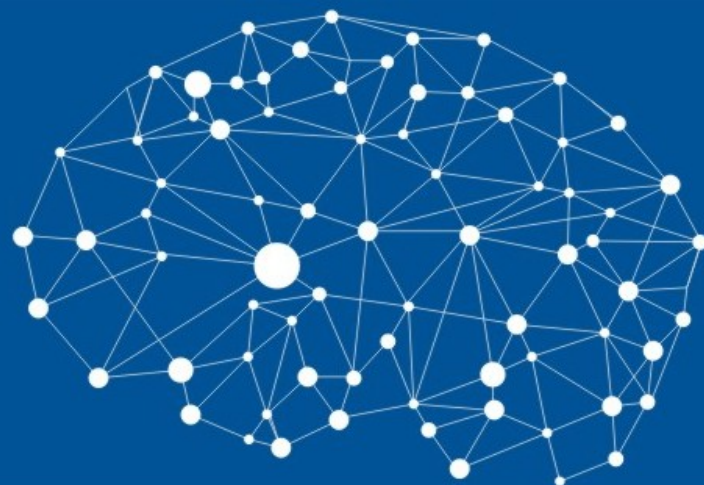




ÍNDICE H DE HIRSCH: análise comparativa entre as bases de dados Scopus, Web of Science e Google Acadêmico

Deise Deolindo Silva¹; Maria Cláudia Cabrini Grácio²

SILVA, D. D.; GRÁCIO, M. C. U. C.. ÍNDICE H DE HIRSCH: análise comparativa entre as bases de dados Scopus, Web of Science e Google Acadêmico In: ENCONTRO BRASILEIRO DE BIBLIOMETRIA E CIENTOMETRIA, 5., 2016, São Paulo. **Anais...** São Paulo: USP, 2016. p. A9

¹Universidade Estadual Paulista Unesp/Marília; ²Universidade Estadual Paulista Unesp/Marília

ÍNDICE *H* DE HIRSCH:

análise comparativa entre as bases de dados *Scopus*, *Web of Science* e Google Acadêmico

Eixo temático: Indicadores em CT&I

Modalidade: Apresentação oral

1 INTRODUÇÃO

A socialização do conhecimento produzido é um fenômeno que cresce ao longo das últimas décadas, principalmente com o advento das Tecnologias da Informação e Comunicação (TIC). Nota-se um crescimento significativo do número de resultados de pesquisas científicas publicados em algum meio, especialmente em bases de dados.

Observa-se, na literatura, uma crescente preocupação em se elaborar bases de dados confiáveis e mecanismos adequados de avaliação de domínios científicos, assim como de se dispor de instrumentos e indicadores para definição de suas diretrizes, alocação de investimentos e recursos, formulação de programas e avaliação de atividades relacionadas ao desenvolvimento científico e tecnológico, como forma de identificar as elites científicas, garantindo um investimento profícuo das agências de fomento à pesquisa. (MUGNAINI *et al.*, 2004; THOMAZ *et al.*, 2011; VANZ, STUMPF, 2010).

Neste sentido, as bases de dados darão suporte para a almejada visibilidade da produção científica, pois a partir dessas fontes de informação são encontrados resultados de pesquisas, de pesquisadores e Instituições, nacionais e internacionais. Dentre as plataformas de dados internacionais pode-se destacar a *Scopus* e a *Web of Science* (Wos). Elas são bases de artigos científicos, multidisciplinares que reúnem um amplo conjunto de revistas de diferentes editores. Além disso, possibilitam o acesso global a grande parte da literatura científica publicada internacionalmente; possuem ferramentas adequadas para a realização de uma revisão bibliográfica em um determinado domínio; são abrangentes e incluem contagem de citações por artigo e outros dados adicionais (SILVA, 2013).

A base *Scopus* foi lançada pela editora *Elsevier* em 2004. No ano de 2013 contava com aproximadamente 21.000 títulos de mais de 5.000 editoras internacionais. A atualização é feita diariamente, a cobertura é se dá desde 1823 e realizam-se as contagens de citações desde 1996 (SCOPUS, 2015; SILVA, 2013). A *Web of Science*, em 2013, contava com mais de 12.000 títulos de pelo menos 3300 editoras internacionais, os quais são atualizados semanalmente, a cobertura e as contagens de citações são realizadas desde 1900 (WEB OF SCIENCE, 2015; SILVA, 2013).

Em novembro de 2004 foi lançado o Google Acadêmico (GA) que, diferentemente da *Scopus* e *Wos*, é uma ferramenta gratuita de busca de publicações científicas, a qual permite localizar trabalhos acadêmicos de vários tipos (teses, dissertações, artigos de congressos e de periódicos tanto de acesso aberto como fechado) e em múltiplas línguas. Ele busca referências em qualquer documento que esteja disponível na *Web*. Essa base de dados não considera citações feitas por livros, embora as citações recebidas por livros sejam consideradas (CAREGNATO, 2011; BAR-ILAN, 2008; VIEIRA, WAINER, 2013).

Dada à importância de se ter publicações indexadas em plataformas de dados confiáveis e da preocupação em se avaliar a produção científica, faz-se necessário estudar o comportamento de tais publicações nos diversos domínios científicos. Dentre as abordagens para a caracterização e avaliação de um domínio, destacam-se os estudos bibliométricos, especialmente as análises baseadas nas citações, por constituir abordagem consolidada, objetiva e confiável (HJORLAND, 2002; OLIVEIRA, GRÁCIO, 2011).

Entre os indicadores de impacto presentes na literatura científica, destaca-se o índice *h*, um parâmetro avaliativo considerado robusto, por avaliar de forma simultânea, os aspectos relativos à produção (quantidade de artigos produzidos) e ao impacto (número de citações).

Diversas bases de dados apresentam o cálculo deste indicador, visto que ele é utilizado tanto na avaliação do comportamento da produção científica de pesquisadores, como também é um critério avaliativo em diversas agências de fomento. No entanto, este indicador pode ter valores discordantes se buscado em diferentes fontes de informação.

Considerando o exposto, esta pesquisa teve como objetivo comparar os valores do índice *h*, fornecidos pelas fontes de informação: *Scopus*, *Wos* e *Google Acadêmico*, com a finalidade de testar a consistência entre os valores obtidos por estas três fontes de dados.

Para isto, tomou-se como universo de aplicação 20 investigadores, considerados mais produtivos na temática “Estudos Métricos”, conforme estudo de Silva e Grácio (2015), sendo que eles deveriam ter pelo menos 20 artigos publicados, em um período de 35 anos (1980 a 2014).

Conforme Oliveira e Grácio (2011) os “Estudos Métricos” podem ser considerados como o conjunto de conhecimentos relacionados à avaliação da informação produzida. Referem-se a estudos de natureza teórico-conceitual quando contribuem para o avanço do conhecimento da própria temática, propondo novos conceitos e indicadores ou reflexões e análises relativas à área. São de natureza metodológica quando se propõem a dar sustentação aos trabalhos de caráter teórico da área onde estão aplicados. Entre seus subcampos, encontram-se: Bibliometria, Cientometria, Webometria, Patentometria, Altmatria e a Informetria, esta última com a maior amplitude e a primeira constituindo a origem dos Estudos Métricos.

Esta pesquisa se justifica, pois se a consistência dos indicadores for observada entre as plataformas analisadas, o pesquisador e/ou avaliador pode optar em buscar as informações em quaisquer bases de dados. Além disso, trará contribuições para a área dos Estudos Bibliométricos, principalmente no âmbito metodológico, visto que trabalhar com dados consistentes é um pré-requisito para se ter resultados significativos e, complementando, os estudos comparativos entre as diversas bases de dados são carentes na literatura científica.

2 PROCEDIMENTOS METODOLÓGICOS

Utilizando o mesmo recorte de Silva e Grácio (2015), tomou-se como universo de pesquisa os 20 autores mais produtivos, considerando a temática Estudos Métricos. Para compor a amostra os pesquisadores deveriam ter pelo menos 20 artigos científicos publicados em um período de 35 anos (1980-2014). Entre os dias 02 e 10 de novembro de 2015 foram coletadas, nas três bases informacionais, as informações do índice *h* para cada pesquisador.

Em seguida procedeu-se a análise descritiva dos dados e para analisar se havia diferença significativa entre os valores obtidos para o índice *h*, com dados obtidos via *Scopus*, *Wos* e *Google Acadêmico*, utilizou-se o Teste de *Kruskal-Wallis*, que conforme Siegel (1975)

e Portal *Action*, (2016), é um teste de hipóteses não-paramétrico usado para comparar a função distribuição de três ou mais populações.

3 ANÁLISE E INTERPRETAÇÃO DOS RESULTADOS

A Tabela 1 apresenta as medidas estatísticas de tendência e dispersão para o índice *h* considerando as bases *Scopus*, da *Wos* e *Google Acadêmico*.

Verifica-se que há uma maior concordância entre os valores obtidos para as fontes de informação *Scopus* e *Wos* e que as estimativas obtidas por meio do *Google Acadêmico* são superiores para todos os autores analisados.

Tabela 1 – Medidas de tendência e dispersão para o índice *h* dos 20 autores mais produtivos na temática Estudos Métricos (1980-2014).

Pesquisador	SCOPUS	WOS	GA
Média	22,6	22,0	38,6
Desvio Padrão	9,0	8,7	15,4
Mínimo	9	9	15
Máximo	38	38	75
Coefficiente de Variação	40,0%	39,4%	39,9%

Fonte: elaborada pelos autores.

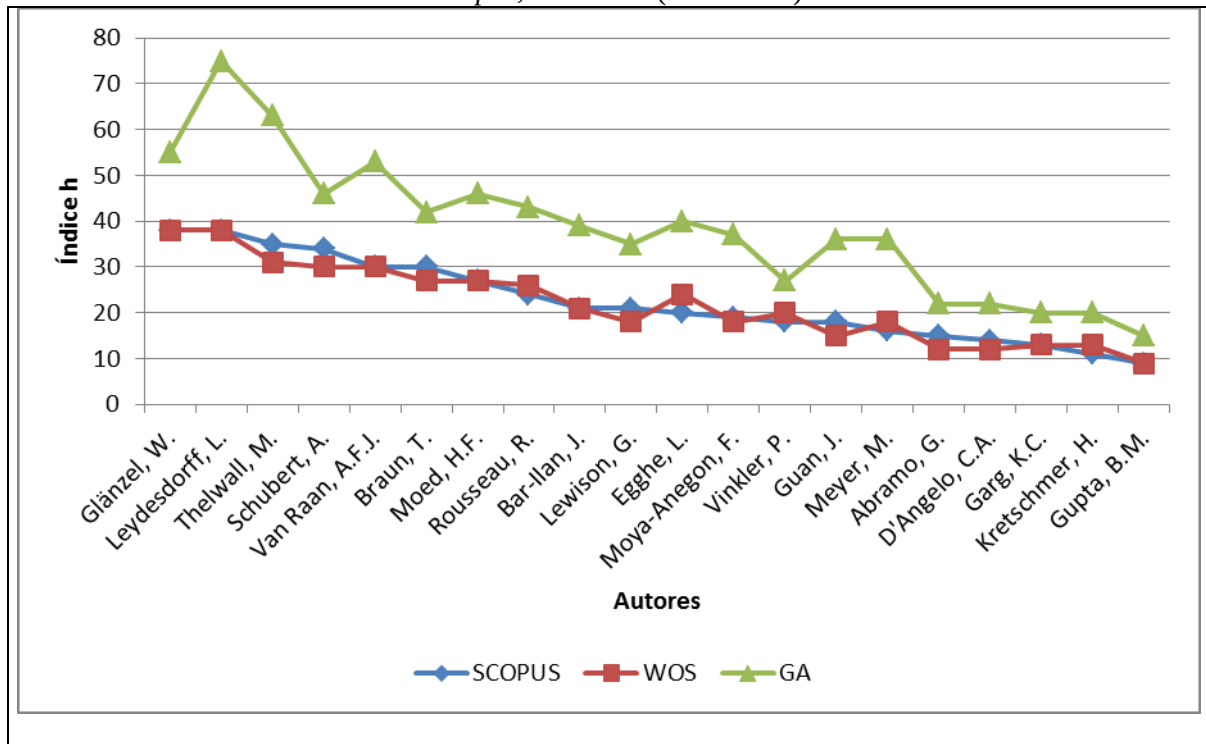
A média do índice *h* para os investigadores desta temática para as diferentes bases foram: $\text{índice-}h_{\text{wos}} = 22,0$, $\text{índice-}h_{\text{scopus}} = 22,6$ e $\text{índice-}h_{\text{GA}} = 378,6$. Verifica-se uma pequena diferença entre os valores médios encontrados para *Scopus* e *Wos*, no entanto, a estimativa para a média do índice *h* para o *Google Acadêmico* é superior às demais.

A Figura 1 apresenta os valores obtidos para o índice *h*, obtidos via *Scopus*, *Wos* e *Google Acadêmico*, para os 20 autores mais produtivos em Estudos Métricos.

Considerando as plataformas *Scopus* e *Wos*, observa-se que os autores Glänzel e Leydesdorff apresentam os maiores índices *h*. Além disso, identifica-se uma pequena oscilação entre as estimativas encontradas para os 20 autores. Os investigadores Schubert e Egghe apresentam uma variação maior entre os valores numéricos encontrados nas plataformas discriminadas.

Para a plataforma Google Acadêmico, o autor Leydesdorff tem o maior índice h ($h = 75$) dentre os pesquisadores analisados, seguido por Thelwall ($h = 63$) e Glänzel ($h = 55$). Gupta apresenta o menor índice h em todas as bases analisadas.

Figura 1. Índice h para os 20 autores mais produtivos em Estudos Métricos, calculados a partir da *Scopus*, *Wos* e *GA* (1980-2014).



Fonte: elaborada pelos autores.

Apesar de existir algumas oscilações entre os valores obtidos para o índice h dos autores analisados, nas bases de dados *Scopus* e *Wos*, conclui-se que, em média, a diferença numérica observada mostrou-se não significativa pelo Teste de *Kruskal Wallis*, considerando um nível de significância de 5% ($p\text{-valor} = 0,0001$), ou seja, independentemente da plataforma utilizada as estimativas obtidas para os índices h não diferem entre si.

Pode-se dizer que as estimativas do índice h oriundas da *Scopus* e *Google Acadêmico*, apresentaram diferença significativa. Similarmente verificou diferença entre os valores dos índices h considerando as bases *Wos* e *Google Acadêmico*, ao nível de significância de 5%.

CONSIDERAÇÕES FINAIS

Nesta pesquisa, realizou-se uma análise comparativa entre as plataformas de dados *Scopus*, *Wos* e *Google Acadêmico*, objetivando verificar a consistência e coerência entre os valores encontrados para o índice *h* de Hirsch.

O Teste de *Kruskal Wallis* foi utilizado para verificar se as distribuições dos indicadores *h* de Hirsch são iguais. Os dados evidenciaram que há consistência entre as estimativas obtidas nas bases de dados internacionais: *Scopus* e *Wos*. Desse modo, apesar da diferença numérica observada, esta não é significativa, considerando uma significância de 5%. Este resultado garante o uso de quaisquer fontes de informação, *Scopus* ou *Wos*, identificará resultados similares para o indicador de citação considerado.

Este mesmo teste apontou que as estimativas dos índices *h* para o Google Acadêmico são superiores às das outras bases, essa característica pode estar relacionada ao modo de indexação dos trabalhos científicos nessas bases de dados. Além disso, como o GA busca referências disponíveis na *Web*, esta informação pode estar duplicada ou ter erros de digitação o que ocasionará uma superestimação da medida.

Diante do exposto, considera-se que esta análise contribui metodologicamente para os Estudos Bibliométricos, pois há carências de estudos relacionados a essa temática. Portanto, sugere-se que a pesquisa em questão seja realizada com outros indicadores; bases de dados e, se estenda a outros domínios científicos.

REFERÊNCIAS

BAR-ILAN, J. **Which h-index? – A comparison of WoS, Scopus and Google Scholar.** *Scientometrics*, Vol. 74, No. 2, p. 257–271, 2008.

COREGNATO, S. E. **Google acadêmico como ferramenta para os estudos de citações: avaliação da precisão das buscas por autor.** Ponto de Acesso, Salvador, V.5, n.3 p. 72-86 dez 2011. Disponível em: <http://www.brapci.ufpr.br/documento.php?dd0=0000011708&dd1=53667>. Acesso em: 20 de janeiro de 2016.

HJORLAND, B. **Domain analysis in information science: eleven approaches traditional as well as innovative.** *Journal of Documentation*, v. 58, n. 4, p.422-462, 2002.

MUGNAINI, R.; JANNUZZI, P. M.; QUONIAM, L. **Indicadores bibliométricos da produção científica brasileira: uma análise a partir da base Pascal.** *Revista Ciência e Informação*. v. 33, p. 123-131, 2004.

OLIVEIRA, E. F. T; GRÁCIO, M. C. C. **Indicadores bibliométricos em ciência da informação: análise dos pesquisadores mais produtivos no tema estudos métricos na base Scopus**. Perspectivas em Ciência da Informação, v.16, n.4, p.16-28, out./dez. 2011.

PORTAL ACTION. **Manual da ferramenta Action para o teste de Wilcoxon – Amostras Independentes**. Disponível em: <http://www.portalaction.com.br/manual-action/532-amostas-independentes>. Acesso em: 19/01/2016.

SCOPUS. **Scopus is the largest abstract and citation database of peer-reviewed literature**: scientific journals, books and conference proceedings. Disponível em: <<http://www.elsevier.com/solutions/scopus>>. Acesso em: 25 de agosto de 2015.

SIEGEL, S. **Estatística não-paramétrica para as ciências do comportamento**. São Paulo: Makron Books do Brasil, 1975.

SILVA, D. **Scopus e web of science – comparação**. (2013). Boletim das bibliotecas da universidade de Aveiro. Disponível em: <http://blogs.ua.pt/bibliotecainforma/?p=1859>. Acesso em: 26/08/2015.

SILVA, D. D.; GRÁCIO, M. C. C. **Indicadores bibliométricos de impacto: estudo comparativo entre os índices h e g aplicados aos pesquisadores mais produtivos na temática Estudos Métricos**. VII Encuentro Ibérico EDICIC 2015: Desafios y oportunidades de las Ciencias de la Información y Documentación en la era digital, 2015.

THOMAZ, P. G; ASSAD, R. S.; MOREIRA, L. F. P. **Uso do fator de impacto e do índice h para avaliar pesquisadores e publicações**. Arquivo Brasileiro de Cardiologia, p. 90-93, 2011.

VANZ, S. A. S.; STUMPF, I. R. C. **Procedimentos e ferramentas aplicados aos estudos bibliométricos**. Informação e Sociedade, v. 20, n.2, p. 67-75, maio/ago. 2010.

VIEIRA, P. V. M.; WAINER, J. **Correlações entre a contagem de citações de pesquisadores brasileiros, usando o Web of Science, Scopus e Scholar**. Perspectivas em Ciência da Informação, v.18, n.3, p.45-60, jul/set. 2013.

WEB OF SCIENCE. **Web of science: the most comprehensive and versatile research platform available**. Disponível em: <http://thomsonreuters.com/en/products-services/scholarly-scientific-research/scholarly-search-and-discovery/web-of-science.html>. Acesso em 25/08/2015.