

# Índice H dos pesquisadores brasileiros: um olhar comparativo entre as bases de dados WoS, Scopus e Google Scholar

H-index of Brazilian researchers: a comparative look at WoS, Scopus and Google Scholar databases

El Índice H de los investigadores brasileños: una mirada comparativa a las bases de datos WoS, Scopus y Google Scholar

Recebido: 08/03/2022 | Revisado: 19/03/2022 | Aceito: 25/03/2022 | Publicado: 31/03/2022

**Alessandra Lacerda Gouvêa**

ORCID: <https://orcid.org/0000-0002-8955-4995>  
Universidade Federal do Triângulo Mineiro, Brasil  
E-mail: [alessandra.gouvea@uftm.edu.br](mailto:alessandra.gouvea@uftm.edu.br)

**Cássia Helene de Ávila**

ORCID: <https://orcid.org/0000-0002-6541-0815>  
Universidade Federal do Triângulo Mineiro, Brasil  
E-mail: [cassia.avila@uftm.edu.br](mailto:cassia.avila@uftm.edu.br)

**Danilo Okimoto Ladislau**

ORCID: <https://orcid.org/0000-0003-3386-692X>  
Universidade Federal do Triângulo Mineiro, Brasil  
E-mail: [d202110021@uftm.edu.br](mailto:d202110021@uftm.edu.br)

**Guilherme Marcelino de Lima**

ORCID: <https://orcid.org/0000-0001-9524-7716>  
Universidade Federal do Triângulo Mineiro, Brasil  
E-mail: [guilherme.marcelino@uftm.edu.br](mailto:guilherme.marcelino@uftm.edu.br)

**Gustavo Henrique Marques Ribeiro**

ORCID: <https://orcid.org/0000-0001-5906-033X>  
Universidade Federal do Triângulo Mineiro, Brasil  
E-mail: [d202020693@uftm.edu.br](mailto:d202020693@uftm.edu.br)

**Juliana Aparecida Vaz**

ORCID: <https://orcid.org/0000-0001-8811-1765>  
Universidade Federal do Triângulo Mineiro, Brasil  
E-mail: [juliana.vaz@uftm.edu.br](mailto:juliana.vaz@uftm.edu.br)

**Lívia Bononi Paiva Tomaz**

ORCID: <https://orcid.org/0000-0002-8798-078X>  
Universidade Federal do Triângulo Mineiro, Brasil  
E-mail: [livia.tomaz@uftm.edu.br](mailto:livia.tomaz@uftm.edu.br)

**Mariana Dias Soares**

ORCID: <https://orcid.org/0000-0001-6283-0542>  
Universidade Federal do Triângulo Mineiro, Brasil  
E-mail: [d202110017@uftm.edu.br](mailto:d202110017@uftm.edu.br)

**Nilson de Moura Netto**

ORCID: <https://orcid.org/0000-0002-0620-0498>  
Universidade Federal do Triângulo Mineiro, Brasil  
E-mail: [d201210944@uftm.edu.br](mailto:d201210944@uftm.edu.br)

**Patrícia Ribeiro Costa**

ORCID: <https://orcid.org/0000-0002-4289-7229>  
Universidade Federal do Triângulo Mineiro, Brasil  
E-mail: [patricia.ribeiro@uftm.edu.br](mailto:patricia.ribeiro@uftm.edu.br)

**Rachel Inês da Silva**

ORCID: <https://orcid.org/0000-0003-1254-2766>  
Universidade Federal do Triângulo Mineiro, Brasil  
E-mail: [rachel.silva@uftm.edu.br](mailto:rachel.silva@uftm.edu.br)

**Wanessa de Sousa Gobbo Vieira**

ORCID: <https://orcid.org/0000-0002-5922-0556>  
Universidade Federal do Triângulo Mineiro, Brasil  
E-mail: [wanessa.gobbo@gmail.com](mailto:wanessa.gobbo@gmail.com)

**Beatriz Gaydeczka**

ORCID: <https://orcid.org/0000-0002-9875-9378>  
Universidade Federal do Triângulo Mineiro, Brasil  
E-mail: [beatriz.gaydeczka@uftm.edu.br](mailto:beatriz.gaydeczka@uftm.edu.br)

**Mônica Hitomi Okura**

ORCID: <https://orcid.org/0000-0002-9875-9378>  
Universidade Federal do Triângulo Mineiro, Brasil  
E-mail: [monica.okura@uftm.edu.br](mailto:monica.okura@uftm.edu.br)

**Ana Claudia Granato Malpass**

ORCID: <https://orcid.org/0000-0001-6487-1225>  
Universidade Federal do Triângulo Mineiro, Brasil  
E-mail: [ana.malpass@uftm.edu.br](mailto:ana.malpass@uftm.edu.br)

**Geoffroy Roger Pointer Malpass**

ORCID: <https://orcid.org/0000-0002-0036-5750>  
Universidade Federal do Triângulo Mineiro, Brasil  
E-mail: [geoffroy.malpass@uftm.edu.br](mailto:geoffroy.malpass@uftm.edu.br)

## Resumo

Este artigo teve como objetivo analisar uma lista de pesquisadores brasileiros, altamente citados, vinculados à UNIFESP, por meio de uma comparação entre a classificação publicada pela Elsevier e as contagens de citações disponibilizadas pelas bases de dados *Web of Science* (da Thomson Reuters), *Scopus* (da Elsevier) e *Google Scholar*. Sendo possível, assim, identificar as semelhanças e diferenças entre as bases de dados e refletir sobre a métrica do impacto científico e produtividade dos pesquisadores. O espaço temporal foi definido por um período de cinco anos, de 2016 a 2021. O resultado da pesquisa foi organizado em tabelas, seguido das respectivas análises e comparações entre as bases consultadas. A partir dos resultados obtidos verificou-se divergências nas contagens de citações de uma base de dados para outra, exatamente por utilizarem fontes distintas para compilação dos dados. Assim, conhecer as características de cada uma das ferramentas disponíveis para a busca de citações e suas limitações pode contribuir na escolha das melhores bases de dados para realização do cálculo do Índice H dos pesquisadores.

**Palavras-chave:** Base de dados; Produção científica e tecnológica; Indicadores de produção científica.

## Abstract

This article aimed to analyze a list of highly cited Brazilian researchers linked to UNIFESP, through a comparison between the classification published by Elsevier and the citation counts provided by the databases Web of Science (Thomson Reuters), Scopus (Elsevier) and Google Scholar. Thus, it was possible to identify the similarities and differences between the databases and reflect on the metrics of scientific impact and productivity of researchers. The temporal space was defined by a five-year period, from 2016 to 2021. The research results were organized in tables, followed by the respective analyses and comparisons between the consulted databases. From the results obtained, there were divergences in the citation counts from one database to another, precisely because they use different sources for data compilation. Thus, knowing the characteristics of each of the available tools for citation search and its limitations can contribute to the choice of the best databases for calculating the H-index of researchers.

**Keywords:** Database; Scientific and technical Activities; Scientific publication indicators.

## Resumen

Este artículo tuvo como objetivo analizar una lista de investigadores brasileños, altamente citados, vinculados a la UNIFESP, a través de una comparación entre la clasificación publicada por Elsevier y los recuentos de citas proporcionados por las bases de datos *Web of Science* (de Thomson Reuters), *Scopus* (de Elsevier) y *Google Scholar*. Así, fue posible identificar las similitudes y diferencias entre las bases de datos y reflexionar sobre las métricas de impacto científico y productividad de los investigadores. El espacio temporal se definió por un período de cinco años, de 2016 a 2021. Los resultados de la investigación se organizaron en tablas, seguidas de los respectivos análisis y comparaciones entre las bases de datos consultadas. A partir de los resultados obtenidos se verificaron divergencias en los recuentos de citas de una base de datos a otra, precisamente porque utilizan fuentes diferentes para la compilación de datos. Em conclusión, conocer las características de cada una de las herramientas disponibles para la búsqueda de citas y sus limitaciones, pueden contribuir a la elección de las mejores bases de datos para el cálculo del Índice H de los investigadores.

**Palabras clave:** Base de datos; Actividades científicas y tecnológicas; Indicadores de producción científica.

## 1. Introdução

Uma das formas mais usuais de avaliar a produtividade e o prestígio dos cientistas é por meio do número de publicações e citações. A quantidade de citações mensura o impacto das publicações, já o *Índice H* - proposto por Jorge Hirsch - faz uma associação entre produtividade e impacto. Esse índice é capaz de fornecer uma estimativa da relevância das contribuições de um pesquisador ao longo do tempo, por meio de um único número (Hirsch, 2005; Marques, 2017). Todavia, essas métricas podem apresentar problemas em sua interpretação e, por isso, devem ser utilizadas juntamente com outros indicadores (Marques, 2017).

Os dados para cálculo são obtidos de diferentes bases de dados, que possuem formas distintas de indexar os documentos, fator que pode ocasionar interpretações equivocadas.

A fim de tentar minimizar esses problemas, a Elsevier<sup>1</sup> publicou em outubro de 2021 a lista dos 100.000 cientistas mais influentes do mundo. Essa classificação pode impactar diretamente a carreira dos cientistas, pois sua visibilidade proporciona uma ampliação da rede de contatos dos pesquisadores e alavanca as possibilidades de obtenção de financiamentos para a pesquisa (Oliveira et al., 2021).

Para compilar os dados dos cientistas mais influentes, a Elsevier utilizou a base de dados *Scopus* para identificar os autores mais citados em todas as áreas do conhecimento. O levantamento foi realizado por meio de um indicador composto, que utiliza seis métricas de citação: (1) citações totais; (2) *Índice H* de Hirsch; (3) *índice hm* de Schreiber ajustado pela coautoria; (4) número de citações de artigos como autor único; (5) número de citações de artigos como autor único ou primeiro autor; e (6) número de citações de artigos como autor único, primeiro ou último autor (Ioannidis et al., 2019). Todos os indicadores são divulgados com e sem autocitações.

Dos 100.000 cientistas mais influentes, 812 são brasileiros. A maioria é vinculada às instituições públicas, como é o caso da Universidade Federal de São Paulo (UNIFESP) que possui 18 pesquisadores listados. Dada a notoriedade desta classificação para o mundo científico, o objetivo deste estudo é analisar a lista dos pesquisadores brasileiros vinculados à UNIFESP altamente citados por meio de uma comparação entre a classificação publicada pela Elsevier e as contagens de citações disponibilizadas pelas bases de dados *Web of Science* (da Thomson Reuters), *Scopus* (da Elsevier) e *Google Scholar*. Dessa forma, foi possível identificar as semelhanças e diferenças entre as bases de dados.

A relevância deste estudo incide no fato de que há diferenças nas contagens de citações de uma base de dados para outra, justamente por utilizarem fontes distintas para compilação dos dados. Este fator pode ocasionar prejuízos aos pesquisadores, como, por exemplo, no financiamento de pesquisas (Bar-Ilan, 2008; Ioannidi et al., 2019). Dessa forma, conhecer as características de cada uma das ferramentas disponíveis para a busca de citações e suas limitações auxiliará na escolha das melhores bases de dados para cálculo do *Índice H*.

## 2. Metodologia

Este estudo teve como ponto de partida a consulta de informações no banco de dados do site da *Elsevier* (<https://elsevier.digitalcommonsdata.com/datasets/btchxktzyw/3>, recuperado em 23, novembro de 2021) para busca dos pesquisadores mais influentes. Foi feito o *download* do arquivo “Table\_1\_Authors\_career\_2020\_wopp\_extracted\_202108.xlsx” e a partir desta lista, foi estabelecido o critério de filtrar apenas os pesquisadores da Universidade Federal de São Paulo (UNIFESP). Assim, foram selecionados os perfis de 18 pesquisadores. Ainda no banco de dados da *Elsevier* foram coletados os dados do ano da publicação do primeiro artigo de cada pesquisador selecionado.

Foi realizada a busca de cada pesquisador selecionado nas bases de dados *Web of Science (WoS)* (<https://www.webofscience.com>, recuperado em 10, novembro de 2021), *Scopus* (<https://www.scopus.com>, recuperado em 10, novembro de 2021) e *Google Scholar* ([scholar.google.com.br](https://scholar.google.com.br), recuperado em 10, novembro de 2021). Para a pesquisa em todas as bases foi definido o período de 05 anos, sendo de 2016 a 2021. O período foi limitado aos últimos 05 anos para comparações justas entre as três bases de dados. Além disso, nas bases de dados da *Scopus* e *Web of Science (WoS)* para alguns pesquisadores era possível visualizar mais de uma opção na pesquisa. Isso pode acontecer devido à semelhança de nomes com outros pesquisadores. Sendo assim, foi definido que para estes casos seria considerado o perfil com mais publicações atreladas.

---

<sup>1</sup> É uma empresa global que trabalha com informações analíticas de pesquisadores e instituições, contribuindo para o progresso da ciência e financiamento de pesquisas (<https://www.elsevier.com/pt-br/about>, recuperado em 08, novembro de 2021).

Após o término da pesquisa, os resultados foram agrupados em forma de tabelas e, em seguida, realizadas as análises e comparações entre as bases consultadas. Como parte dos resultados e discussões também foi analisada qual a influência do ano da publicação do primeiro artigo ao fator de impacto. Essa análise adicional teve como intuito identificar se o período da primeira publicação torna os números do *Índice H* mais equilibrados entre as bases.

### 3. Resultados e Discussão

#### 3.1 Síntese de informações sobre as Bases de Dados: à guisa de comparações

As bases de dados científicas são fontes utilizadas para avaliar pesquisas e pesquisadores por meio de indicadores bibliométricos baseados no índice de citações (ISI), como por exemplo a métrica de impacto do *Índice H*, que busca combinar, em suma, produtividade e impacto. Atualmente, destacam-se três diferentes bases de dados muito utilizados pela comunidade acadêmica: *Web of Science (WoS)*, *Scopus* e *Google Scholar*.

A plataforma *Web of Science*, ferramenta utilizada para pesquisa de periódicos, artigos e outros documentos do meio acadêmico, até meados do ano de 2004 era única base que permitia a obtenção extensa sobre dados de citações.

Segundo Bar-Ilan (2007), a *Web of Science* é uma ferramenta que extrai referências completas dos itens indexados, na qual a interface de referência evidencia em sua totalidade as citações de obras de um autor, independente se estão indexados ou não no *Web of Science*.

A *WoS* é um instrumento que não é fornecido de maneira gratuita, ou seja, é um objeto que para ter acesso às suas funcionalidades, o operador deve pagar para obter um usuário (*login*) na ferramenta digital. A base de dados da *Web of Science* concede continuamente informações de uma vasta gama de disciplinas (em torno de 8700) dos periódicos mais renomados de pesquisa de altíssimo impacto em todo globo terrestre (Bar-Ilan, 2007).

Assim, segundo Vieira e Wainer (2013), ao se avaliar pesquisadores por indicadores bibliométricos como o *Índice H* (combinação de produtividade e impacto), cada serviço como *WoS*, *Google Scholar* ou *Scopus* possui suas próprias premissas e características de pesquisas. No caso em questão, a ferramenta digital de serviço *WoS* é considerada a mais tradicional das três plataformas citadas acima, e caracteriza-se por considerar um nicho pequeno de revistas científicas para seu monitoramento e, através disso, realiza a contagem de todas as citações executadas por meio de artigos publicados nessas mesmas revistas, identificando as fontes das citações.

Para conhecimento, uma revista sendo indexada na *WoS*, apenas por este motivo já é considerada de qualidade, uma vez que a ferramenta segue critérios de indexação extremamente rigorosos neste aspecto (Vieira & Warner, 2013). Segundo Grácio e Silva (2017), a *Web of Science* possui um sistema de cobertura e contagem de citações que alteiam ao ano de 1900, e que, até o ano de 2013 encontravam-se mais de 12 mil títulos entre 3 mil a 3,5 mil editoras internacionais, sendo seus dados atualizados semanalmente.

Entretanto a *WoS* (Vieira & Warner, 2013) apresenta também algumas fragilidades, tanto em produção quanto em citações em documentos relacionados às áreas de engenharia e ciência da computação (com produção científica em conferências), ciências humanas e sociais (produção científica por monografias e outros), sendo que sua utilização é mediante pagamento de assinatura, através de uma licença nacional; contagens de citações são deflacionadas, quando estas citações não são agrupadas, pois o *WoS* contabiliza apenas correspondências exatas.

Já em relação a base *Scopus*, trata-se de uma base de dados que inclui resumos e citações de literatura, abrangendo conteúdos nas seguintes áreas do conhecimento: química, física, matemática, engenharia, ciências da saúde e vida, ciências sociais, psicologia, economia, biologia, agricultura, ciências ambientais e ciências gerais (Elsevier, 2020).

Cerca de 60% dos títulos são de outros países que não os Estados Unidos e 85% do conteúdo é indexado aproveitando os vocabulários controlados para a definição das palavras-chave e/ou descritores. Jacso (2005) afirma que a *Scopus* se encontra, juntamente com o *Google Scholar* e a *WoS*, entre as maiores bases de dados multidisciplinares (Elsevier, 2020).

A base *Scopus* é alterada diariamente, e seu portfólio conta com mais de 21.500 publicações revisadas por pesquisadores atuantes em suas áreas, das quais mais de 4.200 são de acesso aberto. Além disso, disponibiliza mais de 360 publicações comerciais; “*Articles-in-Press*”, ou seja, artigos que foram julgados e aceitos para publicação nas revistas científicas, de mais de 5.000 editoras internacionais (Elsevier, 2020).

A base pesquisa, simultaneamente, o conteúdo dos artigos indexados em sua própria plataforma na *web*, através do buscador *Scirus* da Elsevier para páginas com conteúdo científico e, ainda, bases de patentes sobre o assunto pesquisado. Quanto aos livros, dispõe de mais de 113.000, com cerca de 10.000 adicionados a cada ano. A base possui 7,2 milhões de anais de conferências, oriundos de mais de 88.880 eventos realizados no mundo inteiro. Dessa maneira, essa base de dados possui aproximadamente 60 milhões de registros (Elsevier, 2020).

No que se refere à seleção de conteúdos independentes e títulos recém-sugeridos, estes são revisados pelos membros do *Scopus Content Selection and Advisory Board (CSAB)* que fazem uso de medidas quantitativas e qualitativas para avaliar essas publicações. O *CSAB* é composto por dezesseis pesquisadores e cientistas de renome internacional e cada membro representa um determinado campo do conhecimento (Elsevier, 2020).

As recomendações do *CSAB* influenciam diretamente no direcionamento tomado pelo *Scopus* e na prioridade que será dada aos conteúdos requisitados. Os títulos de periódicos, antes de serem avaliados pelo *CSAB*, passam por uma revisão preliminar, pois devem atender a certos critérios básicos, por exemplo: consistir em uma publicação revisada por pares e conter uma descrição disponível do processo de avaliação pelos revisores; ser publicada em uma base usual; ter um conteúdo que é relevante para um público internacional, entre outros (Figueiredo et al., 2017).

Figueiredo et al. (2017) ainda identificaram que a base de dados *Scopus* se mostrou uma boa fonte de busca. Um fator importante é a ótima organização da interface do site. Com o *Scopus*, a pesquisa se torna mais didática e fluida, apresentando comandos de filtro de resultado, abertura e *download* de arquivos, formas de organização da listagem de arquivos bem simplificada e clara. Além do mais, não existe a possibilidade de haver retorno de artigos repetidos. Segundo pesquisas realizadas pelos autores com a palavra-chave “*Macrophages*”, o *Scopus* mostrou 100% de eficácia, ou seja, todos os artigos retornados se encaixam nos critérios pré-estabelecidos, embora o número total de artigos retornados tenha sido baixo.

Em relação ao *Google Scholar*, consiste em ferramenta do *Google* destina a buscas de informações científicas, lançada em 18 de novembro de 2004 (Payne, 2004 *apud* Bar-Ilan, 2007), denominada no Brasil por *Google Acadêmico (GA)*, que diferentemente da *Scopus* e da *WoS* possui acesso gratuito, e permite localizar trabalhos acadêmicos de vários tipos, como teses, dissertações, artigos de congressos e de periódicos, tanto de acesso aberto como fechado, e em múltiplas línguas, recurso este aperfeiçoado a partir de novembro de 2012 em uma nova versão que oferece classificações gerais por idioma, por área, temas e disciplinas, embora limitadas para revistas em inglês, excluindo idiomas em chinês, português, alemão, espanhol, francês, coreano, japonês, holandês e italiano (Delgado & Repiso, 2013).

Este banco indexa dados de textos completos disponíveis gratuitamente em servidores de pré-impressão ou sites pessoais, como também de editoras dispostas a fornecer pelo menos o resumo do artigo gratuitamente (*Google Scholar*, 2005 *apud* Bar-Ilan, 2007), cuja visualização completa do texto pode ser paga ou por assinatura, donde as referências são automaticamente extraídas dos textos indexados, e quando a referência em si não seja indexada pelo *GA*, retorna apenas o número de citações daquele texto nos resultados da pesquisa (Bar-Ilan, 2007).

O *Google Scholar* busca referências em qualquer documento que esteja disponível na *Web*, inclusive versões eletrônicas de revistas e conferências, possibilitando que qualquer documento citado por algum outro documento disponível na *Web* tenha

as suas citações monitoradas, e dispondo a partir deste monitoramento a funcionalidade de envio de notificações sobre assuntos do interesse do pesquisador. No entanto, o GA não contabiliza as citações feitas por livros, embora considere as citações recebidas por livros (Bar-Ilan, 2007; Vieira & Wainer, 2013), limitando seus serviços para algumas das áreas da ciência, como Ciências Humanas e Sociais, posto que sua produção científica se faz por meio de monografias, livros e capítulos de livros (Vieira & Wainer, 2013).

Esta ferramenta é muito utilizada para recuperação de publicações científicas devido à sua ampla abrangência de pesquisas proporcionada por suas funcionalidades como meta-buscador com fulcro iniciativa de acesso livre à informação científica (*Open Access Initiative - OAI*) e índice de citações, transformando-se num indicador de visibilidade das publicações. No caso do Brasil, significativa parcela da comunidade científica tem acesso aos principais artigos publicados em virtude do investimento da CAPES no Portal de Periódicos (Mugnaine & Strehl, 2008).

Todavia, este processo contrapõe ao formato do ISI, cujas publicações passam por rigoroso processo seletivo para serem indexadas, sendo uma preocupação evidenciada quanto a esta abrangência ser confiável. Por outro lado, pesquisadores reconhecem a sua importância para buscas simples, em casos de procura específica de um documento conhecido ou quando se pretende acessar links de outras fontes (Mugnaine & Strehl, 2008).

Em suma, essas plataformas de busca possuem semelhanças e diferenças relevantes em suas formas de pesquisa. A distinção já pode ser observada no que concerne ao período da cobertura de dados. A *WoS* abrange citações a partir de 1961; a *Scopus*, embora seja uma base de dados recente, disponibiliza citações desde 1996; já o GA oculta a informação sobre o período contemplado na sua base de pesquisa (Bar-Ilan, 2007).

As três bases de dados têm a mesma finalidade de busca, mas cada uma com suas particularidades, conforme mencionado anteriormente, o que resulta em diferenças significativas de resultados na realização de pesquisa de um mesmo cientista. De acordo com Deolindo Silva e Cabrini Grácio (2017, p. 197): “as bases de dados dão suporte para a almejada visibilidade da produção científica, pois a partir dessas fontes de informação são encontrados resultados de pesquisas, de pesquisadores e de instituições, nacionais e internacionais”.

Em geral, é possível observar que o *Google Scholar* apresenta um número consideravelmente maior de respostas de citações de *Índice H* que as demais bases, devido à sua atuação em áreas mais diversificadas, tendo em vista que essa ferramenta busca publicações científicas de vários tipos, tanto de acesso aberto como fechado e em diversas línguas, conforme sua extensa abrangência mencionada anteriormente. Em relação a *Web of Science* e a *Scopus*, nota-se que os resultados de citações de *índice h* são muito similares para cada autor, o que será exemplificado com mais detalhes nas seções seguintes.

### 3.2 Índice H

O *Índice H* foi proposto em 2005 por Jorge E. Hirsch, sendo definido “como o número de artigos com número de citações maior ou igual a *h*, como índice útil para caracterizar a produção científica de um pesquisador” (Hirsch, 2005, p.16569). Em outras palavras, “O *índice-h* de um pesquisador é definido como o maior número “*h*” de artigos científicos desse pesquisador que têm pelo menos o mesmo número “*h*” de citações cada um” (Marques, 2017, p. 35). Nesse sentido, caso um pesquisador tenha muitas publicações, porém pouco citadas, ou possua citações expressivas, contudo provenientes de um número limitado de trabalhos publicados, possuirá pequeno *Índice H*. Silva et al. (2018) esclarecem que o índice avalia fatores relacionados à repercussão científica, representada pelas citações, e à produção do pesquisador, representada pela quantidade de artigos realizados.

O *Índice H* possui papel relevante no meio acadêmico dado sua metodologia para mensurar o impacto e a produção científica de um pesquisador. Para Drei et al. (2021, p.11) o índice “não só, se insere nos trabalhos, artigos e revistas, bem como

as bases que esses estão inseridos, como também está presente na vida acadêmica dos pesquisadores, indicando o impacto de suas produções e como elas se desdobram ao longo dos anos”.

Marques (2017) complementa que além de apresentar o desempenho de um pesquisador, o *Índice H* pode ser medido também para grupos de pesquisa, universidades, países e revistas científicas. O autor também informa que a métrica pode contribuir nas decisões relativas ao destino de verbas e à concessão de prêmios. Contudo, vale destacar que o índice não considera aspectos mais qualitativos da pesquisa e do pesquisador, somando-se ainda a dificuldade de representação de cientistas com produção reduzida ou no começo de carreira (Drei, Carvalho, Mello & Angulo-Meza, 2021). Em relação às críticas elencadas quanto ao indicador, Marques (2017) comenta também a falha ao contrapor o desempenho de pesquisadores que desenvolvem trabalhos em áreas distintas; a possibilidade de manipulação nos dados com autocitações; a dificuldade ao comparar pesquisadores de áreas com cultura de publicações em livros, pois o índice considera o mesmo peso para livros e artigos; além de não considerar o contexto das citações.

Nota-se a importância do *Índice H* no meio acadêmico e sua fácil aplicação, possui cálculo simples e condensa no seu resultado a produtividade e o impacto da produção do pesquisador, todavia, considerando-se algumas limitações, para maior efetividade, é interessante que seja analisado em conjunto a outros indicadores bibliométricos.

As Tabelas 1 e 2 apresentam o *Índice H* dos autores brasileiros selecionados neste trabalho e o *Índice H* para os vinte autores mais influentes, de acordo com a análise e publicação dos autores mais influentes do mundo, da Elsevier. Nota-se que o *Google Scholar* apresenta *índices h* superiores para todos os autores, o que é relacionado ao fato de que a base busca por trabalhos e revistas em toda a internet. A comparação entre ambas as tabelas indica que os autores mais influentes têm *Índice H* superior ao *Índice H* dos autores brasileiros selecionados para realização deste trabalho.

**Tabela 1** – Índices H dos autores brasileiros nas bases de dados *Google Scholar*, *Web of Science* e *Scopus*.

Autor	Ano da primeira publicação	Google Scholar	WoS	Scopus
Tufik, Sergio Brasil	1973	65	18	14
Ferri, Cleusa P.	1994	46	8	9
Ambrósio, Renato	2001	41	15	18
Colombo, Arnaldo L.	1991	53	19	21
Machado, Flavia R.	1999	46	26	29
Ferraz, Marcos Bosi	1989	34	3	4
Carlini, Elisaldo A.	1958	31	3	5
Cavalheiro, Esper A.	1976	36	10	10
Suchecky, Deborah	1986	28	10	10
Gales, Ana C.	1996	43	15	16
Yoshida, Nobuko	1975	24	11	6
Ribeiro, Daniel Araki	2004	28	13	14
Gomes, Jose Alvaro P.	1993	29	15	16
Alves, Alexandre	2003	-	12	12
Ferreira, Lydia Masako	1984	42	0	11
Nahas, Fabio Xerfan	1997	26	7	7
Atallah, Alvaro Nagib	1985	39	9	9
Riske, Karin A.	1997	26	9	9
Guinsburg, Ruth	1988	31	10	11

Fonte: Adaptado de Baas et al. (2021) <https://elsevier.digitalcommonsdata.com/datasets/btchxktzyw/3>



**Tabela 2** – Índices H dos vinte autores mais influentes nas bases de dados.

<b>Autor</b>	<b>Ano da primeira publicação</b>	<b>Google Scholar</b>	<b>WoS</b>	<b>Scopus</b>
Wang, Zhong Lin	1986	201	112	111
Kresse, Georg	1992	84	23	26
Bengio, Yoshua	1988	177	34	54
Altman, Douglas G.	1972	191	50	53
Bandura, Albert	1953	141	4	7
Grätzel, Michael	1971	186	99	96
Hinton, Geoffrey	1979	126	9	16
Newman, M. E.J.	1991	88	25	0
Grimme, Stefan	1992	86	37	35
Kessler, Ronald C.	1975	198	47	49
Ioannidis, John P.A.	1994	157	2	62
Friston, Karl J.	1989	159	185	46
Perdew, John P.	1970	79	24	26
Zadeh, Lotfi A.	1949	65	0	5
Moher, David	1984	125	49	45
Higgins, Julian P.T.	1996	109	36	48
Becke, Axel D.	1977	N/A	8	8
Bartel, David P.	1989	94	17	19
Goodenough, John B.	1953	N/A	59	62
Semenza, Gregg L.	1984	121	28	29

Fonte: Adaptado de Baas et al. (2021) <https://elsevier.digitalcommonsdata.com/datasets/btchxktzyw/3>

### 3.3 Ano da primeira publicação x Índice H

O *Índice H* é cumulativo e tem relação com o número de citações, e com o número de publicações. Considera em sua estrutura aspectos relativos ao impacto científico (citação) e à produtividade (número de artigos produzidos) do pesquisador (Silva et al., 2018, p. 265).

Ainda segundo Thomaz et al., (2011):

Quando dizemos que o índice H de um pesquisador é dez, significa que ele tem, pelo menos, dez artigos publicados, cada um deles com, pelo menos, dez citações. Quanto maior o número de artigos de grande interesse publicado pelo

pesquisador, maior será o número de citações alcançadas, e maior será seu índice H, refletindo a qualidade acadêmico-científica do pesquisador e sua capacidade produtiva. (Thomaz et al., 2011, p. 91).

Conforme os dados levantados da Tabela 2, dos autores pesquisados, estes publicaram seus primeiros artigos entre 1958 a 2004. Em relação ao *Índice H* nas bases pesquisadas o autor Tufik é o que tem o maior *Índice H* na base Google Scholar e sua primeira publicação foi no ano de 1973; nas bases WoS e *Scopus*, o autor Machado tem maior *índice h* dentre os autores analisados e sua primeira publicação foi feita no ano 1999.

Conforme se verifica na Tabela 2, que apresenta as citações da primeira publicação em periódicos, é possível constatar que não existe influência no *Índice H*, devido a um dos fatores fundamentais para aumento do *índice h* ser a quantidade de citações que o artigo atingiu.

Constata-se que quanto maior o período analisado, mais equilibrados tornam-se os números do *Índice H* (Avena & Barbosa, 2017, p. e-03262-7). Na Tabela 2 verifica-se que os autores com data da primeira publicação mais antiga têm o *índice h* maiores e menores no grupo avaliado. Autores com data da primeira publicação mais recente também apresentam variação no *índice h*, não sendo essa data fator determinante para *Índice H*.

A Tabela 3 apresenta o ano da primeira publicação em cada uma das bases de dados analisadas.

**Tabela 3** – Citações da primeira publicação nas bases de dados.

AUTOR	Ano	Google Scholar	Ano	WoS	Ano	Scopus
Tufik, Sergio Brasil	1973	2	1998	20	1973	2
Ferri, Cleusa P.	1994	14	1993	14	1994	9
Ambrósio, Renato	2001	2	1984	4	2001	24
Colombo, Arnaldo L.	1991	25	1991	2	1991	10
Machado, Flavia R.	1999	0	2000	4	1999	0
Ferraz, Marcos Bosi	1989	0	1990	161	1989	1
Carlini, Elisaldo A.	1958	0	1980	6	1958	16
Cavalheiro, Esper A.	1976	0	1980	8	1976	27
Suhecki, Deborah	1986	152	1986	50	1986	47
Gales, Ana C.	1996	11	1996	4	1996	10
Yoshida, Nobuko	1975	12	2010	32	1975	9
Ribeiro, Daniel Araki	2004	1	2005	24	2004	6
Gomes, Jose Alvaro P.	1993	21	1984	1	1993	1
Alves, Alexandre	2003	não encontrado	2013	8	2003	10
Ferreira, Lydia Masako	1984	0	2001	21	1984	4

Nahas, Fabio Xerfan	1997	4	1997	55	1997	44
Atallah, Alvaro Nagib	1985	4	1989	26	1985	29
Riske, Karin A.	1997	35	1997	32	1997	31
Guinsburg, Ruth	1988	0	1998	74	1988	0

---

Fonte: Adaptado de Baas et al. (2021) <https://elsevier.digitalcommonsdata.com/datasets/btchxktzyw/3>

Verifica-se que os fatores que mais influenciam o *Índice H* são o número de publicações, número de citações, natureza da área de atuação do pesquisador, cobertura dos periódicos nas diversas bases (Avena & Barbosa, 2017, p. e-03262-5) de acesso aberto ou fechado ao periódico.

#### 4. Considerações Finais

O presente estudo apresentou uma análise realizada com dados de pesquisadores brasileiros, altamente citados, da UNIFESP, por meio de uma comparação entre a classificação publicada pela Elsevier e as contagens de citações disponibilizadas pelas bases de dados *Web of Science* (da Thomson Reuters), *Scopus* (da Elsevier) e *Google Scholar*. Mesmo tendo a mesma finalidade de busca para avaliação de pesquisas e pesquisadores através de indicadores baseados no índice de citações (ISI) cada base de dados apresenta suas peculiaridades.

Observou-se que o *Google Scholar* apresenta um número superior em relação às citações de *Índice H*, comparado com às outras bases, haja vista sua atuação em áreas mais diversificadas, os variados tipos de publicações científicas encontrados, tanto de acesso aberto como fechado, e a diversidade de línguas, devido à sua extensa abrangência. Em relação à *Web of Science* e *Scopus* verificou-se que os resultados de citações de *Índice H* são semelhantes para cada autor.

Há grande evidência da relevância do *Índice H* na academia devido à simplicidade dos cálculos em mensurar a produtividade e o impacto da produção do pesquisador, entretanto, deve-se também avaliar outros indicadores bibliométricos para aumentar a efetividade dessa ferramenta. Ressalta-se que os fatores que mais influenciam o *Índice H* são: número de citações; número de publicações; natureza da área de atuação do pesquisador; cobertura dos periódicos em diferentes bases de dados; e se o acesso é aberto ou fechado ao periódico.

Além de constatar que os pesquisadores mais influentes (de acordo com o levantamento dos autores mais influentes do mundo, da Elsevier) tem *Índice H* superior à lista de pesquisadores brasileiros, identificou-se que quanto maior o período analisado, mais equilibrados tornam-se os números do *índice h*, pois foi verificada grande variação do *índice h*, tanto em análise mais antigas da primeira data de publicação, quanto mais recente, não sendo, a data fator determinante para o *Índice H*.

Sendo assim, este trabalho apresenta dados, de acordo com o recorte temporal dos últimos cinco anos (2016-2021). Espera-se que as informações apresentadas possam contribuir para pesquisas futuras sobre a confiabilidade das bases de dados e sobre o impacto do *Índice H* na vida do pesquisador. Assim, os autores sugerem que os próximos trabalhos nesta vertente apresentem atualização dos dados apontados neste artigo, assim como, recomendam mais estudos, principalmente, sobre a base de dados do *Google Scholar* devido à sua extensa abrangência.

## Referências

- Avena, M. J., & Barbosa, D. A. (2017). Indicadores bibliométricos das Revistas de Enfermagem sob a ótica das bases indexadoras. *Revista da Escola de Enfermagem da USP*, 51, e03262-1-9. <https://doi.org/10.1590/S1980-220X2017014603262>.
- Baas, J., Boyack, K., & Ioannidis, J. P. A. (2021). August 2021 data-update for "Updated science-wide author databases of standardized citation indicators". *Elsevier BV*. <https://elsevier.digitalcommonsdata.com/datasets/btchxktzyw/3>. <https://doi.org-10.17632/btchxktzyw.3>.
- Bar-Ilan, J. (2008). Which h-index? — A comparison of WoS, Scopus and Google Scholar. *Scientometrics*, 74(2), 257–271. <https://doi.org/10.1007/s11192-008-0216-y>.
- Delgado-López-Cózar, E., & Repiso-Caballero, R. (2013). The impact of scientific journals of communication: Comparing Google Scholar Metrics, Web of Science and Scopus. *Comunicar*, 21(41), 45–52. <https://doi.org/10.3916/c41-2013-04>.
- Drei, S. M., Carvalho, E. B. de., Mello, J. C. C. B. S. de., & Angulo-Meza, L. (2021). Evaluation of the papers' panorama on h-index over the decade of 2011 to 2020. *Research, Society and Development*, 10(11), e360101119785. <https://doi.org/10.33448/rsd-v10i11.19785>.
- Figueiredo, A. R., Wanderley B. G., Vilas Boas, T. S., & Santos, M. C. (2017). Estudo da eficiência dos portais Science Direct, Scopus, Lilacs e Periódicos CAPES, evidenciando seus aspectos positivos e negativos. *Scientia Amazonia*, 6(2):1–10.
- Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National academy of Sciences*, 102(46), 16569-16572. <https://doi.org/10.1073/pnas.0507655102>.
- Ioannidis, J. P., Baas, J., Klavans, R., & Boyack, K. W. (2019). A standardized citation metrics author database annotated for scientific field. *PLoS biology*, 17(8), e3000384.
- Marques, F. (2017). Os limites do índice-h. *Boletim Técnico do PPEC*, 2(1), 35-39. <https://econtentos.bc.unicamp.br/boletins/index.php/ppec/article/view/9048>.
- Mugnaini, R., & Strehl, L. (2008). Recuperação e impacto da produção científica na era google: uma análise comparativa entre o google acadêmico e a web of science 10.5007/1518-2924.2008v13nesp1p92. *Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação*, 13(1). <https://doi.org/10.5007/1518-2924.2008v13nesp1p92>.
- Oliveira, L., Reichert, F., Zandona, E., Soletti, R. C., & Staniscuaski, F. (2020). The 100,000 most influential scientists rank: the underrepresentation of Brazilian women in academia. *bioRxiv*.
- Silva, D. D., & Grácio, M. C. C. (2017). Índice h de Hirsch: análise comparativa entre as bases de dados Scopus, Web of Science e Google Acadêmico. *Em Questão*, 23, 196. <https://doi.org/10.19132/1808-5245230.196-212>.
- Silva, D. D., Almeida, C. C., & Gácio, M. C. C. (2018). Avaliação científica de periódico em ciências sociais: junção dos indicadores fator de impacto e índice h. *Encontro Brasileiro de Bibliometria e Cientometria*, 6(6), 264–271. <http://hdl.handle.net/20.500.11959/brapci/117487>.
- Silva, D. D., Almeida, C. C. de, & Grácio, M. C. C. (2018). Associação do Fato de Impacto e do Índice h para a avaliação de periódicos científicos: uma aplicação no campo da Ciência da Informação. *Em questão*, 24, 132-151, Edição Especial 6 EBBC. <https://doi.org/10.19132/1808-5245240.132-151>
- Thomaz, P. G., Assad, R. S., & Moreira, L. F. P. (2011). Uso do Fator de impacto e do índice H para avaliar pesquisadores e publicações. *Arquivos Brasileiros de Cardiologia*, 96(2), 90–93. <https://www.scielo.br/j/abc/a/kp6RzbscSJt5snkS7XQvsq/?lang=pt>.
- Vieira, P. V. M., & Wainer, J. (2013). Correlações entre a contagem de citações de pesquisadores brasileiros, usando o Web of Science, Scopus e Scholar. *Perspectivas em Ciência da Informação*, 18(3), 45–60. <https://doi.org/10.1590/s1413-99362013000300004>.