

AVALIAÇÃO DOS CIENTISTAS: UTILIZAMOS AS MÉTRICAS CORRETAS?

Josefredo R. Pliego Jr.^{a,*} ^aDepartamento de Ciências Naturais, Universidade Federal de São João del-Rei, 36301-160 São João del-Rei – MG, Brasil

Recebido em 12/08/2022; aceito em 29/08/2022; publicado na web em 07/10/2022

SCIENTISTS' ASSESSMENT: ARE WE USING THE CORRECT METRICS? The use of metrics to evaluate scientists is widespread in the present time, with implications for hiring, fellowships, and research grants. Such fact requires that metrics must be constantly scrutinized to be improved. This work analyzes the use of metrics in the area of chemistry in Brazil and discusses its limitations and shortcomings. The main findings indicate that the use of the impact factor must be complemented by the cited half-life of the journals, and a composite metric named influence factor is proposed, similar to the R-impact. The h-index is not a good metric anymore because does not correct for authorship inflation. The individual h-index (h_i), which takes into account fractional counting of citations, is more reliable than the h-index. An analysis on the use of h_i -index with randomly selected 15 Brazilian chemists among the top 500 more productive shows an important effect on the ranking order.

Keywords: fractional h-index; R-impact; fractional counting; cited half-life; fractional authorship.

INTRODUÇÃO

A atividade científica cresceu intensamente ao longo do último século, ficando cada vez maior e mais complexa. Tal atividade requer uma quantidade considerável de recursos, os quais devem ser bem investidos para que os países possam ter um retorno em termos de avanços na ciência de cada nação. Como tem sido estabelecido, nações cientificamente desenvolvidas tendem a serem economicamente mais bem sucedidas.¹ Para tomar decisões sobre investimentos dos recursos financeiros, é necessário conhecer onde está a excelência em pesquisa científica, para que aqueles grupos de pesquisa ativos e produtivos possam ser apoiados com os recursos necessários. Essa demanda, associada à complexidade do julgamento da atividade científica, levou naturalmente à criação de índices numéricos ou métricas que permitissem ranquear os pesquisadores em ordem de prioridade de recebimento de recursos. A adequação das métricas utilizadas atualmente no Brasil, mais especificamente na área de química, é o objeto de discussão deste manuscrito.

As métricas utilizadas para avaliar cientistas apontam em duas direções: quantidade e qualidade. Tal abordagem é baseada na ideia de que um melhor desempenho do pesquisador corresponde a mais artigos publicados em bons periódicos, bem como o número de citações que esses artigos recebem. Essas têm sido as métricas utilizadas para julgamento em várias partes do mundo, não só no Brasil. Em um artigo de 2010 na *Nature*, foi perguntado a vários cientistas leitores do periódico sobre o uso de métricas para avaliação dos mesmos.² O resultado foi bem interessante: observou-se que as métricas eram amplamente utilizadas na contratação, promoção e decisão sobre posições permanentes dos docentes nas instituições. Uma fatia majoritária (63%) dos que responderam disseram não estar satisfeitos sobre como as métricas estavam sendo utilizadas. Um número considerável (71%) disse estar preocupado de que seus colegas possam trapacear no sistema de métricas. Entretanto, muitos disseram estarem satisfeitos com a transparência das métricas, enquanto outros disseram que as métricas não estavam sendo usadas de forma suficiente ou consistentemente. Um comentário interessante foi de que o uso de métricas pode combater a endogenia em contratações, em que preferências pessoais prevalecem sobre

competência. O mais surpreendente foi que quando os respondedores foram perguntados sobre como queriam ser avaliados, as respostas mais frequentes foram: 1) publicações em periódicos de alto impacto; 2) projetos aprovados; 3) orientações de estudantes; 4) número de citações recebidas. Ou seja, usando as métricas de produção e qualidade que já utilizamos. Assim, concluiu-se que o problema não parece ser sobre o uso das métricas em si, mas sim o uso adequado e justo das métricas, de forma a avaliar corretamente os cientistas. Também é de meu entendimento que as métricas devem refletir de fato a produtividade e excelência científica de cada cientista.

Uma das métricas utilizadas atualmente para ranquear cientistas, o índice h, tem sido criticamente questionado em várias publicações recentes, indicando que não representa mais uma medida adequada da performance do cientista.³⁻⁶ Da mesma forma, o fator de impacto, largamente utilizado para ranquear periódicos, tem importantes limitações, uma vez que mede citação de artigos recentes (2 anos), e não artigos que podem impactar durante décadas. Dessa forma, periódicos que publicam pesquisa mais fundamental e com menor citação imediata acabam tendo um menor fator de impacto. Com base nessas colocações, deve-se dizer que é extremamente importante que os indicadores utilizados para ranquear cientistas seja justo e meça o mérito individual do cientista. Isso nos leva às perguntas que este manuscrito pretende responder: as métricas utilizadas atualmente são adequadas? Medem de fato o mérito de um cientista? Permitem que haja manipulação dos índices para obter uma produtividade e excelência científica aparente, distorcendo a avaliação?

Este manuscrito surgiu de um seminário que ministrei em um simpósio de pós-graduação da Universidade Federal de São João del-Rei em agosto de 2022, o que me levou a muita leitura e reflexão sobre avaliação dos cientistas. Neste manuscrito, pretendo compartilhar com a comunidade minhas impressões sobre este assunto e contribuir para melhorar os critérios de avaliação dos cientistas no país.

PERIÓDICOS: A LIMITAÇÃO DO FATOR DE IMPACTO

O fator de impacto (FI) foi introduzido em 1963 por Garfield e Sher, sendo reportado pela primeira vez em 1975 no *Journal of Citation Reports* (JCR),⁷ e se tornando uma verdadeira obsessão entre editoras dos periódicos, editores e pesquisadores nas últimas décadas. Todos os anos há uma expectativa no mês de junho para verificar os

*e-mail: pliego@ufsj.edu.br

novos índices e como cada periódico em que publicamos está em termos de performance. O FI é usualmente considerado uma medida real da importância e influência de um periódico. Mas será este o único indicador capaz de dizer sobre a importância média dos artigos de um periódico? Para responder a essa questão, vamos analisar a definição do FI de um periódico em um ano N : O FI é o número de citações que os artigos deste periódico publicados nos anos $(N-1)$ e $(N-2)$ receberam no ano N , dividido pelo número de artigos publicados nos anos $(N-1)$ e $(N-2)$. Esse índice é uma média de quanto um artigo publicado naquele periódico é citado a cada ano. Seria um bom índice se a citação anual medida em apenas 2 anos fosse a mesma recebida durante um longo período de tempo. Entretanto, esse não é o caso, e o padrão de citação muda conforme a área e mesmo subárea. Para ver isso melhor dentro da química, vamos analisar as citações recebidas pelos artigos de um grupo de periódicos num ano específico (2021), e verificar quanto tempo no passado esses artigos foram publicados.

A Figura 1a mostra o padrão temporal de citações aos artigos publicados em 4 periódicos da *American Chemical Society*, representando cada um uma das 4 subáreas da química, recebidas no ano de 2021. As citações nesse ano são feitas a artigos publicados em 2021 e em todos os anos anteriores. Observe que o gráfico mostra a porcentagem de citações para cada ano, de forma que a área sob cada curva integre para 100% (estendendo a curva para antes de 2012). Como exemplo, cerca de 10% das citações recebidas pelo periódico AC em 2021 foram feitas aos seus artigos publicados em

2019. Notamos que os periódicos de analítica (AC) e inorgânica (IC) têm um padrão de citação semelhante, com um grande aumento de citações advinda de publicações logo antes do ano de referência, caindo mais acentuadamente para publicações de anos mais distantes. No caso da orgânica (JOC), esta subida inicial é bem mais suave, e na físico-química (JPCA) mais suave ainda, ficando num platô durante o período apresentado. Tal gráfico mostra que as citações de analítica e inorgânica se concentram em artigos publicados em anos recentes, enquanto em orgânica e principalmente na físico-química, tendem a ser mais dispersas. Pelo fato da ordenada ser porcentagem de citação, grande parte das citações de físico-química e orgânica ocorre de artigos com mais de 10 anos, enquanto de analítica e inorgânica, há bem menos tempo. Quando se analisa o fator de impacto de 2021, ele mede as citações a artigos de 2019 e 2020 apenas, ou seja, os periódicos de analítica e inorgânica terão um aumento no seu fator de impacto em relação a orgânica e físico-química devido ao seu padrão de citações. Isso certamente é uma distorção séria no processo de ranqueamento do periódico pelo FI. Entretanto, a situação pode ficar ainda mais distorcida ao incluir mais periódicos.

Na Figura 1b, decidiu-se por incluir periódicos de assuntos que tem crescente importância tecnológica nos dias atuais: catálise ambiental (*Applied Catalysis B*) e energia (*Fuel*). Também se incluiu um dos mais prestigiados periódicos na área de físico-química e física molecular, o *Journal of Chemical Physics*. Como podemos notar, o *Applied Catalysis B* e o *Fuel* tem um crescimento vertiginoso na

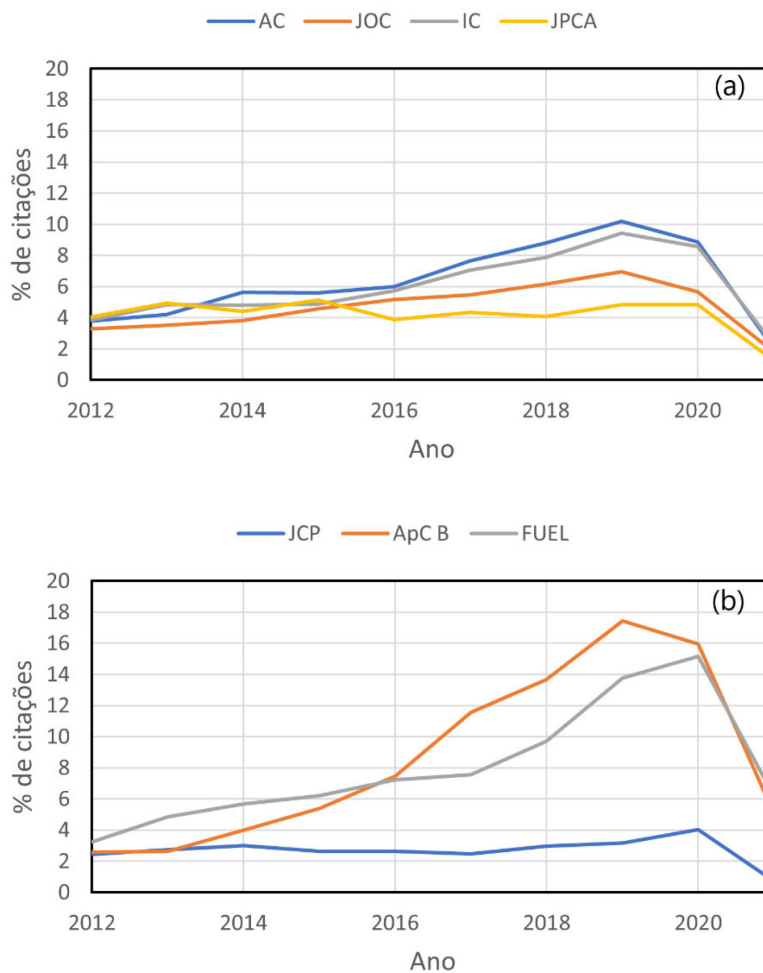


Figura 1. Citações recebidas por um periódico no ano de 2021, relativo a todos os artigos publicados neste periódico durante sua vida registrada no JCR. Está apresentada a porcentagem do total de citações feita ao conjunto de artigos publicados em cada ano. Siglas em (a): AC (Analytical Chemistry), JOC (Journal of Organic Chemistry), IC (Inorganic Chemistry), JPCA (Journal of Physical Chemistry A). Siglas em (b): JCP (Journal of Chemical Physics), ApC B (Applied Catalysis B: Environmental), FUEL (Fuel)

porcentagem de citações logo antes do ano de referência, contribuindo para elevar muito seus FI, e caem rapidamente logo após o pico, sugerindo que os artigos desses periódicos são citados durante um curto período de tempo. Por outro lado, o *Journal of Chemical Physics* tem um padrão de citação em que há um pequeno crescimento inicial, e um longo platô, indicando que os artigos lá publicados vão sendo citados durante muitos anos. Podemos notar que na Figura 1b a diferença de padrão de citação encontrado na Figura 1a é ainda mais exacerbado, mostrando que há áreas de pesquisa cuja citação é mais imediata, enquanto outras áreas as citações são modestas a cada ano, mas podem ficar durante décadas sendo citados. Tal fato não é captado pelo FI, o que distorce a qualidade real (medida por citação) de uma área de pesquisa e pode induzir os cientistas a buscar pesquisas mais imediatas para conseguirem ser mais bem avaliados. Com isso, trabalhos fundamentais e de mais longo prazo são prejudicados, com prejuízo para toda a ciência.

FATOR DE INFLUÊNCIA: INCORPORANDO A MEIA-VIDA DE CITAÇÕES

Uma forma de corrigir a limitação do fator de impacto é incorporar uma métrica que inclua por quanto tempo em média os artigos publicados em um periódico são citados. Tal métrica é disponível no JCR, sendo denominado *Cited Half-Life*, que traduzi como meia-vida de citações (MVC). O MVC em um ano específico é a mediana da idade dos artigos recebendo citações. Uma meia vida de 5 anos significa que metade das citações recebidas por um periódico no ano de referência foram de artigos publicados até 5 anos antes do ano de referência. Quanto maior a MVC, por mais tempo os artigos daquele periódico continuam a ser citados. A incorporação dessa métrica provê uma medida mais justa para se comparar periódicos.

Com o objetivo de propor uma métrica que leve em consideração tanto citações mais recentes, quanto citações de artigos publicados há mais tempo, é proposto o fator de Influência (FI_n), que é definido como:

$$FI_n = (FI \times MVC)/10 \quad (1)$$

Fazendo a multiplicação de ambos o FI e o MVC podemos balancear melhor estes diferentes padrões de citação. A divisão por 10 visa apenas resultar em um número menor. Deve-se mencionar que Kuo e Rupe propuseram um índice similar,⁸ que eles denominaram *Reliability-based citation impact factor*, ou *R-impact*.

Para ver o resultado dessa nova métrica composta, utilizamos os 7 periódicos da Figura 1 mais um oitavo, o *Journal of Catalysis*, para verificar a performance de um periódico com artigos mais fundamentais de catálise. Os resultados estão na Figura 2. Analisando o FI (Figura 2a), vemos que os periódicos de assuntos que estão em grande evidência no presente, como catálise, ambiental e energia tem maiores FI, enquanto as áreas mais fundamentais da química tem menor FI, destacando-se que a analítica ainda tem um bom FI, talvez por ter uma certa proximidade com a área ambiental. Quando olhamos a MVC (Figura 2b), a situação se inverte, e as áreas mais fundamentais de físico-química, orgânica e inorgânica se destacam. O periódico de catálise mais fundamental também se destaca, mas o de maior destaque é o *Journal of Chemical Physics*, reconhecidamente um periódico com muitos artigos fundamentais em físico-química e física molecular. Deve-se mencionar a forte relação do padrão de citação na Figura 1 e esses dados de FI e MVC da Figura 2. Os periódicos com um pico pronunciado nos gráficos da Figura 1 têm baixa MVC, e aqueles que tem um platô tendem a ter uma alta MVC.



Figura 2. Análise do FI (fator de impacto, gráfico a), MVC (meia-vida de citações, gráfico b) e FI_n (fator de influência, equação 1, gráfico c) de 8 periódicos. Dados do JCR do ano de 2021

O último gráfico (Figura 2c) mostra o FI_n, que combina as duas métricas. Nesse caso, vemos que os periódicos são muito mais comparáveis em performance, sem a forte distorção que vemos na Figura 2a. Na minha visão, a Figura 2c reflete de forma muito mais fiel minha própria percepção em relação a importância de cada periódico do que a Figura 2a. Ocorre que em 2c ambas as citações recentes, bem como as citações mais antigas, são contempladas, equilibrando melhor a avaliação do periódico. Com base nesse resultado, uma proposta é que o FI_n seja utilizado na avaliação do periódico, e não o FI. A presente análise aponta que o uso do FI gera uma forte distorção na avaliação, e não deveria ser a métrica a ser utilizada.

AUTORIA: A NECESSIDADE DA CONTAGEM FRACIONÁRIA

O desempenho do cientista é medido por quantos artigos ele publica, qual a qualidade do periódico, e quantas citações recebe, essa última geralmente expressa pelo índice h. Entretanto, a contagem feita atualmente não discrimina se o artigo tem 1 ou 10 autores, sendo um aspecto que tem sido severamente criticado na literatura.⁹⁻¹¹ É evidente que o artigo escrito inteiramente por 1 autor requer uma quantidade consideravelmente maior de trabalho por autor do que o autor de um artigo com 10 autores, principalmente se o autor em comparação não for o primeiro autor. Essa prática gera uma produtividade aparente, que não condiz com a produtividade real do autor em múltipla autoria. Para tornar isso mais claro, vamos definir contagem inteira e contagem fracionária de autoria:

Contagem inteira da autoria: crédito total a cada autor pelo artigo e pelas citações, como se cada um tivesse feito o trabalho sozinho.

Contagem fracionária da autoria: o crédito pelo artigo e pelas citações é feito levando em conta o número de autores no artigo, dividindo o crédito pelo número de autores.

Para ver de forma clara como a autoria inteira distorce a produtividade real, vamos observar a Figura 3. Temos dois grupos de 5 autores, A e B. No grupo A, cada autor produz individualmente 5 artigos. A produção total do grupo A é 25 artigos. No grupo B, cada autor produz 5 artigos em coautoria com os demais, de forma que a produção total do grupo B é 5 artigos. Com esses dados, é evidente que o grupo A é cinco vezes mais produtivo do que o grupo B. O mesmo devemos esperar dos autores de A em relação a B. Entretanto, se fizermos a contagem inteira de autoria, ambos os autores do grupo A e do grupo B são igualmente produtivos, pois cada um publica 5 artigos. Essa distorção é evidente! No entanto, se usarmos a contagem fracionária, cada autor do grupo A publica 5 artigos, enquanto cada autor do grupo B publica apenas 1 artigo. Essa é a contagem correta! Se for perguntado a um gestor de fundos onde os recursos de pesquisa seriam melhor aplicados, quais pesquisadores ou qual grupo ele escolheria? Essa análise mostra que estamos avaliando incorretamente os cientistas, pois a soma de FI das publicações e o índice h não leva em conta a autoria fracionária, e sim a autoria inteira.

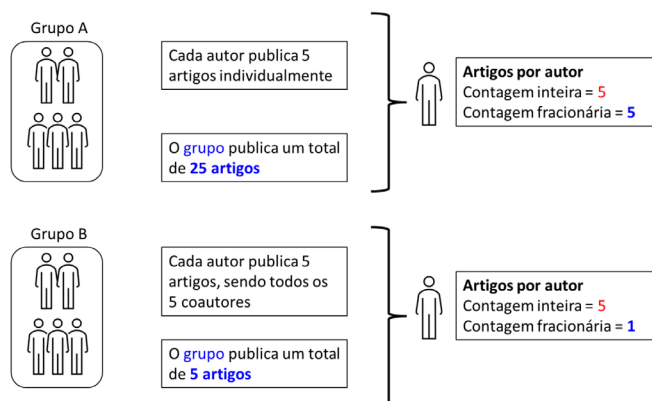


Figura 3. Comparação entre contagens inteira e fracionária de autoria, mostrando como a contagem inteira distorce a produtividade real

Desde pelo menos 1980 tem sido apontado que o número de autores em ciência cresce proporcionalmente mais do que o número de artigos.¹² No caso da química no Brasil de 2012 a 2021 (Figura 4), também vemos um aumento maior no número de autores (104%) do

que de publicações (72%). A média de publicações por autor por ano diminuiu de 0,354 em 2012 para 0,298 em 2021. Estes dados indicam que a nossa produtividade real está caindo.

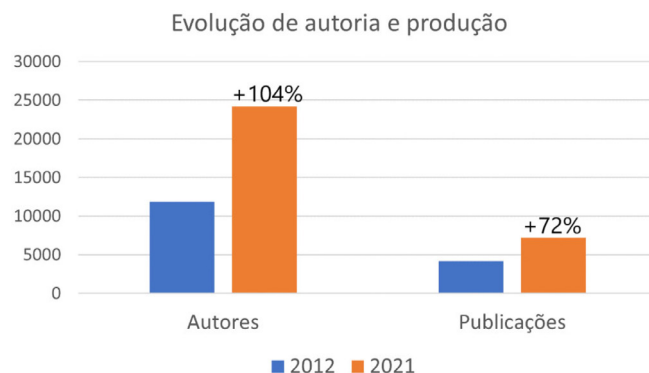


Figura 4. Documentos publicados na área de química por pesquisadores brasileiros em 2012 e 2021, e número de autores. Dados da base SciVal/Scopus

Na visão deste autor, o uso da contagem inteira, associada a disputa por recursos, é um fator que está levando a um aumento no número de autores por artigo e a diminuição da produtividade real. Embora se possa argumentar que a ciência atualmente está mais colaborativa, e isso resultaria em mais autores por artigo, é certo também que muitos pesquisadores estão se juntando em grupos maiores para publicar mais, aumentando assim a produtividade aparente, uma vez que é essa a utilizada para julgar o cientista, e não a produtividade real. A recente divulgação de um documento do comitê assessor da área de química do CNPq, apontando mecanismos de inflação de publicações em candidatos a bolsa de pós-doutorado, é uma evidência de que o uso da contagem inteira de autoria está induzindo a tais práticas. Nesse esquema, é fácil manipular os números. Devo enfatizar aqui a necessidade de mudança. Precisamos incorporar autoria fracionária na avaliação dos cientistas.

A OBSOLESCÊNCIA DO ÍNDICE H

O índice h foi proposto em 2005 por J. Hirsch e se tornou referência para ranquear cientistas.¹³ O índice h de um cientista é definido como correspondendo a ter a h artigos com h ou mais citações. A ideia é que um bom pesquisador precisa ter muitos artigos com muitas citações, ao invés de poucos artigos com muitas citações. Dessa forma, o índice h foi proposto como uma métrica melhor do que o total de citações. Entretanto, o índice h não leva em consideração coautoria, de forma que ao se juntar em grupos maiores, os pesquisadores podem inflar artificialmente o índice h, da mesma forma como o número de artigos. Como discutido na Figura 3, não tem sentido um autor que divide a autoria com outros 4 autores ficar com todo o crédito de citações para si. Outros autores como Vavrycuk⁵ já apontaram que “the primary goal of ranking is to evaluate the scientific credit of researchers but not their collaboration abilities”.

Em um artigo recente, Koltun e Hafner mostraram que o índice h não tem mais uma correlação efetiva com reputação científica,⁴ um problema atribuído por eles a hiperautoria nos artigos. Eles utilizaram um grande conjunto de dado de biologia, física, economia e computação e consideraram os 1000 cientistas mais citados de cada área. Entre seus achados, eles observaram que a correlação do índice h com prêmios em física caiu de 0,34 em 2010 para 0,00 em 2019 (correlação τ de Kendall). Entretanto, quando um índice h fracionário (h_f) foi usado, em que o número de autores nos artigos é considerado, esse declínio observado no índice h foi mitigado. Em suas palavras: “fractional allocation improves the effectiveness and

predictive power of research metrics, and h-frac is consistently the most reliable bibliometric indicator”.

Outro estudo que também aponta a falha crescente do índice h foi publicado por Ioannidis e colaboradores.⁶ Esses autores coletaram dados dos 84116 mais influentes cientistas em 12 campos de estudo, classificados por um indicador composto a ser discutido adiante neste texto. Eles observaram que o índice h tinha uma forte correlação com o total de citações ($R = 0,88$), justamente uma propriedade que o índice h não deveria apresentar, pois a ideia desse índice era substituir o total de citações como métrica de influência científica. De maneira similar ao estudo de Koltun e Hafner, quando um índice h que leva em consideração o número de autores foi utilizado (h_m , diferente do h_i),¹⁴ a correlação dele com o total de citações foi $R = 0,00$, indicando ser um índice muito melhor do que o índice h nesse aspecto. Outro resultado importante foi que dos 1000 cientistas mais influentes segundo o total de citações,⁶ 122 deles não tinha um único artigo como único, primeiro ou último autor, sugerindo que apesar de serem altamente citados, não eram cientistas líderes. Esses resultados apontam de forma clara para necessidade de observar mais atentamente a autoria, não só o número de autores, mas também a posição de autoria nos artigos.

ÍNDICES H FRACIONÁRIOS: MEDINDO MAIS REALISTICAMENTE O MÉRITO INDIVIDUAL DOS CIENTISTAS

Com a percepção de que o índice h falha ao não levar em consideração a coautoria, várias métricas derivadas do índice h incorporando esse aspecto foram propostas. Uma primeira proposta foi feita por Batista e colaboradores em 2006.¹⁵ Eles propuseram o índice h_i (1 maiúsculo), definido por:

$$h_i = \frac{h}{\langle N_a \rangle} \quad (2)$$

Sendo $\langle N_a \rangle$ o número médio de autores nos h artigos usados na computação do índice h . A análise destes autores indicou que o índice h_i permitiria uma melhor comparação entre áreas, como física, química, matemática e biologia/biomédicas. Logo depois em 2008, Schreiber propôs o índice h_m (m de múltipla autoria).¹⁴ Nesse índice, a coautoria é incluída por colocando os artigos em ordem decrescente de citação, mas contando cada artigo como $1/N_a$, sendo N_a o número de autores em cada artigo. Avaliando 8 físicos com alto índice h , a razão h_m/h variou de 0,58 a 0,86, indicando que o ranqueamento dos cientistas pode mudar bastante com a inclusão de coautoria.

Outro índice h fracionário foi proposto por Egghe,¹⁶ denominado h_f (equivalente ao h_i , i minúsculo, de individual), em que a citação de cada artigo (C) é dividida pelo número de autores, ou C/N_a . Isso é feito para cada artigo e usado de maneira similar ao índice h para definir o índice h_f . Em termos operacionais, colocamos os artigos em ordem decrescente de C/N_a , atribuímos uma numeração de cada artigo pelo inteiro r , em ordem crescente, começando por 1 e aumentando de 1 em 1, e escolhemos o maior inteiro r na sequência de artigos tal que:

$$h_i = h_f = \text{maior } r \leq \frac{C(r)}{N_a(r)} \quad (3)$$

Egghe também discutiu o índice h_m , o qual denominou de h_p . Seu trabalho foi teórico e o índice h_i somente foi avaliado com dados concretos mais recentemente, por Koltun e Hafner.⁴ Esses autores mostraram que esse índice é superior aos índices h_i e h_m , e muito superior ao índice h na correlação com prêmios científicos.

Outra investigação do índice h_i foi reportada este ano por Henry H. Bi.³ O autor avaliou o desempenho de 12 prêmios Nobel com relação

ao índice h e h_i . Encontrou-se que a razão h_i/h variou de 0,27 até 0,89, indicando que a introdução de coautoria pode mudar substancialmente o ranqueamento dos cientistas. Observou-se também que quanto mais coautores, maior a tendência de diminuição do h_i em relação a h .

O índice h_i não está disponível até o momento nas plataformas *Web of Science* e *Scopus*. Entretanto, está disponível no programa “*Publish or Perish*”,¹⁷ que pode ser baixado e utilizado de forma gratuita. O programa pode acessar dados do Google Scholar e gerar os índices h e h_i (denominado de $h_{i, \text{norm}}$), bem como várias outras métricas.

Para termos uma ideia do efeito da coautoria no índice h dos químicos brasileiros, realizou-se uma análise com uma amostra de 15 pesquisadores, escolhidos aleatoriamente entre os 500 mais produtivos entre os anos de 2012 e 2021, com dados da base Scopus. Nessa análise, utilizou-se o h_i pelo fato dos estudos citados apontarem este índice como o melhor para ranquear cientistas. Utilizou-se o software *Publish or Perish*,¹⁷ o qual foi usado na computação do h_i . Essa computação baseou-se nos dados de publicações e citações dos pesquisadores disponíveis no Google Scholar. Os resultados podem ser vistos na Tabela 1. Observa-se que na amostra estão pesquisadores com índice h de 15 a 67, uma variação bem considerável. No caso do número médio de autores por artigo, ficou entre 4,16 e 5,85, um intervalo não muito grande. Como esperado, o índice h_i é compactado em relação ao índice h , como podemos ver na coluna h_i/h . Essa razão varia bastante, estando a maior em 0,56 e a menor em 0,31. Tal variação pode causar uma mudança drástica na classificação dos pesquisadores. Por exemplo, o pesquisador 3 perderia 3 posições se fosse usado o índice h_i , enquanto o pesquisador 12 poderia subir 5 posições, empatando com o pesquisador 8. Tal variação pode certamente impactar significativamente em processos decisórios de distribuição de recursos.

Tabela 1. Métricas para um grupo de 15 pesquisadores altamente produtivos da área de química escolhidos aleatoriamente^a

Autor	h	h_i	maa	h_i/h
1	67	33	4,16	0,49
2	58	22	5,17	0,38
3	51	16	4,94	0,31
4	46	21	5,16	0,46
5	45	20	4,56	0,44
6	40	13	5,85	0,33
7	35	18	4,16	0,51
8	34	15	4,78	0,44
9	33	12	4,30	0,36
10	33	11	5,47	0,33
11	27	10	5,50	0,37
12	27	15	4,28	0,56
13	24	8	5,10	0,33
14	23	9	4,62	0,39
15	15	5	5,13	0,33

^aCientistas brasileiros da área de química sorteados entre os 500 mais produtivos de 2012 a 2021 pela plataforma Scopus. Os dados h , h_i e maa (média de autores por artigo) foram obtidos com o programa *Publish or Perish*, usando dados do Google Scholar.

Para obter uma relação mais quantitativa do efeito de múltipla autoria, foi feita uma regressão linear da relação entre h_i e a razão h/h_i (maa), sendo maa = média de autores por artigo. Podemos observar no gráfico da Figura 5 uma ótima correlação, com $R^2 = 0,92$ e a equação obtida foi:

$$h_i = (2,0499) \frac{h}{maa} - 0,6977 \quad (4)$$

Essa equação vale apenas no intervalo testado, com maa entre 4 e 6, mas é suficiente para mostrar como o número de autores impacta de forma significativa o h_i . Na média, com 4 autores o índice h_i é 1/2 do índice h , enquanto com 6 autores essa relação cai para 1/3. Esse resultado é importante, pois mostra que o índice h_i tende a corrigir a inflação de autores, sendo um desincentivo a formação de grupos para inflar os índices de avaliação.

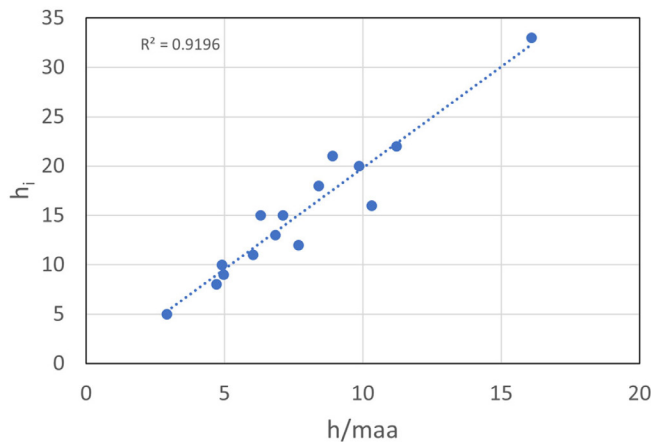


Figura 5. Correlação entre h_i e a razão h/maa , sendo maa a média de autores por artigo

ÍNDICES COMPOSTOS: INCORPORANDO DETALHES DO PERFIL

O uso de uma única métrica no ranqueamento de cientistas é atraente pelo fato produzir um único número que pode ordenar os mesmos. Entretanto, é desejável que múltiplos aspectos da autoria seja avaliada, olhando mais detalhadamente no perfil do pesquisador. Por exemplo, sabemos que usualmente o primeiro autor é quem mais contribuiu para um trabalho, e seria útil saber em quantos artigos de sua autoria um cientista é o primeiro autor, assim como quantas citações estes artigos receberam. Da mesma forma, o último autor é geralmente o idealizador do projeto e chefe do grupo, tendo também um lugar de destaque na publicação. Seria interessante saber quantos artigos o pesquisador é último autor e quantas citações recebeu. Por outro lado, em um artigo com muitos autores, um dos autores do meio provavelmente teve uma contribuição pequena, talvez fornecendo uma análise técnica específica, e muitas vezes nem saberia explicar o conteúdo do artigo. Um perfil com muitos artigos desse tipo pode significar falta de liderança e independência científica. Pensando nesses aspectos, Ioannidis e colaboradores propuseram um índice composto para ranquear os cientistas, que ficou conhecido como *Stanford ranking*.⁶ O ranking tem sido atualizado em anos seguintes.^{18,19} No índice composto, eles consideraram 6 métricas:

NC = total de citações

NS = total de citações (artigos como único autor)

NSF = total de citações (artigos como único ou primeiro autor)

NSFL = total de citações (artigos como único, primeiro ou último autor)

H = índice h

Hm = índice h ajustado pela coautoria, o h_m

O índice composto, nomeado como C, foi obtido pela expressão:

$$C = \frac{\ln(NC+1)}{\ln(NC_{max}+1)} + \frac{\ln(NS+1)}{\ln(NS_{max}+1)} + \frac{\ln(NSF+1)}{\ln(NSF_{max}+1)} + \frac{\ln(NSFL+1)}{\ln(NSFL_{max}+1)} + \frac{\ln(H+1)}{\ln(H_{max}+1)} + \frac{\ln(H_m+1)}{\ln(H_{m,max}+1)} \quad (5)$$

na qual podemos notar que os índices são log-normalizados, com o maior valor encontrado (max) usado como referência. Dessa forma, cada índice contribui com no máximo uma unidade e o valor de C pode ir de zero até seis. Ioannidis e colaboradores analisaram um grande número de dados de mais de 6 milhões de cientistas de áreas diversas, refinando para 84116 cientistas por esse indicador composto, os quais seriam os mais influentes de seus respectivos campos de pesquisa.

Na análise deste indicador C, os autores reportaram uma baixa correlação entre H e C ($R = 0,25$), indicando novamente que o índice h não é uma boa métrica atualmente. Entretanto, ao analisar a correlação entre Hm e C, observou-se que havia uma ótima correlação ($R = 0,92$), novamente sugerindo que um índice h fracionário é mais adequado para ranquear cientistas. Eles estenderam a análise para verificar a habilidade dos indicadores em capturar cientistas com Nobel, analisando os 14150 cientistas mais bem ranqueados por C, H, Hm e NC. O índice C consegue capturar 31 dos 47 cientistas com Nobel, enquanto apenas 18, 26 e 15 foram capturados por H, Hm e NC. É evidente a melhor performance do índice C, e ainda notamos novamente como h_m , um índice fracionário, supera o índice h. Sumarizando, os autores mostraram que o uso de um indicador composto como utilizado por eles provê uma informação mais completa do perfil do cientista, e possibilita uma avaliação mais realista da performance do mesmo.

CONSIDERAÇÕES FINAIS

O uso indiscriminado de métricas na classificação de cientistas e periódicos tem causado bastante preocupação, principalmente pelo uso inadequado delas. Um grupo significativo de pesquisadores tem alertado sobre suas limitações, e a importância de se conhecer a informação que cada métrica fornece.²⁰ O manifesto de Leiden estabelece 10 princípios para guiar o uso de métricas na avaliação científica.²⁰ Um deles é justamente o constante escrutínio da métrica e seu melhoramento. Isso é o que este documento pretende: que nossas métricas, especialmente na química, sejam atualizadas, após as análises aqui apresentadas, e mais discussões na comunidade.

AGRADECIMENTOS

O autor agradece ao CNPq pelo suporte.

REFERÊNCIAS

- Rodríguez-Navarro, A.; Brito, R.; *Scientometrics* **2022**, *127*, 2871. [Crossref]
- Abbott, A.; Cyranoski, D.; Jones, N.; Maher, B.; Schiermeier, Q.; Noorden, R. V.; *Nature* **2010**, *465*, 860. [Crossref]
- Bi, H. H.; *Scientometrics* **2022**. [Crossref]
- Koltun, V.; Hafner, D.; *PLoS One* **2021**, *16*, e0253397. [Crossref]
- Vavrycuk, V.; *PLoS One* **2018**, *13*, e0195509. [Crossref]
- Ioannidis, J. P.; Klavans, R.; Boyack, K. W.; *PLoS Biol.* **2016**, *14*, e1002501. [Crossref]
- Larivière, V.; Sugimoto, C. R. In *Springer Handbook of Science and Technology Indicators*; Glänzel, W., Moed, H. F., Schmoch, U., Thelwall, M., eds.; Springer International Publishing: Cham, 2019, p. 3.
- Kuo, W.; Rupe, J.; *IEEE Trans. Reliab.* **2007**, *56*, 367. [Crossref]
- Poder, E.; *Front. Res. Metr. Anal.* **2021**, *6*, 824518. [Crossref]

10. Lichtman, M. A.; *Blood Cells Mol. Dis.* **2020**, *84*, 102454. [Crossref]
11. Pöder, E.; *J. Am. Soc. Inf. Sci. Technol.* **2010**, *61*, 2593. [Crossref]
12. Persson, O.; Glänzel, W.; Danell, R.; *Scientometrics* **2004**, *60*, 421. [Crossref]
13. Hirsch, J. E.; *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 16569. [Crossref]
14. Schreiber, M.; *New J. Phys.* **2008**, *10*, 040201. [Crossref]
15. Batista, P. D.; Campitelli, M. G.; Kinouchi, O.; Martinez, A. S.; *Scientometrics* **2006**, *68*, 179. [Crossref]
16. Egghe, L.; *J. Am. Soc. Inf. Sci. Technol.* **2008**, *59*, 1608. [Crossref]
17. Harzing, A.-W.; *Publish or Perish*; Version 8; 2021.
18. Ioannidis, J. P. A.; Boyack, K. W.; Baas, J.; *PLoS Biol.* **2020**, *18*, e3000918. [Crossref]
19. Ioannidis, J. P. A.; Baas, J.; Klavans, R.; Boyack, K. W.; *PLoS Biol.* **2019**, *17*, e3000384. [Crossref]
20. Hicks, D.; Wouters, P.; Waltman, L.; de Rijcke, S.; Rafols, I.; *Nature* **2015**, *520*, 429. [Crossref]