

Debate sobre o artigo de Camargo Jr.
 Debate on the paper by Camargo Jr.
 Debate acerca del artículo de Camargo Jr.

<http://dx.doi.org/10.1590/0102-311XCO010913>

Rita Barradas Barata

Faculdade de Ciências Médicas
 da Santa Casa de São Paulo,
 São Paulo, Brasil.
rita.barradasbarata@gmail.com

Medir ou classificar a produção científica de pesquisadores?

A questão de como avaliar o desempenho de pesquisadores tem ocupado espaço crescente na reflexão e nas discussões de agências de fomento, editores científicos, gestores de política científica e dos próprios pesquisadores em todo o mundo.

Aparentemente há um consenso na comunidade de que se deva adotar algum tipo de avaliação. Porém, o consenso parece terminar aí. A forma de realizar essa avaliação apresenta inúmeras controvérsias como bem apontado no artigo ora comentado.

Para início de conversa creio ser importante estabelecer algumas precisões. Embora seja frequente a crítica aos indicadores bibliométricos por serem médias e, portanto, não servirem para representar distribuições tão assimétricas, a rigor, a maioria deles não pode ser definida como tal. O fator de impacto, por exemplo, embora seja descrito como a média de citações recebidas por um periódico, matematicamente falando é uma razão, visto que relaciona duas quantidades distintas: no numerador, o número de citações recebidas em um determinado ano e, no denominador, o número de artigos ou matérias citáveis publicadas nos últimos dois anos. O índice *h* tão-pouco é uma média, correspondendo a uma posição na qual dois *rankings* se encontram: o número acumulado de artigos publicados por um autor e o número de citações recebidas por estes artigos. Do mesmo modo, se considerarmos artigos individualizados ao invés de periódicos não podemos falar em fator de impacto, mas sim no número de citações recebidas por um determinado artigo em um certo período de tempo. Portanto, a questão não está nas médias nem na forma assimétrica das distribuições de interesse.

Outra questão que me parece mais relevante é a permanente confusão entre o significado dessas medidas e os usos que delas se tem feito nos mais diferentes contextos e para as mais diversas finalidades.

Realmente, nenhum dos proponentes desses índices considera que eles sejam capazes de

mensurar a qualidade do trabalho que está sendo feito. O que se espera indicar por meio dessas medidas é a importância ou a possibilidade de um artigo ou periódico influenciar a produção de conhecimentos em determinada área.

Não me parece que seja possível afirmar que a qualidade da Matemática como ciência é inferior à qualidade da Imunologia porque os fatores de impacto de periódicos ou o número de citações por artigo seja muito menor na primeira área.

Há uma nítida confusão entre medidas indiretas de circulação e prestígio/influência e qualidade que não me parece autorizada sob nenhum aspecto.

Para tornar ainda mais complexa a (má)utilização dos indicadores bibliométricos há uma série de aspectos que costumam ser ignorados por aqueles que os utilizam que, no entanto, modificam amplamente a interpretação dos resultados. Em relação ao fator de impacto, por exemplo, há que se considerar a amplitude das bases bibliométricas consideradas em seu cálculo e quanto da produção científica das diferentes áreas de conhecimento está nelas representada; língua, tradição de citação e de autoria em diferentes comunidades, e assim por diante.

Em artigo recente, a ser publicado pela *Revista de Saúde Pública*, Barreto et al.¹ tiveram a oportunidade de demonstrar que o índice *h* de pesquisadores de diferentes áreas das ciências da saúde sofrem o efeito direto da base de citações consideradas em seu cálculo. Para periódicos, todos sabemos que o índice *h* é fortemente influenciado pelo tempo de existência e indexação nas bases correspondentes.

Diante de tantas limitações, seria o caso de se perguntar por que tais medidas continuam a ser amplamente utilizadas para muitos propósitos, em todas as partes do mundo.

Aqui também muitas respostas seriam possíveis. Desde as que enfatizariam as escolhas ideológicas até aquelas que destacariam a questão mais prosaica de que eles são usados porque estão disponíveis.

Poderíamos nos perguntar por que esses índices são tão atrativos. Seria apenas o fetiche dos números e a ideia de objetividade que eles carregam? Em que medida grupos de avaliadores se sentiriam mais confortáveis e, até, relativamente eximidos de emitir juízos de valor, transferindo aos índices e seus *rankings* a tarefa espinhosa de escolher/deliberar sobre a avaliação de pesquisadores, institutos ou instituições?

Nas condições em que se realizam as avaliações e os julgamentos de editais, me parece plenamente compreensível que tais medidas, ainda que tão problemáticas, encontrem seu lugar.

Será que a substituição dessas formas tão imperfeitas de mensuração por classificações qualitativas é a solução?

Creio que a chamada avaliação qualitativa tem também suas próprias dificuldades e inúmeros desafios.

Para ficar apenas no âmbito da avaliação de pesquisadores individuais, embora pareça tentadora a ideia de que o próprio autor determine quais são as suas produções mais significativas e as ofereça à avaliação por pares, na prática não parece tão simples.

Todos os editores científicos e os coordenadores de processos de avaliação de projetos para o fomento têm enfrentado enormes dificuldades com a revisão por pares. Cada vez torna-se mais difícil conseguir avaliadores e avaliações bem feitas que possam ser realmente úteis para as comissões encarregadas da aprovação ou priorização das propostas.

Quando as avaliações envolvem projetos de diferentes áreas e subáreas como é possível hierarquizar as propostas? Como avaliar projetos de áreas ou subáreas que se encontram em diferentes estágios de desenvolvimento? Será que uma produção considerada de excelência em uma determinada área pode ser comparada com outra, avaliada em uma área distinta?

Em campos como a Saúde Coletiva, com uma enormidade de objetos, abordagens, correntes teóricas, temáticas, posicionamentos ideológicos, parece extremamente desafiador a proposição de avaliações classificatórias.

Como garantir que os comitês julgadores sejam adequadamente constituídos para dar conta dessa diversidade sem cometer grandes injustiças?

Enfim, todas essas reflexões foram aqui apresentadas apenas no intuito de problematizar uma questão que, me parece, está muito longe de ser adequadamente equacionada.

Nós todos que nos interessamos pela ciência como objeto de investigação temos muito com o que nos ocuparmos. Sem contar com o enorme desafio que significa tentar avaliar o real impacto da ciência e de seus produtos na melhoria das condições concretas de existência da humanidade.

<http://dx.doi.org/10.1590/0102-311XCO020913>

Paulo Vaz

*Escola de Comunicação,
Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil.
paulorgvaz@uol.com.br*

Por um novo jogo

Como a pesquisa no Brasil depende quase que totalmente de financiamento público, a distribuição de verbas tende a utilizar formas ordenadas de avaliação. É preciso responder prévia e estruturalmente, por meio da adoção de procedimentos, a possíveis denúncias de injustiça nos julgamentos das agências de fomento.

Visando a assegurar legitimidade e eficiência na alocação de recursos escassos, nos últimos 20 anos, dois tipos maiores de indicadores quantitativos paulatinamente dominaram as avaliações: um calcado na hierarquização dos veículos nos quais as pesquisas são publicadas, principalmente a classificação de revistas científicas, e o outro na repercussão do trabalho do pesquisador, mensurada pelas citações que recebeu.

O artigo de Keneth Rochel de Camargo Jr. propõe que é hora de avaliar o uso praticamente absoluto dessas duas métricas nas avaliações das ciências no Brasil. A ênfase do texto recai na exposição e crítica de algumas das razões que nos levam a confiar nesses indicadores. Dentre as razões, podem ser destacadas a crença nos números como garantidor da objetividade, a facilidade de construção de consenso e a possibilidade de os juízes se eximirem do risco de julgar, isto é, de se eximirem da necessidade de justificar suas escolhas perante seus pares.

O artigo também aponta brevemente as consequências negativas. As avaliações, hoje, quando se limitam ao uso dessas métricas, tendem a não ser capazes de apreender a qualidade e favorecem a consolidação de um ambiente marcado pelo privilégio da comunicação em detrimento da informação, ou seja, pelo privilégio de publicar em detrimento de realizar pesquisas de qualidade. Esse “produtivismo” também tende a médio prazo a esgarçar as comunidades científicas, pois o volume de publicações e a necessidade de publicar podem fazer com que os pesquisadores especializem cada vez mais suas leituras e tenham muita dificuldade para saber o que é relevante e inovador devido à deterioração da relação entre sinal e ruído.

Este comentário irá se deter sobre o produtivismo, essa consequência negativa global. Não é difícil compreender por que esses índices foram criados. Ter seu trabalho publicado nas melhores revistas de uma determinada área significaria, em princípio, ter enfrentado uma competição mais dura e ter sido avaliado com mais rigor por pares do que seria o caso numa revista menos importante. E ter sua pesquisa citada é o modo

1. Barreto ML, Aragão E, Souza LEPE, Santana T, Barata RB. Diferenças entre as medidas do índice-h geradas em distintas fontes bibliográficas e engenhos de busca. *Rev Saúde Pública*; no prelo.