



Métricas na ciência

Metrics in science

Cada vez mais estamos cercados de métodos de avaliação de *performance* em ciência. Nem sempre estes métodos, embora matematicamente corretos, são capazes de refletir a realidade. Estes índices avaliam autores, periódicos, classificam estudos. O índice de Hirsch, ou índice h, foi desenvolvido em 2005 pelo físico J. E. Hirsch na Universidade da Califórnia para verificar o impacto cumulativo e a relevância da produção de pesquisas de um indivíduo¹; ou seja, o objetivo do índice h é quantificar a produção científica de um indivíduo por relevância.

A melhor forma de traduzir este índice é por meio do seguinte exemplo: o índice h do indivíduo será igual a 10 se o pesquisador possui 10 artigos que tenham sido citados pelo menos 10 vezes cada um. O efeito é logarítmico, ou seja, para passar de 1 para 2 é muito mais fácil que de 11 para 12, pois neste caso, o número de citações necessárias será bem maior. Existem alguns limites em relação a este índice². O número de autocitações pode aumentar o índice h. Este número pode ser expurgado em algumas bases de dados, mas se isso for feito, perde-se a força de se medir a evolução de uma linha de pesquisa. Foram também encontrados casos de citação propositadamente cruzada entre autores com a finalidade de se alavancar os índices dos autores envolvidos, o que não é correto.

Embora existam outras formas de se avaliar um determinado autor em relação à sua produção científica como: 1) avaliar o número total de artigos publicados; 2) avaliar o número total de citações; 3) avaliar isoladamente o número de citações por artigo; ou ainda 4) isolar o número total de artigos de grande número de citações produzidos por este autor; ainda assim o índice h representa de forma mais global a relevância da produção científica do autor.

Outra limitação é que este índice deve ser utilizado apenas para comparar autores da mesma área ou do mesmo campo de ciência. Os especialistas em Cirurgia Plástica, por exemplo, poderiam ser prejudicados quando comparados aos cardiologistas, área que apresenta uma quantidade maior de revistas indexadas e, portanto, têm uma maior probabilidade de serem citados nestes periódicos sem que estes artigos tenham um valor científico maior. Outra potencial desvantagem do índice h é que esta é uma métrica que não dá relevância a artigos altamente citados. De fato, o artigo mais citado de Hirsch, no qual este autor descreveu o índice h, teve 2196 citações de acordo com a base de dados Scopus enquanto o seu índice h é “apenas” 45 (em março de 2015). Este fato demonstra que o índice h não é perfeito, mas fornece uma ideia geral da produção de um pesquisador.

O Fator de Impacto (FI) de um periódico é medido pela relação entre o número total de artigos deste jornal citados em periódicos nos últimos 2 anos e o número total de artigos publicados no jornal no mesmo período. Ou seja, quanto maior a citação dos artigos relativamente ao número total de artigos, maior será o FI. Certamente, um autor irá preferir ter seu estudo publicado numa revista que terá maior visibilidade. Entretanto, nem sempre um periódico com maior visibilidade terá o maior impacto. Mas, seguramente, todo o autor deseja que seu artigo tenha uma boa repercussão no meio médico, o que pode ser refletido pelo número de citações de determinado artigo. Mas o que faz o artigo ser mais ou menos citado? Isso irá depender da área em que o artigo foi escrito, da qualidade do artigo, mas também de onde foi publicado.

Definitivamente, existem periódicos que não estão em bases de dado com ampla divulgação e, entre os de uma mesma base de dados, existem periódicos mais ou menos influentes. E é aí que entra o FI. Um periódico será, teoricamente, mais influente, na medida em que seus artigos terão maior citação naquela base de dados. Isso é teórico, pois o número de citações do artigo dependerá, basicamente, mais do artigo do que do periódico em si³. Certamente, revisores de periódicos de alta qualidade exigirão um conteúdo melhor

durante a avaliação do artigo que revisores de revistas de menor qualidade; portanto, o fato do artigo ter sido aceito por uma revista de alta qualidade é um bom sinal, mas não é tudo.

Foram encontrados vários defeitos e vulnerabilidades neste índice. Um deles, que já citamos anteriormente, é o fato de que alguns poucos artigos podem elevar o FI de forma importante, enquanto outros nem seriam citados. Uma forma de se aumentar artificialmente o FI é liberar para a publicação artigos de melhor qualidade e com capacidade de gerar alto número de citações no início de janeiro, já que o FI será calculado a partir deste mês e aumenta a possibilidade de citação deste artigo no período. Alguns editores incentivam a criação de suplementos sobre assuntos cujos artigos tradicionalmente tenham um alto número de citação. A citação cruzada, acordada entre editores de revistas, e o fato de alguns editores estimularem os autores a citarem artigos recentes publicados na própria revista são passíveis de punição por parte das bases de dados, o que já ocorreu no passado.

Uma outra forma de tentar qualificar os artigos é a classificação em tipos de estudo de acordo com níveis, usando os princípios da Medicina Baseada em Evidência (MBE). Os artigos são classificados numa pirâmide em 5 tipos de artigos, de acordo com seu nível de evidência. Esta classificação estimula os cirurgiões plásticos a melhorarem o nível de produção científica, já que os próprios autores devem fazer uma classificação do estudo durante a submissão do artigo.

Como sempre, existem os dois lados da moeda. Embora esta classificação seja excelente para artigos novos, quando a MBE é utilizada em revisões sistemáticas, o sucesso não é sempre alcançado⁴. As áreas cirúrgicas e de medicina intervencionista estão em constante evolução. Avaliar uma técnica cirúrgica ou uma técnica intervencionista ao longo de vários anos, pode não ser adequado, já que técnicas podem ser melhoradas e materiais refinados. Isso tudo pode levar a uma diferença de resultados ao longo dos anos. Quando é realizada uma meta-análise sobre uma

técnica, pode-se ter estudos muito bem realizados com nível 1 de evidência na década de 70, 80 e no ano de 2015. No final teremos um estudo (meta-análise) que será considerado como nível 1 de evidência, com resultados pouco confiáveis, já que as técnicas foram realizadas com uma grande diferença de tempo e, portanto, com diferentes técnicas e materiais.

Existem muitos outros índices que buscam avaliar a repercussão de um artigo, mas estes são os que mais têm sido utilizados. Alguns índices estão buscando avaliar artigos baseado em acessos na *web*, como o Altmetrics, mas podem ser criticados do ponto de vista científico, pois o acesso à internet é universal e não apenas restrito ao ambiente médico.

Estas métricas merecem uma análise especialmente voltada à cirurgia plástica, mais aprofundada, para que possamos ter uma ideia do que está ocorrendo nas revistas da especialidade, especialmente em relação à qualidade da produção científica de outras áreas. Avaliações são sempre importantes guias para se ter uma ideia de como e onde estamos.

REFERÊNCIAS

1. Hirsch JE. An index to quantify an individual's scientific research output. *Proc Natl Acad Sci U S A*. 2005;102(46):16569-72. DOI: <http://dx.doi.org/10.1073/pnas.0507655102>
2. Tedesco ACB, Nahas FX, Ferreira LM. The Hirsh (h) Index. *Plast Surg Pulse News*. 2010;2(3).
3. Rodrigues MA, Tedesco AC, Nahas FX, Ferreira LM. Journal impact factor versus the evidence level of articles published in plastic surgery journals. *Plast Reconstr Surg*. 2014;133(6):1502-7. DOI: <http://dx.doi.org/10.1097/PRS.0000000000000214>
4. Nahas FX, Ferreira LM. Evidence-based medicine: the past and the future of plastic surgery. *Plast Reconstr Surg*. 2014;134(3):499e-500e. DOI: <http://dx.doi.org/10.1097/PRS.0000000000000474>

FABIO XERFAN NAHAS

Universidade Federal de São Paulo, São Paulo, SP, Brasil

LYDIA MASAKO FERREIRA

Universidade Federal de São Paulo, São Paulo, SP, Brasil