

Avaliação Bibliométrica de Pesquisadores: não é correta . . . nem mesmo errada

Frank Laloë⁽¹⁾ (lalo@lkb.ens.fr) e Rémy Mosseri⁽²⁾

⁽¹⁾ Laboratoire Kastler Brossel, École normale supérieure, CNRS et UMPC, 24 rue Lhomond, 75005 Paris

⁽²⁾ Laboratoire de Physique Théorique de la Matière Condensée, UMR CNRS 7600, Université Pierre et Marie Curie, 4 Place Jussieu, 75252 Paris Cedex 05

Resumo

Utilizar índices bibliométricos para avaliação de pesquisadores resulta de uma extrapolação injustificada de casos triviais, onde esses índices têm algum sentido, para casos reais onde são postos em prática. As verificações mais elementares não foram feitas para compreender a influência dos métodos de cálculo arbitrariamente escolhidos sobre os resultados obtidos. Atualmente, sua utilização evoca mais a pseudociência, como a astrologia, do que um procedimento científico.

Comenta-se que Wolfgang Pauli, um dos geniais fundadores da mecânica quântica, furioso contra um artigo de física sem nenhum interesse, se pôs a gritar: “Isto não é correto e, pior, nem mesmo errado!”. É verdade que, essencialmente, aquilo que não é correto nem errado não pode ser científico: uma afirmação deve ser refutável para merecer esse qualificativo. A observação de Pauli se aplicaria muito bem a diversas aplicações da bibliometria, promovidas por aqueles cuja crença parece ser que, desde o momento em que se manipulam números, se raciocina cientificamente.

A avaliação bibliométrica dos pesquisadores é na verdade “nem mesmo errada”: sim, quando se compara um pesquisador internacionalmente reconhecido a um pobre coitado que jamais foi citado senão por si próprio, os índices do primeiro são bem superiores aos do segundo; ninguém irá contestar. Se o objetivo do exercício fosse distinguir pesquisadores excepcionais de pesquisadores muito medíocres, sem dúvida poder-se-ia recorrer à bibliometria para descobrir ... o que cada um já sabe. Mas suponhamos que se pretenda extrair alguma informação realmente útil, como por exemplo classificar os pesquisadores no seio de um grupo homogêneo, digamos os pesquisadores de um bom laboratório. Constata-se imediatamente nítidas flutuações de seus índices (H, G ...), e com uma certa surpresa: valores muito diferentes podem ser atribuídos a pesquisadores cuja qualidade da produção científica é considerada muito similar pela comunidade científica. Porquê?

Muita razões explicam porque os métodos bibliométricos dão uma visão simplista das contribuições científicas individuais. São métodos sensíveis à qualidade, certo, mas o “sinal” está submerso num “ruído” criado por uma forte dependência de outras variáveis. Tomemos por exemplo um índice bibliométrico H, função de uma variável X que suporemos ser comparável à qualidade do trabalho científico, da variável Y que é o estilo do pesquisador (trabalha mais só ou em equipe constituída, é mais um pioneiro ou alguém que prefere os domínios já em moda, próximo das aplicações ou não, etc.), da variável Z que é seu estilo de publicação (se dedica mais aos artigos curtos ou aos artigos de fundo, talvez aos livros ou compêndios? é atraído pelas revistas ditas de prestígio, gênero *Nature* ou *Science*, mesmo se essas são menos utilizadas na sua área?), e enfim W (pertence a uma linha de pesquisa bastante reconhecida depois de anos, ou escolheu um pequeno domínio emergente, etc.). Tal lista de variáveis não é limitante; poder-se-ia acrescentar por exemplo a capacidade de comunicação do pesquisador e seu gosto pelos congressos, que nem sempre se relacionam com sua criatividade. Qualquer cientista sabe que, quando se busca o destaque da variável X, as diversas outras

variáveis vão agir como um “ruído estatístico”. A única maneira de se acessar a variável de interesse é eliminar este ruído por procedimentos adequados. Se efetuados sobre grandes amostragens, estatisticamente as variáveis Y, Z, W se distribuirão sobre todos os seus possíveis valores, e sua influência desaparecerá, deixando aparecer a de X. É isto que permite a bibliometria obter valores pertinentes para, por exemplo, uma avaliação comparativa da produção nacional dentro de uma grande área de pesquisa. Ao contrário, utilizar H para conhecer a variável X no nível individual é simplesmente um erro de raciocínio que não seria perdoado em nenhum laboratório sério de pesquisa.

Além disto, quando um autor redige um artigo e nele inclui referências, isto não é um ato destinado à bibliometria: o objetivo principal das citações não é levantar uma espécie de lista de premiação, mas dar ao leitor informações que lhe são úteis para ler o artigo em questão. É então um processo relativo, profundamente contextual. Por conveniência, pode-se por exemplo citar um artigo de revisão em vez das fontes originais, para economizar espaço. Muitas vezes se cita um artigo que permite encurtar sua própria redação, e se escolhe então o texto adequado por uma questão de similaridade. Pode-se mesmo citar um artigo que se considera errado com o objetivo de corrigi-lo! Quando se trata de facilitar a repetição de experimentos em outros laboratórios, privilegia-se nas citações os artigos descritivos dos métodos e do instrumental. Para as idéias científicas mais abstratas, ao contrário, são geralmente os artigos derivados que se citam, não o grande artigo original e fundador. É portanto um emprego muito lateral das citações, para não dizer um desvio de sua função real, que sejam tomadas como elemento básico na avaliação de qualidade científica. Pior ainda, por um efeito muitas vezes assinalado, arrisca-se introduzir mudanças na maneira como as citações serão feitas no futuro, em detrimento da qualidade da redação científica e conseqüentemente da própria pesquisa.

A base utilizada na avaliação bibliométrica nas “ciências duras” é o *Web of Science* do ISI.

Primeiro comentário: as obras científicas (livros, compêndios) não são consideradas no cálculo do fator H. Primeiro paradoxo, considera-se amplamente que uma das melhores maneiras para um pesquisador “deixar seu traço numa área” é precisamente publicar livros ou compêndios.

Segundo comentário: os índices G, H, etc., classicamente utilizados para classificar indivíduos, são tão fundamentalmente tendenciosos quanto a clas-

sificação de Xangai¹. No cálculo dos índices, a contribuição de um autor é a mesma seja ele único signatário ou se há dez autores! Poderia parecer elementar calcular índices G' , H' , com base no número de citações dividido pelo número de autores, o que decorreria da lógica mais elementar, mas ninguém segue este procedimento sobre os dados do ISI². A parcialidade é evidente: se três amigos decidem publicar sempre em conjunto durante toda a sua carreira, cada um de seus índices H dará um pulo para cima.

Terceiro vício: tudo é centrado no curto prazo. Em muitas áreas, pequenos avanços técnicos geram uma avalanche de publicações, rapidamente esquecidas, mas numerosas. Assim, os índices são muito orientados aos temas da moda, mesmo que eles desapareçam logo. Não existe nenhuma razão particular para este vício, poder-se-ia facilmente imaginar cálculos mais adaptados nos quais se leva em conta sobretudo as publicações que têm influência a longo prazo. Mas, novamente, ninguém o faz³.

Quarto vício, deficiência técnica pode ser, mas bem real em alguns casos: a base ISI não é homogênea, flutuou ao sabor dos hábitos de trabalho dos operadores de preenchimento de dados que a alimentaram durante décadas. Isto provocou uma série de correções necessárias, que não cabe discutir aqui, mas que demandam um trabalho de especialista. Apenas um trabalho preciso permitiria reconstruir os índices reais, mas evidentemente ninguém se dá ao trabalho: é muito mais simples produzir uma classificação com os falsos dados obtidos com três “clics” de computador.

Para terminar, estes índices de avaliação individual:

- 1) Jamais foram testados racionalmente para correlacioná-los com outras avaliações; são métodos que se apresentam como métodos de avaliação científica, mas paradoxalmente não passaram eles próprios pela avaliação científica de sua confiabilidade e de sua pertinência. Mesmo verificações das mais elementares, como as assinaladas acima, jamais foram feitas para compreender a influência dos métodos de cálculo escolhidos nos resultados obtidos.

¹Não levam em conta o tamanho dos estabelecimentos.

²Não se trata de afirmar que esta maneira de fazer seria boa, mas simplesmente que ela seria menos errada que o método original.

³No caso do fator de impacto das revistas faz-se justamente o contrário, uma vez que apenas os dois últimos anos são levados em conta; para algumas áreas de pesquisa isto pode ter eventualmente algum sentido, mas não para o núcleo das grandes disciplinas científicas.

- 2) Ninguém parecer ter investido tempo na tentativa honesta de melhorá-los para obter valores que sejam mais pertinentes. Promover o debate sobre a pesquisa inteligente da qualidade real da pesquisa científica não parece ser prioritário.
- 3) São “nem mesmo errados”, porque embora contenham um pouco de informação sobre os indivíduos, na maioria das vezes esta informação é trivial e já conhecida. Utilizados numa população homogênea de pesquisadores (processos num Comitê Nacional, por exemplo), definem mais um estilo de trabalho (voltado para o trabalho em equipe ou líder em seu domínio, etc.), enquanto todos estes estilos são necessários para uma pesquisa equilibrada e eficaz. Utilizá-los na avaliação de pesquisadores é portanto um erro de raciocínio, uma espécie de extrapolação injustificada de casos triviais, onde os índices têm algum sentido, para casos reais onde serão utilizados.
- 4) Seu sucesso se deve à falsa facilidade que proporcionam: obter um número com poucos “clicks” é evidentemente bem cômodo.
- 5) A fé nestes índices se transformou numa espécie de crença que escapa ao racional. A comparação que vem à mente é a astrologia ou a numerologia, que também se adornam das virtudes científicas, mas jamais passaram com sucesso pelo crivo de uma verdadeira avaliação científica.

Receita

Se você é um bom pesquisador e deseja melhorar seu índice H, eis alguns conselhos:

- 1 Trabalhe num grupo de pelo menos 5 ou 6 pessoas, se possível mais, cujas publicações são sistematicamente compartilhadas; isso permitirá pelo menos dobrar seu índice, talvez mais. Ademais, este agrupamento de esforços permite compartilhar os “meios” (materiais e humanos, posdocs por exemplo) e talvez pode mesmo aumentar sua produtividade real, o que não é desprezível. Inútil acrescentar que, quanto mais brilhantes forem seus colegas, mais você se beneficiará; portanto, escolha-os bem!
- 2 Favoreça os grandes domínios; constata-se uma correlação entre as taxas de citações e o tamanho do domínio científico, devido ao fato dos artigos de pequenos domínios citarem muitos artigos gerais e não o inverso. Evite trabalhos à margem da corrente geral de seu domínio, mesmo se você é genial: passarão 10 anos para que seus trabalhos sejam realmente reconhecidos, e então serão os trabalhos derivados dos seus que serão citados. Enfim, não se deixe deslumbrar pelo interesse científico de suas pesquisas: arriscar a tentativa de um furo científico elucidativo raramente é compensador antes de décadas!
- 3) Não perca tempo sobretudo a publicar livros ou compêndios, qualquer que seja seu impacto internacional; são peso morto para os indicadores.
- 4) Não dê muita importância à missão fundamental do pesquisador, a produção de conhecimento, em particular ao redigir seus artigos; é a comunicação que passa na frente.