Forget the Ideology of Mass Production in Medical Research! Quality, not Quantity is Decisive

Schluss mit der „Tonnenideologie“ in der medizinischen Forschung! Qualität, nicht Quantität ist wichtig

EDITORIAL

ACCEPTED: March 2018
PUBLISHED ONLINE: April 2018
DOI: 10.5960/dzsm.2018.326

Böning D. Schluss mit der „Tonnenideologie“ in der medizinischen Forschung! Qualität, nicht Quantität ist wichtig.Dtsch Z Sportmed. 2018; 69: 85-86.

Prof. Dr. Dieter Böning
Vorsitzender Verein zur Förderung der Sportmedizin (VFSM) e.V

Every researcher wants to be important. If not manifestly, then at least frequently quoted by peers. This is an important basis for an academic career.

But how is the importance of a scientist really recognized? In the case of Albert Einstein and his theory of relativity or Alexander Fleming and penicillin, it's easy. But what about the many others who always advance in small steps? If they are often cited by colleagues, that may be one sign. So how does one get there? It is necessary to publish as many articles as possible in journals which are widely read and cited. Various factors have been invented to quantify this. The Impact Factor and recently often the Hirsch-Index (h-Index) have been used as a measure of the frequency of citation. This is, however, being criticized more and more, and with good reason.

In 2013, I already wrote a critical editorial for this journal (2) about the Impact-Factor. Last year, three scientific academies in Europe (Académie des Sciences, Leopoldina and Royal Society) published a joint statement on the evaluation of scientists and their achievements (1). Similar ideas have been in the Support Guidelines of the German Research Foundation for some time, as the President, Peter Strohschneider, made clear in a speech in 2017 (7). In Forschung und Lehre, the journal of the Hochschullehrerverband (Association of University Lecturers), a sort of Professors' Union, there have also been many critical commentaries (for example (8)).

Essential Statement of the Academies

„Evaluation requires peer review by acknowledged experts working to the highest ethical standards and focusing on intellectual merits and scientific achievements. Bibliometric data cannot be used as a proxy for expert assessment. Well-founded judgment is essential. Overemphasis on such metrics may seriously damage scientific creativity and originality. Expert peer review should be treated as a valuable resource“ (1).

What are Bibliometric Data?

These are, among other things, the quantities mentioned above: They measure how often articles of one journal or a certain researcher are cited. They have been used for years, very often still today, without reservation to “calculate” the importance of scientists.

The Impact Factor

The best-known quantity is the Impact Factor. It describes how often an article from the preceding two years in a journal was cited on average in the report year; less often, the citations for articles in the last five years were assessed. It is, thus, actually a measure of the importance of a journal. If I manage to publish in Nature, Science, The Lancet or The New England Journal of Medicine, I am assigned a very high Impact Factor (40.1; 37.2; 47.8 and 72.4 in 2017), even if I, personally, was not quoted. The presumed importance of a researcher is deduced from the sum of the Impact Factors for his articles.

Why is the Impact Factor so high for the journals named above? They only accept articles, or even brief notices, which appear to be particularly important; correspondingly, they are often cited. In Nature and Science, the articles are mostly from natural sciences. Clinical medicine appears much more rarely. Since 1869, in a total of 389 000 articles in Nature, 1 878 dealt with Internal Medicine, 377 with Sports Medicine. And what do the editors consider important in Sports Medicine? Primarily doping and genetics. In the interdisciplinary medical journal The Lancet (founded in 1820), there is also very little about Sports Medicine: The word “doping” appeared 27 times in the title of an article, “physical training” 26 times, and “sports injury” seven times , when all volumes are evaluated .

Moreover, the Impact Factor depends in great measure on the size of the field. Most journals in small special fields like Sports Medicine, in which overall less is published than, say, in Internal Medicine, have difficulty reaching the rating 2. Language also plays a great role; these days, one must write in English to achieve international recognition.

Which Articles are Cited especially often (though not Necessarily Quickly)?

1. Important Knowledge or Theories

These articles often take time to become recognized. Occasionally a presumably important bit of news turns out after a few years to be a flop, like supposed autism after a measles vaccination in a Lancet article.

2. Methodical Advances

Here, too, the advances often take a long time before they are applied in other laboratories and can then be cited. ›

QR-Code scannen und Artikel online lesen.

KORRESPONDENZADRESSE:
Univ. Prof. a. D. Dr. med. Dieter Böning
Institut für Physiologie
Charité-Universitätsmedizin Berlin
Charitéplatz 1
10117 Berlin
✉: dieter.boening@charite.de

3. Reviews

These are often cited; but they are only important when new relationships are recognized (often, but not only by means of meta-analyses).

All of this reduces the usability of the Impact Factor, primarily for 2 years.

One particular problem is the skewed distribution of the citations (6). Few articles are usually cited very frequently, most rarely or not at all. With an article in Nature or The Lancet, I will still get a rating of 40 or 48 for my Impact sum, even for an article which is never cited at all.

Can the Impact Factor be enlarged in a not-quite honest manner? In an institute with various working teams, researchers who were not actually involved are often included mutually as authors. And the boss is almost always in the list: correctly, if he suggests and checks, incorrectly if he does nothing. A particularly bad way are the so-called quotation cartels, when various research groups agree to quote each other.

In past years, the greater the sum of Impact Factors, the greater the probability of obtaining a professorship or funding. Sociologists have coined the term "Tonnenideologie" (~ Ideology of Tons) for the fondness for scientific mass production (8).

The Hirsch-Index

The American physicist Jorge E. Hirsch suggested an Index (Hirsch-Index h) which is person-related as a way to avoid the various disadvantages of the Impact Factor (4). The publications of an author are sorted by frequency of citation. The most-cited article is assigned number 1, the others follow in decreasing citation frequency. A scientist's Hirsch-Index is the value for which the number agrees with the number of citations. For example, h=10 means that the 10th publication was cited ten times. The Index is thus not similar to the arithmetic mean, but rather the median value.

The advantage over the Impact Factor is that the Hirsch-Index is not journal-related but author-related. But that is almost the end of its usability. The first disadvantage is that there are 3 different versions. The lowest value is obtained with SCOPUS, an abstract and citation database, which (with some exceptions) only goes back to 1996. Care has to be taken by the user that various versions of the name (in my case, not only Böning, but also Boning, Boening, Bœning in English articles) are included. The Web of Science is used most frequently, it usually delivers somewhat higher ratings; but it doesn't include any books or book chapters. The highest value is obtained with Google Scholar.

Other characteristics make the Hirsch-Index almost as unsuitable as the Impact Factor. It preferentially describes mediocracy, not necessarily excellence. Whether the author's most important publications were cited 1000 times or 50 times is not apparent from the Hirsch-Index. And whether the theses are correct or false also cannot be determined from the frequency of citation. Over the years, I have found many typical errors in some highly-cited publications (3). And if a claimed effect can't be reproduced, it can often be discussed in the literature.

Certainly, many well-known scientists, especially in Physics or Biosciences, have a high Hirsch-Index of about 100 (for example Stephen Hawking). But even Nobel Prize winners for Medicine have a modest rating of 20 even after receiving the prize (5). And Peter Higgs, who predicted the Higgs Particle (Nobel Prize for Physics in 2013) has the pathetic value of only 11!

I therefore suggest forgetting the Hirsch-Index – although it affects me personally. My h-rating happens to be higher than that of some Nobel Prize winners.

Literatur

(1) **ACADÉMIE DES SCIENCES LARS.** Statement by three national academies (Académie des Sciences, Leopoldina and Royal Society) on good practice in the evaluation of researchers and research programmes. 2017, p. 1-4. https: //www.leopoldina.org/uploads/tx_leopublication/2017_Statement_3Acad_Evaluation.pdf [16th March 2013].

(2) **BÖNING D.** Publizieren in der DZSM lohnt sich! Dtsch Z Sportmed. 2013; 64: 95. doi:10.5960/dzsm.2012.066

(3) **BÖNING D.** Scientific progress or regress in Sports Physiology? Int J Sports Physiol Perform. 2016; 11: 1106-1110. doi:10.1123/IJSPP.2016-0289

(4) **HIRSCH JE.** An index to quantify an individual's scientific research output. Proc Natl Acad Sci USA. 2005; 102: 16569-16572. doi:10.1073/pnas.0507655102

(5) **KREINER G.** The slavery of the h-index—measuring the unmeasurable. Front Hum Neurosci. 2016; 10: 556. doi:10.3389/fnhum.2016.00556

(6) **OSTERLOH MF, BRUNO S.** Absurde Mess-Manie. Der fragwürdige Impact des Impact-Faktors. Forschung & Lehre. 2017; 24: 876-878.

(7) **STROHSCHNEIDER P.** Über Wissenschaft in Zeiten des Populismus. In: Jahrestagung der DFG. Halle/Saale: 2017. http: //www.dfg.de/dfg_magazin/querschnitt/171218_rede_des_jahres/index.jsp

(8) **STRÜBING J.** Problem, Lösung oder Symptom? Zur Forderung nach Replizierbarkeit von Forschungsergebnissen. Forschung & Lehre. 2018; 25: 102-105.