

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

Forget the Ideology of Mass Production in Medical Research!

Quality, not Quantity is Decisive

Jeder Forscher möchte gern bedeutend sein. Wenn das nicht offensichtlich ist, möchte man wenigstens oft von den Fachgenossen zitiert werden. Dies ist eine wichtige Grundlage für die akademische Karriere.

Aber woran erkennt man die Bedeutung eines Wissenschaftlers wirklich? Bei Albert Einstein mit der Relativitätstheorie oder bei Alexander Fleming mit dem Penicillin ist das einfach. Aber wie ist das bei den vielen anderen, die immer nur kleine Schritttchen vorangekommen sind? Wenn sie oft von Kollegen zitiert wurden, könnte das ein Hinweis sein. Wie schafft man das? Man muss möglichst viele Artikel in Zeitschriften, die viel gelesen und zitiert werden, unterbringen. Um das zu quantifizieren, hat man verschiedene Faktoren erfunden. Als Maße für die Zitierhäufigkeit sind der Impact-Faktor und neuerdings oft der Hirsch-Index (h-Index) in Gebrauch. Dies wird allerdings mehr und mehr und mit guten Gründen kritisiert.

Zum Impact-Faktor habe ich bereits 2013 ein kritisches Editorial für diese Zeitschrift verfasst (2). Im vergangenen Jahr haben drei wissenschaftliche Akademien in Europa (Académie des Sciences, Leopoldina und Royal Society) eine gemeinsame Stellungnahme über die Bewertung von Wissenschaftlern und deren Leistung abgegeben (1). Ähnliche Vorstellungen findet man schon seit längerer Zeit in den Förderungsrichtlinien der Deutschen Forschungsgemeinschaft, die ihr Präsident Peter Strohschneider 2017 in einer Rede (7) erläuterte. Auch in „Forschung und Lehre“, der Zeitschrift des Hochschulverbandes, sozusagen der Professoren-gewerkschaft, erscheinen viele kritische Kommentare (z. B. (8)).

Wesentliche Aussagen der Akademien

„Die Bewertung erfordert Überprüfung durch anerkannte Fachleute („peer review“), die mit höchsten ethischen Standards arbeiten und auf intellektuelle Verdienste und wissenschaftliche Fortschritte fokussieren. Bibliometrische Daten können nicht als Ersatz für Überprüfung durch Experten benutzt werden. Eine begründete Beurteilung ist wesentlich. Überbewertung von bibliometrischen Messgrößen kann wissenschaftliche Kreativität und Originalität schwer schädigen“ (1).

Was sind bibliometrische Messgrößen?

Es sind u. a. die oben erwähnten Größen: Sie messen, wie oft Artikel einer Zeitschrift oder eines bestimmten Forschers zitiert werden. Jahrelang, sehr oft noch heute, hat man sie recht bedenkenlos benutzt, um die Bedeutung von Wissenschaftlern zu „berechnen“.

Der Impact-Faktor

Die bekannteste Messgröße ist der Impact-Faktor. Er beschreibt, wie oft ein Aufsatz aus den vorangegangenen beiden Jahren in einem Journal im Durchschnitt im Berichtsjahr zitiert wurde; seltener werden die Zitationen für Artikel aus den letzten fünf Jahren gewertet. Er ist also eigentlich ein Maß für die Bedeutung einer Zeitschrift. Gelingt es mir, eine Veröffentlichung in Nature, Science, Lancet oder New England Journal of Medicine unterzubringen, bekomme ich dafür einen sehr hohen Impact-Faktor (40,1; 37,2; 47,8 und 72,4 im Jahr 2017) angerechnet, auch wenn ich selbst überhaupt nicht zitiert wurde. Die angebliche Bedeutung eines Forschers wird aus der Summe der Impact-Faktoren für seine Artikel abgeleitet.

Warum ist der Impactfaktor der obengenannten Zeitschriften so hoch? Sie nehmen nur Aufsätze oder auch Kurzmitteilungen an, die besonders wichtig zu sein scheinen; entsprechend werden sie häufig zitiert. In Nature und Science sind dies überwiegend Arbeiten aus den Naturwissenschaften. Die klinische Medizin kommt viel seltener vor. Seit 1869 gab es in Nature bei insgesamt 389 000 Beiträgen nur 1 878 aus der Inneren Medizin, 377 aus der Sportmedizin. Und was halten die Herausgeber bei der Sportmedizin für wichtig? Vor allem Doping und Genetik. In der fachübergreifenden medizinischen Zeitschrift Lancet (bestehend seit 1820) findet man ebenfalls sehr wenig Sportmedizin: 27 Mal erscheint in einem Artikel Doping, 26 Mal Physical Training, sieben Mal Sports Injury.

Außerdem hängt der Impactfaktor in erheblichem Ausmaß von der Fachgröße ab. Die meisten Zeitschriften in kleinen Fächern wie der Sportmedizin, wo insgesamt weniger publiziert wird als z. B. in der Inneren Medizin, haben Mühe, den Wert 2 zu erreichen. Ebenfalls spielt die Sprache eine große Rolle; heutzutage muss man in Englisch schreiben, um international Beachtung zu finden.

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



Article incorporates the Creative Commons Attribution – Non Commercial License.
<https://creativecommons.org/licenses/by-nc-sa/4.0/>



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

Welche Veröffentlichungen werden besonders häufig (aber nicht unbedingt bald) zitiert?

1. Bedeutende Erkenntnisse oder Theorien

Diese Arbeiten brauchen oft Zeit, bis sie anerkannt werden. Manchmal erweist sich eine angeblich wichtige Neuigkeit nach ein paar Jahren auch als Flopp wie der angebliche Autismus nach Masernimpfung in einem Lancet-Artikel.

2. Methodische Fortschritte

Auch hier dauert es oft, bis sie in anderen Labors eingesetzt und danach zitiert werden können.

3. Übersichtsarbeiten (Reviews)

Sie werden oft zitiert; bedeutend sind sie aber nur, wenn neue Zusammenhänge erkannt werden (häufig, aber nicht nur durch Meta-Analysen).

All dies verringert die Brauchbarkeit des Impact-Faktors vor allem für 2 Jahre.

Ein besonderes Problem ist die Schiefverteilung der Zitationen (6). Wenige Artikel werden üblicherweise sehr häufig zitiert, die meisten selten oder gar nicht. Mit einem überhaupt nicht zitierten Artikel in Nature oder Lancet bekomme ich trotzdem den Wert 40 bzw. 48 für meine Impactsumme.

Wie kann man den Impact-Faktor auf nicht ganz redliche Weise vergrößern? In einem Institut mit verschiedenen Arbeitsgruppen werden oft gegenseitig Forscher als Autoren aufgenommen, die eigentlich nicht beteiligt sind. Und der Chef ist sowieso fast immer dabei: Richtig, wenn er anregt und überprüft, unangebracht, wenn er nichts dazu tut. Ein besonders schlechter Weg sind sogenannte Zitierkartelle, wenn sich verschiedene Forschergruppen gegenseitig zum Zitieren verabreden.

Je höher die Summe der Impact-Faktoren, desto höher war in den vergangenen Jahren häufig die Wahrscheinlichkeit, eine Professur oder Drittmittel zu bekommen. Soziologen haben für die Vorliebe für wissenschaftliche Massenproduktion den Begriff der „Tonnenideologie“ erfunden (8).

Der Hirsch-Index

Um verschiedene Nachteile des Impact-Faktors zu vermeiden, hat der amerikanische Physiker Jorge E. Hirsch einen Index vorgeschlagen (Hirsch-Index h), der personenbezogen ist (4). Man sortiert die Veröffentlichungen eines Autors nach der Zitierhäufigkeit. Die höchstzitierte Arbeit bekommt die Nummer 1, die anderen folgen nach absteigender Zitierhäufigkeit. Ein Wissenschaftler bekommt als Hirsch-Index den Wert, für den die Nummer mit der Zahl der Zitationen übereinstimmt. Z. B. bedeutet $h=10$, dass die 10. Veröffentlichung zehnmals zitiert wurde. Der Index ähnelt daher nicht dem arithmetischen Mittel, sondern eher dem Medianwert.

Der Vorteil gegenüber dem Impact-Faktor ist, dass der Hirsch-Index nicht zeitschriften-, sondern autorbezogen ist. Damit endet aber schon fast die Brauchbarkeit. Der erste Nachteil ist, dass es 3 verschiedene Versionen gibt. Die niedrigsten Werte bekommt man mit SCOPUS, einer Abstract- und Zitationsdatenbank, die (mit Ausnahmen) nur bis 1996 zurückverfolgt. Man muss auch selbst darauf achten, dass verschiedene Namensversionen (in meinem Fall nicht nur Böning, sondern auch Boning, Boening, Bøning in englischen Artikeln) einbezogen werden. Am häufigsten wird das Web of

Science benutzt, das meist etwas höhere Werte liefert; es berücksichtigt aber keine Bücher oder Buchkapitel. Die höchsten Werte bekommt man bei Google Scholar.

Weitere Eigenschaften machen den Hirsch-Index fast genauso ungeeignet wie den Impact-Faktor. Er beschreibt bevorzugt die Mittelmäßigkeit, nicht unbedingt die Exzellenz. Ob die wichtigste Veröffentlichung eines Autors 1000 Mal oder 50 Mal zitiert wurde, ist aus dem Hirsch-Index nicht ersichtlich. Und ob seine Thesen richtig oder falsch sind, erschließt sich ebenfalls nicht aus der Zitierhäufigkeit. Ich habe im Laufe der Jahre eine Reihe von typischen Fehlern in teils hochzitierten Veröffentlichungen gefunden (3). Und lässt sich ein behaupteter Effekt nicht reproduzieren, kann das auch öfters in der Literatur diskutiert werden.

Sicherlich haben viele bekannte Wissenschaftler, besonders in Physik oder Biowissenschaften, einen hohen Hirsch-Index um die 100 (z. B. Stephen Hawking). Aber selbst bei Nobelpreisträgern der Medizin gibt es auch bescheidene Werte von 20 noch nach der Preisverleihung (5). Und Peter Higgs, der das Higgs-Teilchen vorhergesagt hat (Nobelpreis für Physik 2013) hat den kläglichen Wert von 11!

Ich schlage daher vor, auch den Hirsch-Index zu vergessen – obwohl es mich persönlich trifft. Mein h -Wert ist nämlich höher als der einiger Nobelpreisträger. ■

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/IJSP.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.