EDITORIAL

Not everything that matters can be measured and not everything that can be measured matters

Bryan D. Choi, MD, PhD,¹ Peter E. Fecci, MD, PhD,² and John H. Sampson, MD, PhD^{2–4}

¹Department of Neurosurgery, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts;

²Duke Brain Tumor Immunotherapy Program, Division of Neurosurgery, Department of Surgery; ³Departments of Pathology and Radiation Oncology; and ⁴The Preston Robert Tisch Brain Tumor Center, Duke University Medical Center, Durham, North Carolina

THE use of bibliometrics to rank the most academically productive programs in neurosurgery is a trend that is growing in popularity. These analyses are often based on the number of publications accumulated over time as well their respective citations as a proxy for quality. Using this approach, an index or value can be ascribed to an individual, a group (e.g., department or university), or even a scholarly journal, in an effort to capture its relative productivity and impact on the academic community.

Taylor and colleagues have undertaken a commendable effort to develop a way to assess the academic productivity of our peers in neurosurgery using a 5-year institutional *h*-index (*ih*[5]-index), which they propose is tailored to more accurately reflect a program's recent research achievements.² They also demonstrate that the *ih*(5)-index has predictive capacity in characterizing intradepartmental publishing equity—that is to say, the degree to which individuals within a department contribute equally—which in turn may be correlated with greater productivity overall. Subsequently, Lozano and colleagues employed the *ih*(5)-index in a separate study to demonstrate that the University of Toronto, when compared to all United States neurosurgery programs, ranks first in terms of generating the highest number of publications that are cited most frequently.¹

Certainly, much like bibliometric profiles that have preceded it, the *ih*(5)-index seems reasonable insofar as it uses many if not all of the fundamental elements to provide a score. It is similar to the more general *h*-index, which

also refers to the number of papers (*h*) that receive *h* or more citations, except that it is restricted to a 5-year time course and institutional data. Other more recently developed indices, including the *g*-index and *e*-index, are meant to complement the *h*-index and are designed to capture highly cited publications or “ignored” citations, whereas the *m*-index reflects an *h*-index that is corrected for the number of years since a first published paper, which essentially eliminates bias against more junior scholars who have been active for a shorter period of time.

Despite their differences, the various strategies that have been developed to analyze bibliometric data are more similar than not, and thus some of the same caveats should be considered when interpreting the *ih*(5). Because the field of neurosurgery is relatively small, one of the primary sources of bias in the *ih*(5) is the degree to which neurosurgeons choose to publish in neurosurgical specialty journals—which de facto garner a smaller audience—versus more broadly in other fields. In addition, the *ih*(5) conversely favors certain publications like those that develop clinical guidelines; one such manuscript referenced by Lozano and colleagues entitled “Guidelines for the management of spontaneous intracerebral hemorrhage” was published in *Stroke*, and as the authors correctly identify, it is the most-cited manuscript from the University of Toronto neurosurgery program. Wider appeal can often be achieved through close collaboration with faculty and scientists in nonneurosurgical departments; incidentally, a prime example of this is the same highly cited manuscript

ACCOMPANYING ARTICLES See pp 547–570. DOI: 10.3171/2014.10.JNS141025. DOI: 10.3171/2014.12.JNS142553.

INCLUDE WHEN CITING Published online June 26, 2015; DOI: 10.3171/2015.2.JNS142977.

DISCLOSURE The authors report no conflict of interest.

from University of Toronto whose first author, Dr. Lewis B. Morgenstern, despite having a faculty appointment in the Department of Neurosurgery, happens to be a clinical neurologist by training.

As such, it is not unreasonable to interpret the *ih(5)*-index, at least in some cases, to be a mark of successful academic collaboration. Supporting this is the finding by Taylor and colleagues that all bibliometric indices, including the *ih(5)*, were positively correlated with the number of faculty in a neurosurgical department; however, this certainly begs the question of whether disparities between programs persist after controlling for faculty size. To reiterate the point, not only the number, but the type of faculty present in a department needs to be considered when interpreting the *ih(5)*-index. Depending on the organization of a particular department, the presence of nonclinical basic science faculty—or those with varying degrees of clinical responsibilities—further confounds the conclusions to be gleaned from index publication data. That is to say, if a neurosurgeon is listed as an author on a particular paper that was primarily driven by oncologists, neurologists, or basic scientists, that paper may in fact elevate the reputation of a neurosurgical program without accurately reflecting the involvement of a neurosurgeon in its success.

Overall, there does seem to be some truth to the saying that “not everything that matters can be measured and not everything that can be measured matters.” While the *ih(5)*-index represents a useful new tool, there is certainly more to academic productivity than publications and their respective citations. Indeed, a successful department needs to balance different talents and contributions in the realms of teaching activities, politics, and the production of work relative value units (RVUs). Toward that end, it is important to note that neurosurgery program rankings by the *ih(5)*-index do not directly recapitulate those published by sources such as the Doximity Residency Navigator and *U.S. News & World Report*. Specifically, the Doximity rank list identifies a number of programs that are not rated within the *ih(5)* top 10; these include NewYork-Presbyterian Hospital (Columbia Campus), Mayo Clinic College of Medicine, Washington University, Massachusetts General Hospital, University of Washington, and Emory University.

Bibliometric indices are of emerging importance in neurosurgery and represent one way to assess academic productivity among departments while allowing for reassessments of this activity over time. As these metrics continue to be refined, perhaps the most interesting questions remain; for instance, what are the drivers that are most important for academic productivity? Do the manuscripts that rank departments also give us a roadmap to develop or choose a promising program? Many successful programs produce high-impact publications that are centered on a few individuals, often at the expense of clinical volume (i.e., the Pareto principle). Does this trade-off ultimately benefit neurosurgery faculty and residents? There is some component of academic mission here that is important to consider for surgical training. Lastly, there are several financial models that differ in reimbursement pathways, not only at the individual faculty level but also institutionally. The manner in which incentives

are aligned—vis-à-vis hospital-driven or university-based salary or promotions—will skew the balance between RVUs versus academic productivity and in turn influence the interpretation of program ranking by bibliometric index alone.

<http://thejns.org/doi/abs/10.3171/2015.2.JNS142977>

References

1. Lozano CS, Tam J, Kulkarni AV, Lozano AM: The academic productivity and impact of the University of Toronto Neurosurgery Program as assessed by manuscripts published and their number of citations. *J Neurosurg* [epub ahead of print June 26, 2015. DOI: 10.3171/2014.12.JNS142553]
2. Taylor DR, Venable GT, Jones GM, Lepard JR, Roberts ML, Saleh N, et al: Five-year institutional bibliometric profiles for 103 US neurosurgical residency programs. *J Neurosurg* [epub ahead of print June 26, 2015. DOI: 10.3171/2014.10.JNS141025]

Response

Douglas R. Taylor, MD,^{1,2} L. Madison Michael II, MD,^{2,3} and Paul Klimo Jr., MD, MPH^{2,3}

¹University of Tennessee Health Science Center and Department of ²Neurosurgery, University of Tennessee; and ³Semmes-Murphey Neurologic & Spine Institute, Memphis, Tennessee

Like an engine that requires the integration of multiple components, there are numerous facets of a neurosurgical residency that theoretically should work synchronously to produce highly educated, competent, and safe neurosurgeons. Research and its end result—publications—are just some of those components, now mandated as part of the ACGME neurosurgery milestones curriculum.⁶ A neurosurgical program's ability to attract the most innovative and productive resident or faculty applicant relies on the accurate and transparent display of the programs' attributes. We attempted to define the most accurate method of measuring the contemporary academic productivity of each US neurosurgical residency program. We found the *ih(5)*-index to be a simple and reproducible metric that may answer our quest.

When utilizing the *ih(5)*-index, there are several issues to consider. First, the *ih(5)*-index is best used when comparing departments within the same field and using the same methodology. Choi et al. are incorrect in labeling it a biased metric if such departments include neurosurgeons that publish in broader journals. In fact, such departments should be commended, and their impact (publication and associated citations) would be accurately captured with the *ih(5)*-index. The overall impact such academic neurosurgeons have is likely to be small, as we demonstrated in a pilot study that only 10% of neurosurgeons have at least 1 publication in a journal with an impact factor greater than 8.0.³ Citations take time to accrue, so publications at the end of the 5-year period are underrepresented in the total citation count. While the effect of this drawback on our current analysis is negated by the fact that the *ih(5)* calculation was uniformly applied to all programs, this could be remedied by repeating the 5-year analysis at in-

tervals shorter than 5 years and overlapping: e.g., 2009–2013, 2011–2015, 2013–2017, etc.

Highly cited manuscripts—such as review articles, especially those that are part of guidelines, such as the one cited by Choi et al.—may skew a program's true academic impact because the *ih(5)* cannot distinguish these articles from original research. Such secondary research, when done properly, is a time-consuming and valuable endeavor and should rightfully be recognized, but it should not be touted as the highlight of a program's academic achievement, nor should it overshadow original contributions. Future departmental analysis may be improved if these articles were counted separately or all together excluded.

Wider appeal can also be achieved through collaboration with nonneurosurgical faculty or scientists. While our *ih(5)*-index incorporated only publications in which at least 1 neurosurgeon was listed as an author, the neurosurgeon's degree of involvement in the project was not assessed. An author's position in the authorship list can often be a surrogate measure of the degree of involvement that author had in the project.⁵ We have coined the phrase “authorship value” and have previously evaluated this on a small scale.³ Primary, secondary, or senior authors likely had more direct contribution or oversight than all other authors, which either had lesser contributions or were granted “gift authorship.” Gift authorship has become, in the opinion of some, an unhealthy commonality in medical publications.^{1,4,7} Analysis of authorship value would be highly labor intensive but would improve the clarity of future departmental analysis.

Choi et al. bring up the issue of faculty number. While the *ih(5)*-index was less correlated to faculty numbers when compared to other analyzed indices, larger institutions tended to have an advantage over those institutions with less faculty. To correct for number of faculty, a ratio of *ih(5)*-index to faculty (or mean *h*-index per faculty) within a department was calculated. As one might expect, smaller departments were now favored with this method. In fact, 6 of the top-10 ranked programs ranked in this manner had a faculty number <10. Inherently, correcting the *ih(5)* by faculty number shifts the focus to the individual level rather than the institution. The *ih(5)*-index limits the influence of highly productive individual outliers and instead emphasizes the academic productivity of the institution as a whole.

Further, there is concern that many successful programs are the product of high-impact publications from a few individuals who pursue research at the expense of clinical volume. We attempted to study this by using what are known as Lorenz curves and Gini coefficients. We found that although it might be expected that publication equality would play a role in the success of an institution's academic enterprise, other factors such as faculty professional priorities, compensation incentives, and departmental leadership may have a greater influence. This is demonstrated by the fact that only 3 programs included in the top 10 most academically productive programs, ranked by *ih(5)*, met our goal Gini coefficient of ≤ 0.5 , which can be viewed as a more “socialistic” research milieu.

The emphasis placed on research in each respective neurosurgical department will be reflected in their ranking using the *ih(5)*-index. Research for some departments, such as ours, is voluntary, whereas for others it is a requirement. Choi et al. cite other rankings (i.e., Doximity and *U.S. News and World Report*) that are in disagreement with ours. This is an observation we too have published on before,² and is not surprising, given the completely different methodology and criteria, making it impossible to compare their lists to ours. While not without limitations and criticisms, we have created a rank list of nearly all US neurosurgical programs based on academic output for the 5 years 2009–2013. To Choi et al., we counter their title and conclude our response with the following, “If you don't know where you are, you don't know where you're going.”

References

1. Drenth JP: Multiple authorship: the contribution of senior authors. **JAMA** 280:219–221, 1998
2. Khan N, Thompson CJ, Choudhri AF, Boop FA, Klimo P Jr: Part I: The application of the h-index to groups of individuals and departments in academic neurosurgery. **World Neurosurg** 80:759–765, 2013
3. Khan NR, Thompson CJ, Taylor DR, Gabrick KS, Choudhri AF, Boop FR, et al: Part II: Should the h-index be modified? An analysis of the m-quotient, contemporary h-index, authorship value, and impact factor. **World Neurosurg** 80:766–774, 2013
4. King JT Jr: How many neurosurgeons does it take to write a research article? Authorship proliferation in neurosurgical research. **Neurosurgery** 47:435–440, 2000
5. Romanovsky AA: Revised h index for biomedical research. **Cell Cycle** 11:4118–4121, 2012
6. Selden NR, Abosch A, Byrne RW, Harbaugh RE, Krauss WE, Mapstone TB, et al: Neurological surgery milestones. **J Grad Med Educ** 5 (1 Suppl 1):24–35, 2013
7. Smith J: Gift authorship: a poisoned chalice? **BMJ** 309:1456–1457, 1994

Response

Christopher S. Lozano,¹ Joseph Tam, PhD,¹
Abhaya V. Kulkarni, MD, PhD,^{1,2} and
Andres M. Lozano, MD, PhD^{1,3}

¹Division of Neurosurgery, Department of Surgery, University of Toronto;

²Division of Neurosurgery, Hospital for Sick Children; ³Division of Neurosurgery, Toronto Western Hospital, Toronto, Ontario, Canada

We thank Drs. Choi, Fecci, and Sampson for their interest in our work and for their thoughtful comments.

Making an appraisal of the academic stature of neurosurgical programs is important for at least two reasons. First, when assessed over time, the trend can tell you whether you are gaining or losing ground, or simply maintaining your position. Second, it provides a measure of how a particular program stacks up in relation to others.

How to assess the academic ranking of a neurosurgical department is a controversial topic and can be a hornet's nest, particularly if subjective measures come into play. To get away from the “beauty contest” concept, it is important to establish objective and reproducible metrics.

There will be, even here, however, variations in opinion as to how reliably any chosen measure represents the true academic value of any particular program. Having said that, we feel that, as opposed to the volume of patients treated or cases operated or the number and nature of grants received and their monetary value, the number of papers written and their citations constitute a leading metric for academic productivity. The thesis is that papers that are cited have impact in that they influence what we practice and move the field conceptually. Papers that are less cited may have a less enduring impact and influence. It is for this reason that we chose not only the papers but also how often they are cited as a surrogate for academic impact. There is nothing like the power of data when it comes to sifting through the quagmire, particularly when it comes to the potentially emotional responses that are

elicited by rankings. We realize that there are many other factors that come into play and that this measure does not capture clinical excellence, teaching accomplishments, suitability as a training program, etc., but the data for numbers of papers and citations are readily available and can be verified using resources available in the public domain.

We hope that our work serves to have individual neurosurgical programs ascertain their own academic productivity with papers and citations and to compare their own program with others both regionally and indeed internationally. The process can be repeated at intervals to see which way things are heading and whether it's full speed ahead without adjusting the sails and rudder, or whether a course correction would be beneficial for getting where you want to go.