Pediatric Academic Productivity: Pediatric Benchmarks for the h- and g-Indices

Megan M. Tschudy, MD, MPH, Tashi L. Rowe, BA, George J. Dover, MD, and Tina L. Cheng, MD, MPH

Objective To describe h- and g-indices benchmarks in pediatric subspecialties and general academic pediatrics. Academic productivity is measured increasingly through bibliometrics that derive a statistical enumeration of academic output and impact. The h- and g-indices incorporate the number of publications and citations. Benchmarks for pediatrics have not been reported.

Study design Thirty programs were selected randomly from pediatric residency programs accredited by the Accreditation Council for Graduate Medical Education. The h- and g-indices of department chairs were calculated. For general academic pediatrics, pediatric gastroenterology, and pediatric nephrology, a random sample of 30 programs with fellowships were selected. Within each program, an MD faculty member from each academic rank was selected randomly. Google Scholar via Harzing's Publish or Perish was used to calculate the h-index, g-index, and total manuscripts. Only peer-reviewed and English language publications were included. For Chairs, calculations from Google Scholar were compared with Scopus.

Results For all specialties, the mean h- and g-indices significantly increased with academic rank (all $P < .05$) with the greatest h-indices among Chairs. The h- and g-indices were not statistically different between specialty groups of the same rank; however, mean rank h-indices had large SDs. The h-index calculation using different bibliographic databases only differed by ± 1.

Conclusion Mean h-indices increased with academic rank and were not significantly different across the pediatric specialties. Benchmarks for h- and g-indices in pediatrics are provided and may be one measure of academic productivity and impact. *(J Pediatr 2016;169:272-6)*.

Academic promotion committees increasingly are using bibliometrics in assessment.[1,2] The h- and g-indices originally were designed for basic scientists and have predictive properties of future achievement.[3] They can be easily calculated with the use of online tools.[4-6] These indices have become used more widely in medicine.[7,8] Known to vary by discipline, studies have assessed the h-index in academic neurosurgery,[8] otolaryngology,[9] and radiation oncology.[10] However, an index of citation count alone may not recognize quality or importance of the papers and has additional limitations related to self-citation.[11]

The Hirsch, or h-index, is one of the most widely adopted of these bibliometrics.[3] The h-index reflects the number of publications and citations per publication.[12] Caution has been urged in the use of Hirsch's index, because initial calculations were done only in a small sample size of researchers who were very accomplished in their field and cumulative measures like the h-index contain intrinsic properties that can result in overestimates of their predictive powers.[13,14]

The h-index is the number of papers published by an author which have received h citations. Hirsch argues that in comparison with traditional journal impact factor metrics, the h-index is a relatively unbiased assessment that is highly predictive of future output[12]; however, others have questioned the value of the h-index for multiple reasons, including failure to account for number of authors on a paper and author placement (eg, first or last author) and varying citation in different disciplines.[15,16]

The g-index is another often-used bibliometric measure.[2] The g-index is calculated by arranging a list of manuscripts by the author in "decreasing order of the number of citations that they received, the g-index is the largest number such that the top g articles received (together) at least g^2 citations."[2] It is highly correlated with the h-index but is always equal to or greater than the h-index.

Although bibliometric indices, including the h- and g-indices, are being used as measures for academic advancement at many institutions, there have not been any benchmarks presented for pediatrics. The objective of this study was to describe h- and g-indices in 2 pediatric subspecialties and general academic pediatrics and to describe these bibliometrics for pediatric chairs and academic pediatricians by rank.

Methods

Methods for sample identification and bibliometric calculation were replicated from those described by a paper on use of the h-index in neurosurgery.[8] A list of all pediatric department chairs (N = 136) was taken from the Association of

From the Department of Pediatrics, The Johns Hopkins University School of Medicine, Baltimore, MD

The authors declare no conflicts of interest.

http://dx.doi.org/10.1016/j.jpeds.2015.10.030

Medical School Pediatric Department Chairs Web site in 2012. A random sample of 30 was chosen for bibliometric analysis.

Three pediatric specialties were chosen for analysis. Sampling strategy within each specialty also was replicated from a neurosurgery publication.[8] For pediatric gastroenterology and pediatric nephrology, 30 divisions from each specialty were selected randomly from fellowship programs accredited by the Accreditation Council for Graduate Medical Education. Because the Accreditation Council for Graduate Medical Education does not accredit fellowship programs in General Academic Pediatrics, 30 divisions were selected randomly from the Academic Pediatric Association's listing of fellowship programs. For each specialty (gastroenterology, nephrology, and general academic pediatrics) an Assistant Professor, Associate Professor, and Professor were chosen randomly from each of the 30 institutions. Names of faculty were gathered from institutional Web sites and phone calls to departments. Random samples were chosen by a computerized random number generator. Only MD full-time faculty was included. In cases in which an institution did not have a faculty member of a particular rank, the field was omitted.

We used Google Scholar via Harzing's Publish or Perish[6] as the primary database to compute bibliometrics. Harzing's Publish or Perish (www.harzing.com/pop.htm) analyzes raw data from Google Scholar to calculate bibliometrics.

The h- and g-indices were used as benchmarks because they are widely used indices in academic promotion. To calculate h- and g-index values and gather statistics on total numbers of articles, total number of citations, number of citations of the highest cited article, and g-index value, each author's surname and first initial was entered into the bibliographic database. The searches included subject fields of "Biology, Life Sciences, Environmental Sciences," "Medicine, Pharmacology, Veterinary Sciences," and "Social Sciences and Humanities." The initial list of articles was then sorted to include only those of the target author. If an initial inquiry resulted in more than 100 articles, then the full first name of the author was used to further focus the results.

Each reference was reviewed, linking to the article if necessary, to ensure that it was authored by the intended faculty member. Isolating the target author's works included analyzing subject matter (title), publication source (name of journal), year published, coauthors, and author institutional affiliation. Results were filtered to exclude all nonpeer-reviewed journal publications such as patents, book chapters, and presentations. Non-English language publications also were excluded. Incorrect and excluded references were removed and bibliometrics recalculated. If a name produced more than 150 articles, the results were double-checked by a second investigator. Investigators then met to reach a consensus. If a name produced more than 1000 articles, the author's middle initial was then included to ensure more accuracy in results. These results also were double-checked by a second investigator.

Recognizing that h-index calculation using different bibliographic databases of academic publications may compile different references and citations that result in different indices, we calculated the bibliometrics for the 30 chairs in Google Scholar as well as another bibliographic database, Scopus, as a comparison. Scopus is a subscription database that is available from Elsevier (www.scopus.com) that cross-references citation and references from 15 000 peer-reviewed journals, conference proceedings, and book series which were published after 1995.

Statistical Analyses

Means with SDs, range, median and mode of total articles, total citations, highest citations, h-index, and g-index were calculated for each faculty member. Differences in h-index between rank within each specialty and between specialties at the same rank were compared by the use of ANOVA analyses. The h-indices were normally distributed and bound by a lower limit of zero. The correlation coefficient between the h-indices from Publish or Perish and Scopus was calculated for the sample of pediatric department chairs. The correlation coefficient between h- and g-indices was calculated from Publish or Perish for all individuals analyzed for each subspecialty.

Results

For each of the specialties, all bibliometric measures increased with academic rank (**Table**). For general academic pediatrics, gastroenterology, and nephrology, the goal of an n of 30 schools with 3 professors each (1 full, 1 associate, 1 assistant) could not be reached. In all 3 specialties, some of the randomly selected schools did not have a professor for each of the desired ranks.

For all academic ranks, there was a wide range in all bibliometric measures. The mean h- and g-indices significantly increased with academic rank (all $P < .001$) with the greatest indices among chairs. There was no statistical difference between the mean h- and g-indices at the same rank in different pediatric subspecialties (Assistant Professors, $P = .07$; Associate Professors, $P = .14$; Professors, $P = .69$) (**Figure 1**). Pediatric chairs had higher means in every bibliometric measure than the professors for each specialty. The mean h-index for pediatric chairs was 26.5 (SD 19.5).

To compare calculations between bibliographic databases, Google Scholar via Publish or Perish and Scopus were both used for pediatric chairpersons. In Scopus the findings were as follows: h-index of 15.6 (4), total articles 83.7 (74.3), total cites 2616.1 (3335.3), highest cited 258 (383). The g-index was not calculated in this software. The correlation coefficient between the h-index of the chairs between Scopus and Publish or Perish was 0.79. The correlation coefficient between h- and g-indices from Google Scholar via Publish or Perish for all individuals analyzed for each subspecialty was 0.98 (**Figure 2**).

Discussion

This study provides initial benchmarks for the h- and g-indices in pediatrics and selected pediatric subspecialties.

Table. Bibliometrics for pediatric specialties and chairs

	Total articles (SD)	Total cites (SD)	Highest cited (SD)	h-index (SD)	g-index (SD)
General Academic Pediatrics					
Professors (n = 28)	85.7 (71.7)	2097.5 (2667.1)	283.8 (346.9)	19.4 (12.8)	35.5 (26.8)
	Range: 2-305	Range: 5-10 440	Range: 4-1455	Range: 1-48	Range: 2-101
	Median: 86	Median: 946	Median: 173	Median: 16	Median: 29
	Mode: 10	Mode: NA	Mode: NA	Mode: 25	Mode: 10
Associate (n = 29)	30.3 (43.5)	429.3 (810.7)	86.9 (105.1)	7.8 (6.9)	14.5 (12.8)
	Range: 2-176	Range: 2-4306	Range: 1-512	Range: 1-34	Range: 1-64
	Median: 15	Median: 168	Median: 56	Median: 6	Median: 11
	Mode: 14	Mode: 208	Mode: NA	Mode: 5	Mode: 14
Assistant (n = 29)	4.7 (8.1)	41.4 (76.4)	17.5 (27.7)	1.8 (2.4)	3.2 (4.3)
	Range: 0-41	Range: 0-356	Range: 0-112	Range: 0-11	Range: 0-18
	Median: 2	Median: 11	Median: 6	Median: 2	Median: 2
	Mode: 0	Mode: 0	Mode: 0	Mode: 0	Mode: 0
Gastroenterology					
Professors (n = 27)	109.8 (69.8)	2126.3 (1206.9)	263.1 (203.7)	21.5 (9.6)	40.3 (17.5)
	Range: 8-251	Range: 41-7611	Range: 24-1015	Range: 3-45	Range: 6-81
	Median: 91	Median: 1493	Median: 230	Median: 20	Median: 37
	Mode: NA	Mode: NA	Mode: 276	Mode: 17	Mode: 47
Associate (n = 25)	34.4 (18.5)	505.3 (472.7)	145.3 (147.6)	9.6 (5.4)	18.8 (9.6)
	Range: 3-76	Range: 21-1748	Range: 7-500	Range: 2-23	Range: 3-39
	Median: 28	Median: 475	Median: 90	Median: 8	Median: 21
	Mode: 22	Mode: 21	Mode: NA	Mode: 5	Mode: 23
Assistant (n = 28)	5.4 (6.7)	78.5 (214.5)	32.2 (73.3)	1.8 (2.6)	3.6 (5.7)
	Range: 0-27	Range: 0-913	Range: 0-314	Range: 0-10	Range: 0-24
	Median: 4	Median: 5	Median: 3	Median: 1	Median: 2
	Mode: 1	Mode: 0	Mode: 0	Mode: 0	Mode: 0
Nephrology					
Professors (n = 27)	83.9 (52.5)	2073.3 (1643.1)	271.5 (267.6)	21.8 (11.5)	37.6 (20.2)
	Range: 4-206	Range: 27-5589	Range: 23-1312	Range: 2-43	Range: 4-70
	Median: 95	Median: 5589	Median: 203	Median: 25	Median: 42
	Mode: 95	Mode: NA	Mode: 275	Mode: 25	Mode: 16
Associate (n = 25)	32.1 (20.9)	513.8 (461.0)	88.1 (68.6)	11.1 (5.4)	19.2 (9.3)
	Range: 2-101	Range: 25-1858	Range: 12-313	Range: 1-24	Range: 2-41
	Median: 25	Median: 369	Median: 73	Median: 11	Median: 19
	Mode: 24	Mode: 126	Mode: 123	Mode: 11	Mode: 24
Assistant (n = 28)	8.7 (6.2)	82.3 (77.4)	33.4 (31.2)	3.2 (2.8)	8.5 (8.3)
	Range: 1-39	Range: 0-2439	Range: 0-1822	Range: 0-14	Range: 0-39
	Median: 10	Median: 64	Median: 33	Median: 3	Median: 7
	Mode: 1	Mode: 63	Mode: 15	Mode: 3	Mode: 7
Pediatric Department Chair (n = 30)	111.9 (93.4)	3436.6 (4318.5)	379.1 (526.2)	26.5 (19.5)	51.1 (37.4)
	Range: 11-297	Range: 19-15 101	Range: 9-2704	Range: 2-63	Range: 4-121
	Median: 82	Median: 1502	Median: 205	Median: 23	Median: 41
	Mode: 129	Mode: 12 578	Mode: 95	Mode: 35	Mode: 17

NA, not applicable.

The h-index is becoming increasingly recognized as a measure to quantify a researcher's academic productivity.[17] The h-index attempts to combine impact and output and has been shown to have utility in predicting future academic achievement.[3] Benchmarking bibliometric indices, like the h- and g-index, is vital if these metrics are being used as a measure of productivity, particularly for academic promotion.

Our results show an increase in h- and g-indices with increasing academic rank for General Academic Pediatrics and the 2 specialties examined. It was not surprising that there was a significant difference in h-index between different faculty ranks with the highest indices among Professors and lowest among Assistant Professors.

Further, our results demonstrate a wide range of values for all bibliometric indices within all ranks. This is likely due to different standards for promotion across institutions. Further, it may be that some faculty members were on promotion tracks that placed less emphasis on publications.

One limitation of the h-index is the difficulty in comparing it across fields. Hirsch found that there was a wide difference in citation rates and numbers of publications among fields with the h-index being highest in the life sciences. He reported greatest values for physicists with the average for associate professor of 12 and full professor 18.[3] Jeang[18] calculated h-indices between 12-39 for the editorial board for the *Journal of Retrovirology*. Kelly and Jennions[19] found a mean h-index of 45 for "highly cited: evolutionists and ecologists." Thus, it has been suggested that in the basic sciences there is value of using this bibliometric within fields or in fields that are closely related.

Although bibliometrics are known to vary by specialty and subspecialty related to different number of citations, size of publishing groups, and size of the academic field, this was not found among our study of pediatric nephrologists, gastroenterologists, and general academic pediatricians.

For the field of medicine, the h-index has been examined in the fields of neurosurgery, academic otolaryngology, and

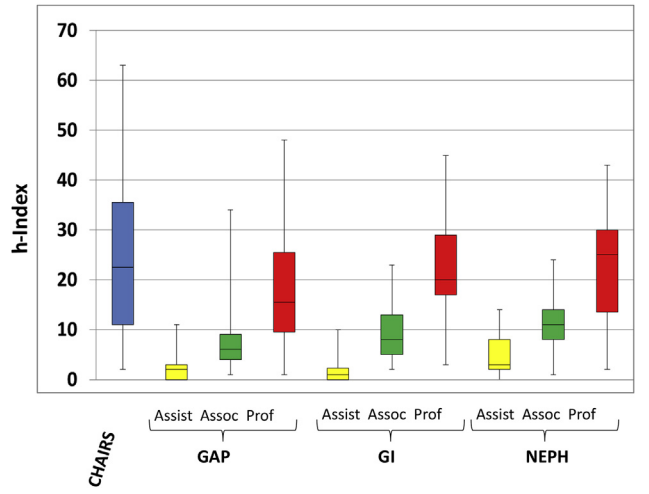

Figure 1. The h-index by Rank and Specialty: box and whisker plot showing median (*line in box*); 25th and 75th percentiles (*box ends*) and min and max (*whiskers*). *Assist*, Assistant Professor; *Assoc*, Associate Professor; *GAP*, General Academic Pediatrics; *GI*, Gastroenterology; *NEPH*, Nephrology; *Prof*, Professor.

radiation oncology. Lee et al[8] found mean indices for neurosurgery of 5.1 for Assistant Professors, 10.7 for Associate Professors, 16 for Professors, and 24.7 for Chairpersons. In an analysis of academic otolaryngologists via the Scopus database we also found a greater h-index with increasing academic rank, including 4.6 (0.26 SEM) for Assistant Professor, 8.13 (0.43 SEM) for Associate Professor, and 15.6 (0.74 SEM) for Professor.[9] These benchmarks for neurosurgery were greater than those that we found for pediatrics, and the benchmarks for academic otolaryngologists were lower. It may not be appropriate to compare h-indices across medical specialties. There are multiple reasons that may account for these differences, including percent of clinical practitioners in comparison with researchers, size of the field,

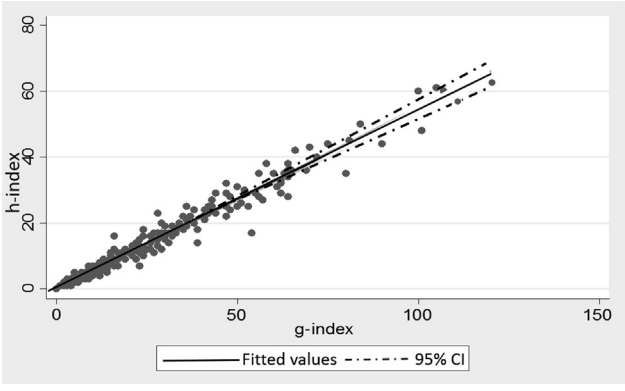

Figure 2. Correlation matrix between h- and g-indices from Publish or Perish for all individuals analyzed for each subspecialty (correlation coefficient = 0.98).

value of journal publications, and number of journals in the field.

Despite the potential advantages of enumerating academic productivity in one number, there are potential drawbacks to using the h-index. For example, the h-index can be misleading without knowing the context of citations. A paper discussed frequently because of inaccuracies or exaggerated claims may result in a disproportionately high measure of impact. For example, the article by Wakefield et al in the *The Lancet* citing a link between the Measles, Mumps, and Rubella vaccine and autism that was later retracted received many citations.[20] There is also the potential for the "Matthew effect," in which a renowned researcher may be disproportionately cited more than a lesser-known researcher.[21] Other concerns about the h-index include self-citation, lack of accounting for placement in the author list, bias toward more years in practice, and a potential sex bias as the result of more name changes among women who have married that may not be captured in bibliometric calculation.[19,22] Further, because it is a whole number, there is a reduced ability for discrimination.

There was a high correlation between g- and h-indices. Because the h-index incorporates number of papers and citations, it could be argued that given this high concordance the h-index may be a superior. Further, the h-index has been studied more widely and has more established benchmarks.

It was interesting to note the significantly lower h-indices generated by Scopus vs Google Scholar via Harzing's Publish or Perish. This is likely because Scopus includes few papers published before 1996. Thus the h-indices were lower especially for faculty members who had publications before that time. This limited coverage of older dates of publication and smaller breadth of content coverage, including less of the social sciences and humanities, may be drawbacks of using Scopus and other search engines like Web of Knowledge for bibliometric calculations.

There are potential limitations to this study. By randomly selecting individuals from a random list of accredited training programs we attempted to include academic faculty and limit selection bias. We believe that this cohort is representative of these pediatric specialties; however, this has not been validated by an independent sample. Faculty lists were obtained from institutional Web sites or phone calls. Some Web sites may have been out of date, although it is unclear how this would systematically bias the results.

In the bibliometric calculation process error could have been introduced by searching the incorrect individual or including papers by another author with a similar name. However, the search database often allowed for the search to include institution name. Common surnames and individuals changing surnames throughout their career are challenges to ensuring bibliometric calculation accuracy. Further, because some nonpeer-reviewed literature is included in Google Scholar, it is possible that some of these nonpeer-reviewed manuscripts were included unintentionally in the calculations.[23] This possible overcounting of publications would not be a detriment because the overcounting

would be unbiased across researchers if all measures (h-index and g-index) were derived from Google Scholar. We attempted to address these limitations by having a second investigator check searches with greater than 150 results and manually deleting nonpeer-reviewed results.

The h-index and g-index are indictors of productivity and impact to be considered in the context of a larger portfolio in the performance review and promotion process. Our data combined with the outcomes from previous studies on bibliometric indices in other specialties indicate that indices used for promotion should be considered within the context of an individual specialty and institution. ■

Submitted for publication Dec 19, 2014; last revision received Aug 10, 2015; accepted Oct 8, 2015.

Reprint requests: Megan M. Tschudy, MD, MPH, Department of Pediatrics, The Johns Hopkins University School of Medicine, Rubenstein Health Building, 200 North Wolfe Street, Suite 2021, Baltimore, MD 21287. E-mail: mtschud1@jhmi.edu

References

1. Costas R, Bordons M. Is the g-index better than h-index? An exploratory study at the individual level. Scientometrics 2008;77:267-88.
2. Egghe L. Theory and practice of the g-index. Scientometrics 2006;69:131-52.
3. Hirsch JE. Does the *h* index have predictive power? Proc Natl Acad Sci U S A 2007;104:19193-8.
4. Scopus. http://www.scopus.com/home.url. Accessed March 10, 2014.
5. Web of knowledge. http://wokinfo.com. Accessed March 10, 2014.
6. Harzing's publish or perish. http://www.harzing.com/pop.htm. Accessed March 10, 2014.
7. van Haselen R. The h-index: a new way of assessing the scientific impact of individual CAM authors. Complement Ther Med 2007;15:225-7.
8. Lee J, Kraus KL, Couldwell WT. Use of the h index in neurosurgery. J Neurosurg 2009;111:387-92.
9. Svider PF, Choudry ZA, Choudry OJ, Baredes S, Liu JK, Eloy JA. The use of the h-index in academic otolaryngology. Laryngoscope 2012;123:103-6.
10. Quigley MR, Holliday EB, Fuller CD, Choi M, Thomas CR Jr. Distribution of the h-index in radiation oncology conforms to a variation of power law: implications for assessing academic productivity. J Cancer Educ 2012;27:463-6.
11. Thomson Reuters Social Sciences Citation Index. http://thomsonreuters.com/products_services/science/science_products/a-z/science_citation_index/#tab4. Accessed July 10, 2015.
12. Hirsch JE. An index to quantify an individual's scientific research output. Proc Natl Acad Sci U S A 2005;102:16569-72.
13. Penner O, Pan RK, Petersen AM, Kaski K, Fortunato S. On the predictability of future impact in science. Sci Rep 2013;3:3052.
14. Penner O, Pan RK, Petersen AM, Kaski K, Fortunato S. The case for caution in predicting scientist's future impact. Phys Today 2013;66:8-9.
15. Butson MJ, Yu PKN. The first author h-index (h(fa)-index): Leveling the field for small and large institute medical science scholars. Australas Phys Eng Sci Med 2010;33:299-300.
16. Romanovsky AA. Revised h index for biomedical research. Cell Cycle 2012;11:4118-21.
17. Ball P. Academic index climbs the ranks. Nature 2007;448:2007.
18. Jeang KT. Impact factor, H index, peer comparisons, and retrovirology: is it time to individualize citation metrics? Retrovirology 2007;4:42.
19. Kelly CD, Jennions MD. H-index: age and sex make it unreliable. Nature 2007;449:403.
20. Eggertson L. Lancet retracts 12-year-old article linking autism to MMR vaccines. CMAJ 2010;182:E199-200.
21. Wendl MC. H-index: however ranked, citations need context. Nature 2007;449:403.
22. Bartneck C, Kokkelmans S. Detecting h-index manipulation through self-citation analysis. Scientometrics 2011;87:85-9.
23. Petersen AM, Fortunato S, Pan RK, Kaski K, Penner O, Rungi A, et al. Reputation and impact in academic careers. Proc Natl Acad Sci U S A 2014;111:15316-21.