Hirsch Index Value and Variability Related to General Surgery in a UK Deanery

Tarig Abdelrahman, BM, Josephine Brown, MBBCh, Jenny Wheat, MBBCh, Charlotte Thomas, MBChB, MSc and Wyn Lewis, MD, DSc

Department of General Surgery, University Hospital of Wales, Cardiff, Wales, United Kingdom

OBJECTIVES: The Hirsch Index (*h*-index) is often used to assess research impact, and on average a social science senior lecturer will have an *h*-index of 2.29, yet its validity within the context of UK General Surgery (GS) is unknown. The aim of this study was to calculate the *h*-indices of a cohort of GS consultants in a UK Deanery to assess its relative validity.

DESIGN: Individual *h*-indices and total publication (TP) counts were obtained for GS consultants via the Scopus and Web of Science (WoS) Internet search engines. Assessment of construct validity and reliability of these 2 measures of the *h*-index was undertaken.

SETTING: All hospitals in a single UK National Health Service Deanery were included (14 general hospitals).

PARTICIPANTS: All 136 GS consultants from the Deanery were included.

RESULTS: Median *h*-index (Scopus) was 5 (0-52) and TP 15 (0-369), and strong correlation was found between *h*-index and TP ($\rho = 0.932$, $p < 0.001$), with the intraclass correlation between Scopus and WoS *h*-index also significant (intraclass correlation coefficient = 0.973 [95% CI: 0.962-0.981], $p < 0.001$). Academic GS consultants had higher *h*-indices than nonacademic University Hospital and District General Hospital consultants (Scopus 12 vs 7 vs 4 [$p < 0.001$] and WoS 10.5 vs 7 vs 4 [$p < 0.001$]). *h*-Index was >2.29 in 57.4% of consultants. No subspecialty differences were apparent in median *h*-indices ($p = 0.792$) and TP ($p = 0.903$).

CONCLUSIONS: *h*-Index is a valid GS research productivity metric with over half of consultants performing at levels equivalent to social science Senior Lecturers. (J Surg Ed 73:111-115. © 2015 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: general surgery, training, bilbiometrics, academic, research

COMPETENCIES: Medical Knowledge, Professionalism, Practice-Based Learning and Improvement

INTRODUCTION

The Hirsch Index (*h*-index), a measure of research impact, was first introduced in 2005 by Jorge E. Hirsch to objectively quantify an individual's cumulative research productivity and address the limitations of other traditional bibliometric indicators such as journal impact factor, publication number, and citations.[1] It has rapidly gained favor for its emphasis on rewarding authors who publish work with a measurable effect in their field, and has been shown to outperform other bibliometrics when evaluating research performance within surgical specialties.[2-10]

The main advantage of the *h*-index is that it incorporates 2 traditional measures of research productivity; the number of publications (quantity) and citations (quality). An author has an index h, if h of their N_p articles have at least h citations each, and their other articles ($N_p - h$) have less than or equal to h citations each.[1] Therefore an author with an *h*-index of 10 will have 10 published articles that have each been cited at least 10 times. Because of this and its apparent insensitivity to authors that publish a high volume of low-impact articles or a low volume of high-impact articles, *h*-index has been adopted in many fields of science as a robust bibliometric. Although self-citation can have an effect, the *h*-index appears to be more resilient to this than other traditional metrics,[10-12] and can be readily calculated using online bibliographic search engines, the most common being Elsevier's Scopus and Thomson Reuters' Web of Science (WoS). A UK study of 120 academics within the Social Sciences showed that on average full professors had an *h*-index of 4.97 and senior lecturers had 2.29.[13]

Within the context of UK General Surgery (GS) *h*-index validity has not been examined; indeed only 1 Canadian study reports its use in GS and found significant differentiation related to academic rank and institutional affiliation.[2] The aim of this study was to assess the value, variability, and validity of the *h*-index as a marker of

Correspondence: Inquiries to Tarig Abdelrahman, BM, Department of General Surgery, University Hospital of Wales, Heath Park, Cardiff, Wales CF14 4XW, UK; e-mail: tarig007@yahoo.com

1931-7204/$30.00
http://dx.doi.org/10.1016/j.jsurg.2015.07.014

research performance and academic training potential for a cohort of GS consultant trainers within a single UK Deanery. This is of particular contemporary relevance in light of the recent publication of the 2013 iteration of the UK Joint Committee on Surgical Training GS Curriculum, which mandates that Higher Surgical Trainees must possess at least 3 peer-reviewed publications for successful award of a Certificate of Completion of Training in GS.[14]

MATERIALS AND METHODS

All consultant general surgeons involved in training within the Wales Deanery were identified through trainee portfolios, hospital departmental websites and deanery records. Surgeons that had retired or moved out of region within the last 2 years were included. All data were collected in January 2015. Surgeons were categorized into: Academic (professor, reader, and senior lecturer, including honorary titles); University Hospital (all nonacademic National Health Service GS working at a University Hospital); or District General Hospital (DGH) (all nonacademic National Health Service DGH GS). The collection of data and calculation of the *h*-index were obtained from 2 online bibliographic databases, Elsevier's SciVerse Scopus and Thomson Reuters' WoS.

Scopus

Using the author search function in Scopus, the surgeon's surname and initials were entered, initially using only 1 initial to keep the search as broad as possible. All publications were reviewed and added to the My List function using the search results format. Using the view citation overview function the *h*-index, publication number and citation number were recorded. Mean citation per publication was calculated independently from the above numbers.

Web of Science

Using the author finder function in WoS, the same search was conducted. Subject categories were limited to life sciences, social sciences, biomedicine, multidisciplinary science, and technology. Abstracts and conference articles were excluded. Relevant publications were then identified and the create citation report function was used to calculate the *h*-index, publication number, citation number, and a mean citation per publication.

If an author was not found on either search engine then the *h*-index was assumed to be 0. Each search was conducted on the same day to minimize bias, and all searches were completed within 3 days, by 2 authors T.A. and J.B. independently. Significant differences in the final *h*-index and publication numbers obtained from both search engines were reviewed, repeated, and cross-checked before inclusion in the study.

An assessment of construct validity for this metric was undertaken by evaluating the differences in *h*-index among Academic, University Hospital, and DGH GS with the expectation that any valid marker of scholarly productivity would provide a higher value in surgeons with academic interests. As a secondary outcome measure, the impact of academic status and institution on the other bibliometrics was measured, including publication numbers, citation numbers, and mean citations per publication.

Statistical Analysis

Data were collected and analyzed in SPSS version 20 (SPSS, Chicago, IL). Traditional measures of research productivity were compared and assessment of the intervariable reliability of the databases used for *h*-index calculation was performed. The null hypotheses were (1) *h*-index would not relate to academic status, (2) *h*-index would not relate to institutional affiliation, and (3) no intervariable disagreement in *h*-index existed related to database used. Continuous data were explored for normality using the Kolmorgorov-Smirnov test. As data sets did not conform to a normal distribution, analyses appropriate for nonparametric data were used in the form of Mann-Whitney U and Kruskal-Wallis H tests. An assessment of reliability for the 2 measures of *h*-index was obtained using a Bland-Altman plot in which both the mean bias (the average difference of the *h*-index between the 2 measures) and the 1.96 SD agreement limits (a measurement of the range of differences) were calculated. This test provides a comparative analysis of 2 methods used to measure the same variable when neither is recognized as the gold standard. The use of the Intraclass Correlation Coefficient allowed quantification of the degree of consistency between the 2 search engines.

RESULTS

Data were available relating to 136 GS consultants in 14 hospitals across Wales. Subspecialty interests numbered 7, the largest cohort was colorectal surgery (47 consultants), and the smallest endocrine (3 consultants); 16 consultants were in possession of an academic title (5 substantive university status and 11 honorary titles). The overall median number of publications declared by Scopus and WoS was 15 (0-369) vs 12.5 (0-291), respectively; *h*-indices were 5 (0-52) vs 5 (0-49); median citation counts were 119 (0-1,2233) vs 130 (0-1,0239), and citations per article were 8.89 (0-46.45) vs 8.85 (0-43.58), respectively. Median *h*-indices (Scopus vs WoS) by subspecialty were: endocrine 7 vs 7, vascular 6.5 vs 7, upper gastrointestinal 6.5 vs 6, hepatobiliary 6.5 vs 5, transplant 5 vs 5, breast 5 vs 3, and colorectal 4 vs 4 ($p = 0.802$ vs $p = 0.541$).

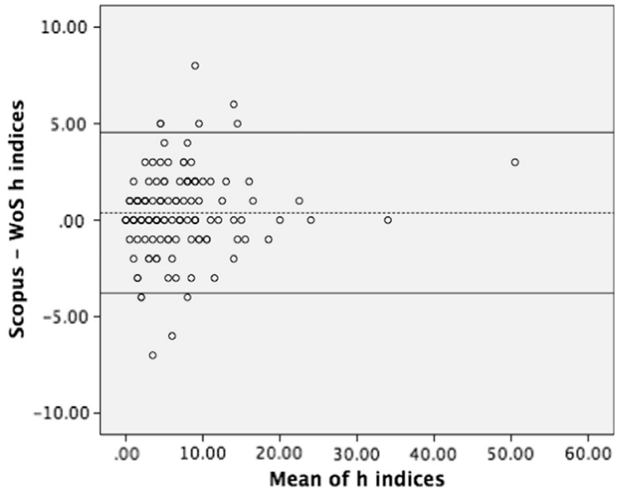

FIGURE. The Bland-Altman plot for Scopus and Web of Science *h*-indices. The mean bias (the mean of the differences between the *h*-indices calculated by the 2 search engines) is represented by the dotted line with the solid lines representing ±1.96 standard deviation of this bias.

Assessment of Reliability

Using the Bland-Altman box plot, the mean bias (agreement) between Scopus and WoS was fair at 0.382 *h*-index points (95% CI: −3.777554 to 4.542), but the range in agreement was considerable (−7 to 8, Fig.). The Intraclass Correlation Coefficient was 0.973 (95% CI: 0.962-0.981, $p < 0.001$), which indicates a high degree of reliability between measurements (Fig.).

Assessment of Construct Validity

The Table shows that the median *h*-indices of GS consultants in possession of an Academic title was significantly higher than University Hospital and DGH nonacademics.

When Academic surgeons were excluded the difference in the *h*-indices of University Hospital vs DGH GS alone reached statistical significance 7 vs 4 ($p < 0.001$) with both Scopus and WoS. The difference between Academic and University Hospital GS was also highly significant (Scopus, $p = 0.003$ and WoS, $p = 0.015$).

Comparison With Other Bibliometric Measures

h-Index correlated with publication number (Scopus: $\rho = 0.927$, $p < 0.001$; WoS: $\rho = 0.917$, $p < 0.001$), total citations (Scopus: $\rho = 0.959$, $p < 0.001$; WoS: $\rho = 0.939$, $p < 0.001$), and mean citations per publication (Scopus: $\rho = 0.656$, $p < 0.001$; WoS: $\rho = 0.573$, $p < 0.001$). Publication numbers were significantly higher in Academic than in University Hospital and DGH GS using both Scopus ($p < 0.001$), and WoS ($p < 0.001$) (Table).

Academic GS consultants had significantly higher total citation numbers than University Hospital and DGH consultants using both search engines (Table).

Mean citations per article were also higher in Academic GS consultants using Scopus ($p = 0.001$), but no discernable difference was apparent using WoS, 9.93 vs 10 vs 7.89 ($p = 0.155$) (Table).

DISCUSSION

This is the first UK study to report the relative academic productivity related to a cohort of general surgeons as measured by the Hirsch Index, a bibliometric that attempts to evaluate both the quantity and influence of an author's academic publications. The principal findings were that 85% of surgeons (116) had published 3 or more articles with the median numbers being 15 (Scopus) and 12.5 (WoS), respectively. The median *h*-index of the full cohort of Wales Deanery general surgeons was 5 using both the Scopus and WoS search engines. Publications were not apparent for 8% (Scopus) and 3% (WoS) of consultants. Higher *h*-indices were associated with surgeons who were in possession of substantive and honorary academic and University titles when compared with their DGH surgeon counterparts, which is in keeping with reports from a number of other surgical and medical specialties allied to surgery (such as ophthalmology, neurosurgery, and anesthesia).[2-10,15-17]

Differences between the *h*-indices calculated using the 2 online databases were apparent, but overall statistical agreement between the 2 was good, with both providing very similar results for the different groups of surgeons. Sharma et al.[2] also compared Scopus and WoS in their cohort of general surgeons, and although average agreement between the 2 search engines was excellent, the range of agreement was considerable with few outlying cases that risk skewing concordance, particularly relating to publications that were older in certain journals. This variation can be

TABLE. *h*-Index Construct Validity and Bibliometrics Related to Surgeon Status

	Scopus				WoS			
Consultant (*n* = 136)	**HI**	**TP**	**C**	**C/P**	**HI**	**TP**	**C**	**C/P**
Academic (*n* = 16)	12 (1-52)	38	520	13.01	10.5 (1-49)	29.5	332.5	9.93
University Hospital (*n* = 39)	7 (0-24)	18	172	9.23	7 (0-24)	18	174	10
DGH (*n* = 81)	4 (0-18)	10	80	6.5	4 (0-19)	9	93	7.89
p Value	<0.001	<0.001	<0.001	0.001	<0.001	<0.001	<0.001	0.155

Values are medians. Ranges in parentheses. C, citations; C/P, citations per paper; HI, Hirsch index; TP, total publications.

explained by differing database journal coverage and date ranges of citations that have been found to produce quantitatively different citation counts and hence *h*-indices. Scopus for example, only accounts for publications that appeared after 1995, disadvantaging senior authors.[15,18,20,21]

The *h*-index as a marker of academic achievement appears impervious to authors that publish frequently low-impact research or those that may have a misleadingly high citation rate reflective of only 1 or 2 influential articles. However, the *h*-index has been criticized for a number of shortcomings, despite its obvious advantages over traditional metrics. The inability to compare scientists belonging to different fields or subspecialties, and to differentiate authorship order within the final manuscript can be significant.[19] Researchers with little input to an article can earn the same recognition as the primary author.[1] Also the type of research is not accounted for, so a potentially influential multicentre randomized control trial is not differentiated from small retrospective case series.[1] The *h*-index will also increase with time as older articles are more frequently cited, which will tend to favor senior researchers, a factor that needs to be considered when comparisons are made.

This study has a number of potential limitations. Difficulty in obtaining reliable bibliographic outputs for individuals with common names was a practical challenge. Individual authors can be identified based on name and affiliation but researchers frequently change or use different initials and affiliations confounding search results. Potential inaccuracies were minimized in this study by the use of 2 reviewers performing each search separately with differences cross-referenced and anomalies excluded. The study focused on 1 UK Deanery and geographical region but a UK wide study would produce higher-powered results. Moreover, higher numbers of academic surgeons, would facilitate further classification related to academic seniority allowing a more precise measure of validity to be performed. Honorary university titles are frequently awarded according to other professional contributions such as medical education rather than research expertise and this arguably represents a potential flaw in using this model as an assessment of construct validity.[2,23] Differences in database citation counts obtained can also produce *h*-indices that vary quite significantly, and several studies have reported potential disagreement between databases.[2,15,18,20,21]

CONCLUSION

h-Index value has been reported to predict future academic productivity in a number of studies.[5,22] The findings of this study suggest that the use of the *h*-index provides a more robust measure of a surgeon's academic profile than traditional bibliometrics. It is evident that the level of published scholarly activity among General Surgeons is high, with *h*-indices equivalent to or above that of a Social Sciences Senior Lecturer (2.29) and Professor (4.97) in 57% and 40% of consultants. In the context of contemporary UK surgical training this metric may be used to identify where opportunities lie for trainees to engage in academic activity likely to result in publication. It seems self-evident that rotations that include time on an academic consultant firm or at a University Teaching Hospital would be associated with improved chances of a trainee achieving publications. With the most recent iteration of the Joint Committee on Surgical Training general surgical curriculum (2013) setting an indicative target number of 3 peer-reviewed publications during Higher Surgical Training, data such as this are valuable in allowing training program directors and health education providers alike access to information to plan and tailor individual rotational programs appropriately so that Certificate of Completion of Training requirements are achieved.[14]

REFERENCES

1. Hirsch JE. An index to quantify an individual's scientific research output. *Proc Natl Acad Sci U S A*. 2005;102(46):16569-16572.
2. Sharma B, Boet S, Grantcharov T, Shin E, Barrowman NJ, Bould MD. The h-index outperforms other bibliometrics in the assessment of research performance in general surgery: a province-wide study. *Surgery*. 2013;153(4):493-501.
3. Colaco M, Svider PF, Mauro KM, Eloy JA, Jackson-Rosario I. Is there a relationship between National Institutes of Health funding and research impact on academic urology? *J Urol*. 2013;190(3):999-1003.
4. Svider PF, Lopez SA, Husain Q, Bhagat N, Eloy JA, Langer PD. The association between scholarly impact and National Institutes of Health funding in ophthalmology. *Ophthalmology*. 2014;121(1):423-428.
5. Lee J, Kraus KL, Couldwell WT. Use of the h index in neurosurgery. Clinical article. *J Neurosurg*. 2009;111(2):387-392.
6. Paik AM, Mady LJ, Villanueva NL, et al. Research productivity and gender disparities: a look at academic plastic surgery. *J Surg Educ*. 2014;71(4):593-600.
7. Ponce FA, Lozano AM. Academic impact and rankings of American and Canadian neurosurgical departments as assessed using the h-index. *J Neurosurg*. 2010;113(3):447-457.
8. Benway BM, Kalidas P, Cabello JM, et al. Does citation analysis reveal association between h-index and academic rank in urology? *Urology*. 2009;74(1):30-33.

9. Svider PF, Pashkova AA, Choudhry Z, et al. Comparison of scholarly impact among surgical specialties: an examination of 2429 academic surgeons. *Laryngoscope*. 2013;123(4):884-889.

10. Eloy JA, Svider PF, Mauro KM, Setzen M, Baredes S. Impact of fellowship training on research productivity in academic otolaryngology. *Laryngoscope*. 2012;122(12):2690-2694.

11. Callaham M, Wears R, Weber E. Journal prestige, publication bias, and other characteristics associated with citation of published studies in peer-reviewed journals. *J Am Med Assoc*. 2002;287(21):2847-2850.

12. Engqvist L, Frommen JG. The h-index and self-citations. *Trends Ecol Evol*. 2008;23(5):250-252.

13. *The London School of Economics*. Impact of social sciences—3: key measures of academic influence. Available at: ⟨http://blogs.lse.ac.uk/impactofsocialsciences/the-handbook/chapter-3-key-measures-of-academic-influence/⟩ Accessed 03.04.15.

14. *Joint Committee on Surgical Training*. Guidelines for the award of a CCT in general surgery. Available at: ⟨http://www.jcst.org/quality-assurance/documents/cct-guidelines/general-surgery-cct-guidelines⟩; 2013.

15. Bould MD, Boet S, Sharma B, Shin E, Barrowman NJ, Grantcharov T. h-Indices in a university department of anesthesia: an evaluation of their feasibility, reliability, and validity as an assessment of academic performance. *Br J Anaesth*. 2011;106(3):325-330.

16. Pagel PS, Hudetz JA. Scholarly productivity of United States academic cardiothoracic anesthesiologists: influence of fellowship accreditation and transesophageal echocardiographic credentials on h-index and other citation bibliometrics. *J Cardiothorac Vasc Anesth*. 2011;25(5):761-765.

17. Choi M, Fuller CD, Thomas CR Jr. Estimation of citation- based scholarly activity among radiation oncology faculty at domestic residency-training institutions: 1996-2007. *Int J Radiat Oncol Biol Phys*. 2009;74(1):172-178.

18. Falagas M, Pitsouni E, Malietzis G, Pappas G. Comparison of PubMed, Scopus, web of science, and Google scholar: strengths and weaknesses. *FASEB J*. 2008;22(2):338.

19. Cronin B, Meho L. Using the h-index to rank influential information scientists. *J Am Soc Inf Sci Technol*. 2006;57(9):1275-1278.

20. Patel VM, Ashrafian H, Almoudaris A, et al. Measuring academic performance for healthcare researchers with the H index: which search tool should be used? *Med Princ Pract*. 2013;22(2):178-183.

21. Kulkarni A, Aziz B, Shams I, Busse J. Comparisons of citations in Web of Science, Scopus, and Google Scholar for articles published in general medical journals. *J Am Med Assoc*. 2009;302(10):1092.

22. Hirsch J. Does the h-index have predictive power? Proceedings of the National Academy of Sciences. 2007;104(49):19193-19198.

23. Fuller CD, Choi M, Thomas CR. Biliometric analysis of radiation oncology departmental scholarly publication productivity at domestic residency training institutions. *Am Coll Radiol*. 2009;6(2):112-118.