Relative value of adapted novel bibliometrics in evaluating surgical academic impact and reach

David B.T. Robinson[1], Luke Hopkins[1], Chris Brown[1], Tarig Abdelrahman[1], Arfon G. Powell[2], Richard J. Egan[3], Wyn G. Lewis[1]

[1] Wales Deanery PGMDE School of Surgery, Health Education and Improvement Wales, Cefn Coed, Nantgarw, CF15 7QQ

[2] Cardiff University School of Medicine, Division of Cancer and Genetics, Heath Park, Cardiff, U.K. CF14 4XW.

[3] Department of Surgery, Morriston Hospital, Heol Maes Eglwys, Swansea, U.K. SA6 6NL.

Corresponding author: Mr David Robinson, Wales Deanery PGMDE School of Surgery, Health Education and Improvement Wales, Cefn Coed, Nantgarw, CF15 7QQ.

T: 029 2074 3268

E: David.robinson4@outlook.com

F: 029 2074 4553

Running Head: RVNB to assess surgical academic impact

Conflict of interest: None.

Funding: N/A

Keywords: Bibliometrics, Altmetrics, Surgery

Word count (excluding title page, abstract, figures/tables and references): 2498

Abstract (Word Count: 249)

Background: The Hirsch Index, often used to assess research impact, suffers from questionable validity within the context of General Surgery and consequently adapted bibliometrics and altmetrics have emerged, including the r-index, m-index, g-index and i10-index. This study aimed to assess the relative value of these novel bibliometrics in a single UK Deanery General Surgical Consultant cohort.

Method: Five indices (h, r, m, g, i10) and altmetric scores (AS) were calculated for 151 general surgical consultants in a UK Deanery. Indices and AS were calculated from publication data via the Scopus search engine with assessment of construct validity and reliability.

Results: The median number of publications, h-index, r-index, m-index, g-index and i10-index were 13 (range 0-389), 5 (range 0-63), 5.2 (range 0-64.8), 0.33 (range 0-1.5), 10 (range 0-125) and 4 (range 0-245) respectively. Correlation coefficients of r-index, m-index, g-index and i10-index with h-index were 0.913 ($p < 0.001$), 0.716 ($p < 0.001$), 0.961 ($p<0.001$) and 0.939 ($p<0.001$) respectively. Significant variance was observed when the cohort was ranked by individual bibliometric measures; the median ranking shifts were: r-index -2 (-46 to +23); m-index -6.5 (-53 to +22); g-index -0.5 (-24 to +13); and i10-index 0 (-8 to +11) respectively ($p<0.001$). The median altmetric score and AS-index were 0 (range 0 - 225.5) and 1 (range 0-10) respectively; AS-index correlated strongly with h-index (correlation coefficient 0.390, $p<0.001$).

Conclusions: Adapted bibliometric indices appear to be equally valid measures of evaluating academic productivity, impact and reach.

Introduction

Nothing declares academic reach and scholarly performance more than bibliometrics; the statistical analysis of published works.[1] Paul Otlet, originally described the term bibliométrie in 1934, with the anglicised version, bibliometrics, first used by Alan Pritchard in his 1969 article Statistical Bibliography or Bibliometrics?[2,3] Arguably, the most ubiquitous contemporary bibliometric tool is the Hirsch Index (h-index), a research impact measure, reported by Jorge E. Hirsch in 2005, to quantify tangibly an author's cumulative research output, and address the limitations of erstwhile indicators such as journal esteem, impact factor, publication number, and citations.[4]

The h-index has gained rapid favor because of incorporating two traditional measures of research yield; publication number (quantity) and citations (quality).[5-10] An author has an index h, if h of their Np (number of publications) have at least h citations each, and their other articles (Np - h) have less than or equal to h citations. Its perceived resilience to authors that publish high volumes of low-impact articles or a low volumes of high-impact articles, has fuelled its robust adoption in many scienctific arenas.[7-10] Yet this scenario is not without controversy,[11] the h-index, akin to many metrics, has been criticized for shortcomings including; an inability to compare scientists from diverse fields; differentiating authorship hierarchy; researchers providing little article input can earn similar recognition to the first author; research evidence level is unaccounted for; and the h-index will be boosted by time with older articles are more frequently cited, favoring senior researchers.[12-14] Because of the above, a number of adapted bibliometrics have emerged, including the r-index, m-index, g-index and i10-index and most recently the social media phenomenon of Altmetric Scores.[8] The aim of this study was to quantify the validity of these novel bibliometric markers, in a cohort of UK General Surgical Consultants. This is of particular interest because of the 2013 iteration of the UK Joint Committee on Surgical Training General Surgery Curriculum, which demands a minimum of 3 peer-reviewed publications for Certification of Completion of Training (CCT).[15]

Material and methods

Consultant general surgeons within the Wales Deanery with training responsibilities were identified. Surgeons' status was categorized as: Academic (professor, reader, and senior lecturer, including honorary titles); University Hospital (all non-academic National Health Service General Surgeons working at a University Hospital); and District General Hospital (DGH) General Surgeons. Data and h-indices were obtained from Elsevier's SciVerse Scopus bibliographic database (June 30, 2018). The indices that were collected/calculated and compared were number of publications, h-Index, r-Index, m-Index, g-index and i10-Index. Altmetric scores were also collected for all articles published since 2011, and this methodology has been described previously.[5]

r-index (Revised h-Index)

Proposed by Andrej A. Romanovsky, the r-index allocates a multiplying factor to first and last author publications compared with middle author publications. This is based on the fact that in a biomedical field, the first and last authors of a publication are in general the biggest contributors to the publication's design, data collection, analysis and write-up. The r-index calculation is restricted to publications constituting an author's h-Index, by the formula $r = 1.6a + 0.4(h\text{-}a)$, where a is the number of first or last author publications, and h is the author's Hirsch index.[12]

m-Index

The m-Index is the h-index divided by the number of years since the author's first publication.[14]

g-index

"The *g*-index is the (unique) largest number, such that the top *g* articles received (together) at least g^2 citations" [16]. An author with 4 publications with citations ranked in descending order of 5,4,4,0 would have a g-index score of 3 as the sum of the first three citations is >9, but the sum of all 4 citations is not ≥16.

i10-Index

Introduced by Google Scholar, the i10-index is defined as the number of publications receiving at least 10 citations.[17,18]

Altmetric Scores

Altmetric Scores (AS) for all papers published since 2011 were collected and the methodology used has been described previously (https://www.altmetric.com/products/ free-tools/bookmarklet/).[19-21]

An adapted calculation or AS-index was also calculated akin to the h-index, substituting the number of citations for the AS. Hence an individual has an AS-index of 10, if they have 10 articles that have received at least 10 altmetric citations.

Statistical Analysis

Data were collected and analyzed in SPSS version 25 (SPSS, IBM Corp, Armonk, NK, Chicago, IL), and analyses appropriate for nonparametric data were used.

The null hypotheses were:

(1) h-index is unrelated to academic status;

(2) h-index is unrelated to institutional affiliation;

(3) Ranking by bibliometrics is unrelated to index type.

Results

Data were available for 151 General Surgical consultants in 12 training hospitals; 136 male and 15 female. Subspecialty interests were 7 (largest cohort colorectal surgery (n=47), and smallest endocrine (4)). University hospital consultants comprised 62, and DGH 89 of the total cohort; 18 consultants held academic titles (5 substantive university and 13 honorary).

The overall median publication number was 13 (range 0 to 389) with 87.4% (n=132) achieving a minimum of 3 publications. Median h-index was 5 (0 to 63); r-index 5.2 (0 to 64.8); m-index 0.33 (0 to 1.5); g-index 10 (0 to 125); and i10-index 4 (0 to 245). Significant inter bibliometric index correlations were observed (Table 1).

Analysis by surgeon status and construct validity

Table 2 demonstrates median bibliometric indices of consultants holding Academic titles, which were significantly higher than University Hospital and DGH non-academics. On excluding Academic surgeons, the h-index, r-index and i10-index of University Hospital vs. DGH Surgeons were similar, 5.5 vs. 5 (p=0.06), 5.2 vs. 4.8 (p=0.235) and 4 vs. 2 (p=0.056) respectively. Both m-index and g-index showed significant variance 0.394 vs. 0.294 (p=0.016), and 12 vs. 8 (p=0.011) respectively. Bibliometric indices related to subspecialty are shown in table 3. With regard to gender, no significant differences in the distribution were observed for the h-index (p=0.081), m-index (p=0.528), g-index (p=0.306) or i10-index (p=0.275), but with regard to the r-index the variation was significant (p=0.034, table 4).

Influence of adapted bibliometrics on consultant rank

To evaluate the influence of adapted bibliometric variance on consultant rank, the cohort was ranked by the h-index to determine the upper quartile, and the variance in rank calculated for each of the adapted indices. Significant variance was observed when consultant surgeons were ranked by each individual bibliometric measure; the median shift in ranking position observed were: r-index -2 (range -46 to +23); m-index -6.5 (-53 to +22); g-index -0.5 (-24 to +13); and i10-index 0 (-8 to +11) respectively ($p<0.001$).

Altmetric Scores

The consultant cohort published 1,064 articles between 2011 and the time of writing; the median AS was 0 (0 - 508), reflecting the fact that 51.5% (n=548) of these articles had an AS of zero. The median AS-index was 1 (0-10). Median AS did not correlate with h-index (correlation coefficient 0.106, p=0.197). Bivariate analysis showed strong correlations between h-index and AS-index (correlation coefficient 0.390, $p<0.001$). Strong correlations were also found between AS-index and all other bibliometrics (Table 1). The only significant correlations found with regards to the median AS were with the m-index and the AS-index (table 1). There was no significant variation in the median AS when academic consultants were compared with University hospital and DGH consultants (3 vs. 1 vs. 0, p=0.500).

Discussion

This represents the primary UK study to report the relative academic output of a cohort of general surgeons measured with the full raft of contemporary bibliometrics. The salient findings were that 87.4% of surgeons (n=132) had published at least 3 articles (median 13). Median h-index, r-index, m-index, g-index, i10-index, AS and AS-index were 5, 5.2, 0.33, 10, 4, 0 and 1 respectively, with strong inter-variable correlation. Surgeons holding academic titles, whether substantive or honorary, had two-fold better indices than their NHS only contemporaries, which is in keeping with reports from other specialties allied with surgery.[5] Even so, significant variance was observed when consultant surgeons were ranked by individual alternative bibliometric measures; with the adapted measures almost universally having a relative negative effect on ranking position, ranging from -0.5 (g-index), to -6.5 (m-index). All of the null hypotheses were therefore rejected.

Periodical journals have represented the chief voice publicising science for over three-and-a-half centuries, with established conventions of integrity including independent peer review, yet have come under intensifying scrutiny. The root of the issue is that journal publication now frequently plays a role as a surrogate marker of research skill, and therefore career potential.[22] Analysing publication productivity is a hot topic of debate and determines financial streams for academic institutions by means of the RAE process. Yet there seems no really good way to do it; clearly a simple point-statistic isn't ideal. The questions to be faced are first; how to rate a very highly cited article with 100 authors? If it receives 1,000 citations, is it realistic to state that all 100 authors can claim 1,000 citations? Second, an author's age will influence their citation rate, and should citation numbers be divided by years? Third, should the first or corresponding author receive the same credit as one buried in the middle? Fourth, the balance between quantity and quality; high numbers of articles suggest the ability to get work done, but science is influenced by key, insightful publications. Therefore, it seems most sensible to summarise citation data in as many ways as possible. At the extreme, as many as 13-point statistics are possible: total citations, median citations per article, median citations / number of authors, mean citations per paper, h-index, g-index, m-index, 1st author h-index, last author h-index, 1st or last author h-index, 1st or 2nd author h-index, i10-index, and recent i10-index. Arguably, a combined algorithm may be more equitable, particularly when faced with the assessment of applicants for academic biomedical positions.

Since the dawn of the new millennium, evaluation and bibliometrics have wreaked havoc in the arena of higher education and scientific research. Governments seek to evaluate everything using quantitative indicators, which have led to the amplification of indicators of excellence and quality, terminology often used with little concern for exact meaning, and pride of place has gone to bibliometrics. The distribution of publications, citations, and research grants follow, in large measure, what is known as Pareto distributions, power laws of the form aX^{-y}, of which Lotka's law is a particular case of such skewed distributions with y=2. These distributions tend to follow a 20/80 rule of thumb i.e. about 20% of authors are responsible for 80% of citations. Publications are less concentrated, with 20% of authors producing only 60% of articles.[23] These trends are in keeping with the "Matthew effect"
"For whosoever hath, to him shall be given, and he shall have more abundance: but whosoever hath not, from him shall be taken away even that he hath." [24]
Arguably, an article's real quality is best measured by the number of times it is cited elsewhere (not self-citation), but building citations takes time, and other faster valuations are welcome, which has lead to the emergence of alternative metrics, now termed ‘‘altimetry’’. Altimetry expands the concept of citation beyond references in scientific articles to include for example, article electronic downloads, incorporation in clinical guidance, or inclusion in course curricula. The London based Altmetric.com , one of the original companies in this field, has since 2011, tracked mentions of published articles in sources ranging from social media, Wikipedia, to government policy.[19,20,21]
A number of potential limitations and confounding factors are inherent to studies of this type and have been highlighted previously. [5] Obtaining precise bibliographic metrics is dependent on accurately identifying authors and affiliations, but

researchers often change or use different initials and affiliations risking skewing results. This potential effect was minimized in this study by using of two reviewers searching independently, with triangulation to exclude anomalies. Only one UK Deanery and geographical region was studied; a UK wide approach should produce more powerful results. A bigger cohort of surgeons holding academic titles, would allow categorization according to academic seniority, allowing a more precise measure of validity. Moreover, honorary University titles may be awarded for a number of professional contributions including, medical education rather than research, which makes conclusions regarding construct validity potentially flawed.[7,25] Variability in database citation counts can also result in h-indices' contrariety, and several reports exist in this regard.[5,7,9,26] Only one on-line publication database was used (Scopus) and clearly it is possible that other databases may have resulted in different bibliometric calculations. However, a previous recent study reported that there was a good degree of reliability between Scopus and Web of Science, the other mainstream database, with Scopus, on average, identifying a greater number of both publications and citations.[5]

In conclusion, the findings of this study demonstrate that six novel bibliometric indices provide robust, though variable, measures of author academic output and profile, in addition to the ubiquitous h-index.[10,27] In the context of contemporary surgical training, all of these metrics may be used to identify opportunities to engage in scholarly activities likely to be rewarded with academic recognition and scientific publication. With the most recent iteration of the UK Joint Committee on General Surgical Training curriculum (2013) setting a target of 3 peer-reviewed publications by completion of training, statistics such as this are valuable in allowing health education directors and providers to develop informed and tailored rotational programs so that the non-operative technical skills demanded for demonstration of research competence and completion of training are achieved.[1]

Table 1: Spearman’s correlations between bibliometric indices

	r-index	**m-index**	**g-index**	**I10-index**	**AS-index**	**AS median**
h-index	0.913	0.716	0.961	0.939	0.390	0.106
	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(0.197)
r-index		0.618	0.864	0.859	0.292	0.038
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(0.644)
m-index			0.686	0.654	0.634	0.358
			(<0.001)	(<0.001)	(<0.001)	(<0.001)
g-index				0.942	0.432	0.139
				(<0.001)	(<0.001)	(0.088)
I10-index					0.371	0.129
					(<0.001)	(0.115)
AS-index						0.682
						(<0.001)

Table 1: Values are correlation coefficients. p-values in parentheses.

Table 2: Bibliometrics related to surgeon status

Consultant (n = 151)	h-index	r-index	m-index	g-index	I10-index
Academic (n = 18)	12.5 (2-63)	10.0 (0.8-64.8)	0.662 (0.22-1.5)	22.5 (4-125)	14.5 (0-245)
University Hospital (48)	5.5 (0-24)	5.2 (0-19.2)	0.394 (0-1.18)	12.0 (0-45)	4.0 (0-40)
DGH (85)	5.0 (0-20)	4.8 (0-20.8)	0.294 (0-1.5)	8.0 (0-43)	2.0 (0-30)
p-value	<0.001	0.001	<0.001	<0.001	<0.001

Table 2: Values are medians. Ranges in parentheses.

Table 3: Analysis by subspeciality

Specialities	h-index	r-index	m-index	g-index	I10-index
Upper GI (n = 29)	5.0 (2-39)	5.6 (0.8-27.6)	0.462 (0.15-1.5)	9.0 (2-62)	2.0 (0-80)
Colorectal (n = 51)	5.0 (0-25)	5.2 (0-26.8)	0.323 (0-1.18)	8.0 (0-47)	3.0 (0-38)
Breast (n = 27)	5.0 (1-63)	4.8 (0.4-64.8)	0.318 (0.8-1.5)	10.0 (1-125)	3.0 (0-245)
Vascular (n = 26)	7.0 (0-20)	7.0 (0-20.4)	0.360 (0-1.15)	18.0 (0-43)	6.5 (0-29)
Transplant (n = 7)	5.0 (3-12)	4.4 (1.6-12.8)	0.421 (0.24-0.79)	10.0 (6-28)	4.0 (1-13)
HPB (n = 7)	4.0 (0-9)	5.2 (0-9.6)	0.308 (0-0.56)	12.0 (0-14)	4.0 (0-9)
Endocrine (n = 4)	8.5 (1-16)	8.2 (0.4-13.6)	0.602 (0.06-0.75)	15.5 (0-27)	9.0 (1-24)
p-value	0.624	0.324	0.234	0.675	0.585

Table 3: Values are medians. Ranges in parentheses.

Table 4: Bibliometrics related to gender

Gender	**h-index**	**r-index**	**m-index**	**g-index**	**I10-index**
Male (136)	5 (0-63)	5.6 (0-64.8)	0.357 (0-1.5)	10 (0-125)	4 (0-245)
Female (15)	3 (0-14)	2.8 (0-10)	0.286 (0-1.18)	8 (0-30)	3 (0-18)
p-value	0.081	0.034	0.528	0.306	0.275

Table 4: Values are medians. Ranges in parentheses.

References

1. OECD and SCImago Research Group (CSIC) (2016) *Compendium of Bibliometric Science Indicators*. Available at: http://oe.cd/scientometrics [Accessed 28/06/18]
2. Otlet P (1934), Traité De Documentation: Le Livre Sur Le Livre, Théorie Et Pratique, Editiones Mundaneum: Mons, Belgium
3. Pritchard A (1969) Statistical Bibliography or Bibliometrics? *Journal of Documentation.* 25(4): 348-349.
4. Hirsch JE (2005) An index to quantify an individual's scientific research output. *Proc Natl Acad Sci U.S.A.* 102(46): 16569-16572.
5. Abdelrahman T, Brown J, Wheat J *et al* (2016) Hirsch Index Value and Variability Related to General Surgery in a UK Deanery. *J Surg Educ*. 73(1): 111-115
6. Garner RM, Hirsch JA, Albuquerque FC *et al* (2018) Bibliometric indices: defining academic productivity and citation rates of researchers, departments and journals. *J NeuroInterv Surg.* 10: 102-106
7. Sharma B, Boet S, Grantcharov T *et al* (2013) The h-index outperforms other bibliometrics in the assessment of research performance in general surgery: a province-wide study. *Surgery* 153: 493-501
8. Svider PF, Pashkova AA, Choudhry Z *et al* (2013) Comparison of scholarly impact among surgical specialties: An examination of 2429 academic surgeons. *Laryngoscope* 123: 884-889
9. Bould MD, Boet S, Sharma B *et al* (2011) h-indices in a university department of anaesthesia: an evaluation of their feasibility, reliability, and validity as an assessment of academic performance. *Br J Anaesth* 106(3): 325-330

10. Hirsch J (2007) Does the h-index have predictive power? *Proc Natl Acad Sci U.S.A.* 104(49): 19193-19198
11. Sidropoulos A, Katsaros D, Manolopoulos Y (2007) Generalised Hirsh h-index for disclosing latent facts in citation networks. *Scientometrics.* 72: 253-80
12. Romanovsky AA (2012) Revised h index for biomedical research. *Cell Cycle* 11(22): 4118-4121
13. Baerlocher MO, Newton M, Gautam T*, et al* (2007) The Meaning of Author Order in Medical Research. *J Investig Med* 55: 174-180
14. McInerney J (2011) h-Index, M-Index and Google Citations. Available at: *http://mcinerneylab.com/research/h-index-m-index-and-google-citations/* [Accessed 02/07/2018].
15. Joint Committee on Surgical Training (2017/2018) *Guidelines for the award of a CCT in General Surgery.* Available at*:* *https://www.jcst.org/quality-assurance/certification-guidelines-and-checklists/* [Accessed 05/08/2018]
16. Egghe L (2006) Theory and practice of the g-index. *Scientometrics* 69: 131
17. Sarkar S, Seshadri D (2015) Citation Indices: Measuring the “Impact” of Published Work. *Indian J Psychol Med* 37(3): 376.
18. Kozak M, Bornmann L (2012) A New Family of Cumulative Indexes for Measuring Scientific Performance. *PLoS One* 7(10): e47679
19. Altmetric Support. *How is the Altmetric Attention Score calculated?* Available at: *https://help.altmetric.com/support/solutions/articles/6000060969-how-is-the-altmetric-attention-score-calculated-* [Accessed 30/08/18]
20. Altmetric Support. *When did Altmetric start tracking attention to each attention source?* Available at: *https://help.altmetric.com/support/solutions/articles/6000136884-when-did-*

altmetric-start-tracking-attention-to-each-attention-source- [Accessed 28/08/18]

21. Powell A, Bevan V, Brown C *et al* (2018) Altmetric Versus Bibliometric Perspective Regarding Publication Impact and Force. *World J Surg* 42(9): 2745-2756
22. The Economist Explains (2017) *The problem with scientific publication, and how to fix it*. Available at: https://www.economist.com/blogs/economist-explains/2017/03/economist- explains-23 [Accessed 02/07/2018]
23. Gingras Y (2016) *Bibliometrics and Research Evaluation Uses and Abuses.* MIT press: Massachusetts
24. Matthew 13:12, Holy Bible: King James Version
25. Fuller CD, Choi M, Thomas CR (2009) Bibliometric analysis of radiation oncology departmental scholarly publication productivity at domestic residency training institutions. *J Am Coll Radiol* 6(2): 112-118
26. Falagas ME, Pitsouni EI, Malietzis GA et al (2008) Comparison of PubMed, Scopus, Web of Science, and Google Scholar: Strengths and Weaknesses. FASEB Journal 22(2): 338-342
27. Lee J, Kraus KL, Couldwell WT (2009) Use of the h index in neurosurgery. Clinical article. J Neurosurg. 111(2): 387-392.