BJUI Comments
BJU INTERNATIONAL

INDICATORS FOR RESEARCH PERFORMANCE EVALUATION: AN OVERVIEW

Saied Froghi, Kamran Ahmed, Adam Finch*, John M. Fitzpatrick†, Mohammed Shamim Khan, and Prokar Dasgupta – MRC Centre for Transplantation, NIHR Biomedical Research Centre, King's Health Partners, King's College London and Department of Urology, Guy's Hospital, London, UK, *Solutions Consultant, Research Performance Analyst, Thomson Reuters, Australia, and †Mater Misericordiae Hospital and University College Dublin, Ireland

Accepted for publication 9 November 2011
S.F. and K.A. contributed equally to this manuscript.

Since the advent of Garfield's science citation index in the 1950s [1], bibliometrics have been formally used to assess academic productivity and performance [2]. Although the use of bibliometric studies predates Garfield, the need to assess mathematically and statistically the quantity and quality of published work did not arise until the 1960s. This was attributable to the exponential growth in scientific publications resulting from the culture of 'publish or perish' [3,4] that overwhelmed different scientific fields, and the need to identify top quality research. With the development of many performance indices over the years, bibliometrics are increasingly becoming a political tool among scientists and policymakers [5], to the extent that they can be used in important decisions regarding appointments [6], promotions and funding.

One of the important roles of bibliometrics is the objective assessment of the dissemination, impact and quality of articles published by a particular author, journal or institution [7,8]. Currently, quantity and quality assessment is multi-dimensional and has extended beyond the three mentioned domains to include many more, such as subject, country and region. The result is the identification of high performers within a given scientific field. In addition, the results can help a faculty in the submission of their manuscripts to a particular journal, assess the quality of an individual researcher or a group of researchers, measure research focus and identify its misallocation [8], and help in admissions into elite societies [9]. An example of this is the recent study by Moppett and Hardman [10], who were able to identify the top ranking anaesthesia departments in the UK using bibliometric indices. Such studies can have an impact on the way research is funded and policies are compiled.

Given their importance, a number of indicators have been developed that are used by databases, such as the Institute of Scientific Information (ISI) Web of Science (accessible via: http://www.isiknowledge.com) and Scopus (http://www.scopus.com/), to compile results on the performance and productivity of researchers. An example would be the total number of articles published by a particular author. This is a simple quantitative indicator, and would not assess the quality (performance). Other more complicated indices have been introduced to the scientific community. Such performance indicators include the h-index [11,12], g-index [12,13], the age-weighted citation ratio (AWCR) [13], impact factor (IF) [14] and many more, which are used to reduce error and increase accuracy of assessment [13,15]. These have been challenged, however, by critics who question the use and validity of citation indices as indicators of quality, utility or even impact [16,17].

The use of bibliometrics by several governments, as part of the assessment of research productivity [18,19], shows the emerging importance of bibliometrics in shaping the future of academia; however, because of the lack of strong and extensive evidence on the accuracy and effectiveness of such indicators, their incorporation into the assessment of research has been slow. The problem that one faces with bibliometrics is that there is no particular index that is superior to the others; no particular index can be used alone as a primary indicator for quality and allocation of funds. The fact that each index has prejudicial peculiarities poses a challenge in deciding which specific one to use.

Perhaps the index that has attracted more attention than other citation-based metrics is the h-index, originally proposed by Hirsch [11] in 2005. This index has the ability to assess the ranks of a researcher's achievements objectively [12]. The h-index was considered to be a more reliable indicator than the total number of publications of an author. Not only does it account for the number of publications of a particular scientist, but it also assesses their impact on the scientist's peers [12]. This new method of assessment drew much attention among leading journals, such as *Nature* and *Science*, who realised the importance and potential of this new index [20]. It was not long before these journals started using the h-index to assess an individual or department's research achievements. Despite the interest in the h-index, it has several flaws, including, most notably, the fact that it depends on the age of the faculty or the length of an author's scientific career, hence, a higher h-index would be associated with older faculty members. Furthermore, it produces variable results from one discipline to another. Other more accurate indices soon followed the h-index. These were introduced to overcome the disadvantages of the h-index either by normalizing the results through reduction or by removal of

TABLE 1 Advantages and disadvantages of various bibliometric indicators

Bibliometric indicator	Definition	Advantage	Disadvantage	Reference
Crown Indicator	Developed by Centre for Science and Technology Studies (CWTS) at Leiden University. Average number of received citations (from a researcher or a research group) divided by the average number that could be expected for publication of the same type published in journals within same field.	• Allows comparison of researchers in different fields • Controls for citation rate in research field, document type and publication year; thus overcoming the limitations of IF	• Not readily available • Does not take into account the fact that articles from one field are published in journals of different field • Only allows comparison of equal-sized research groups	[8,21–24]
h-index	Proposed by Hirsch [11]: 'A scientist has index *h*, if *h* of his/her (N) papers have at least h citations each, and the other (N-h) papers have no more than h citations each'	• Readily available • Insensitive to extremely rare or frequently cited articles • Allows comparison of faculty of different ranks	• Depends on the age of the researcher/scientific career • Variable from one discipline to another • Does not take into account the position in the author list • Sensitive to homonym conflicts • Insensitive to highly cited work	[25–27]
Eigenfactor Score	A journal indicator that is 'an estimate of the percentage of time that library users spend with that journal'	• Easily available	• Does not take into account the scientific value of a journal	[8,28–30]
Article Influence Score	Measures the average influence, per article, of the papers in a journal; provides a standardized Eigenfactor score	• Reduces or removes large differences between fields evident in IF	• Dependent on the number of articles published	[30–33]
g-index	The highest number *g* of papers that in total received *g*2 or more citations	• Takes into account the citations that are ignored by h-index • Avoids subsequent counting of top cited h papers	• Puts more weight on highly cited papers	[12,13,34,35]
AWCR	A measure of average number of citations for an entire body of work, adjusted for the age of each individual paper.	• Actual number of citations are taken into account • Makes use of age of publication • Can be used with h-index to complement its accuracy	• Does not apply to source items where age has no meaning	[36,37]
m quotient	$m = h$/year where h = h-index and yr = number of years since publishing the first paper	• Allows h-index to compare faculties of different rank	• Insensitive to highly cited work • An unstable index for junior researcher as it takes into account the year of publication; thus large changes in *m* quotient can result from small changes in h-index	[12,13,36,38]

the difference incurred by the errors of the h-index.

In addition to author indicators, variables such as the frequency of citation and number of publications are used in calculating the IF of a particular journal (Table 1) [8,11–13,21–38]. Information yielded from this area of bibliometrics can help authors in selecting a particular journal for publication. Also, using parameters set by the ISI and Journal Citation Reports (JCR), one can attain information about the ranks of journals publishing in a particular speciality, in addition to assessing the academic profile of that particular speciality.

To date, there has been no specific use of such indices in evaluating the productivity and performance of different urology faculties and departments. Neither have they been used specifically in identifying the right candidates for their matched programme or for allocation of resources to the right faculty and department. Surely parameters such as number of papers published by a particular author are too simplistic a means of assessment of urology clinicians, a viewpoint that is shared by many academics of different disciplines. This necessitates the identification of new and improved methods of assessment. To assess its relevance, Benway *et al.* [39] explored the predictive power of the h-index in academic urology departments of the USA. Despite the limitations of the study, they were able to identify the h-index as a valuable measure for appraising the contributions of academic urologists. In addition, their study was able to identify top urology programmes in the USA based on rankings of departmental h-indices.

Even though a recent study revealed that the field of urology has satisfactory measures created by the JCR in comparison with various fields of medicine [40], there are essential gaps to be filled when considering bibliometrics in assessment of urologists. Academic urology lacks specific author and journal indicators that would allow the identification of high performance clinicians based on realistic statistical parameters of quantity and quality. Similarly, the identification of high quality urology journals for publication would help in the evidence-based assessment process of different faculty members and departments.

The design and development of databases that allow for use of bibliometric indices are examples of how technology aids assessment of ranks and allocation of funds. Thus, allowing for an objective assessment of different scientific fields and aiding its progress through identification of high performers and allocation of resources to the world of academia. The evolution of peer-review appraisal systems in assessing an individual researcher for promotions or for admission into elite societies is not yet complete. It is only with the development of more accurate bibliometric indicators that assessment and evaluation of research is becoming more objective and robust. Despite this, the production of bona fide results with high accuracy is still a challenge because of database variability and problems associated with normalizing self-citations.

With the expansion of bibliometrics, it is believed a better system of appraisal of urology clinicians is at hand, not only to assess the individual candidate's suitability for a position, but also to assess the merits of researchers and departments. Patient safety can only be secured by ensuring that evidence-based practice originates from quality researchers and institutions. Overall, we believe the use of bibliometrics is beneficial since it provides a statistical framework for assessment of research quality; however, to obtain optimum results, there is a need to modify and design new indices of assessment. Indeed, further studies are required to appraise the urology faculty and department's productivity and performance.

ACKNOWLEDGEMENTS

Prokar Dasgupta acknowledges financial support from the Department of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy's & St Thomas' NHS Foundation Trust in partnership with King's College London and King's College Hospital NHS Foundation Trust. He also acknowledges the support of the MRC Centre for Transplantation.

CONFLICT OF INTEREST

None declared.

REFERENCES

1 **Garfield E.** Citation indexes for science; a new dimension in documentation through association of ideas. *Science* 1955; **122**: 108–11
2 **Sen BK, Shailendra K.** Evaluation output of recent scientific research by a bibliometric method. *Scientometrics* 1992; **23**: 31–46
3 **Refinetti R.** Publish and flourish. *Science* 2011; **331**: 29
4 **De Rond M, Miller AN.** Publish or Perish. *J Manage Inquiry* 2005; **14**: 321–9
5 **Traynor M.** Bibliometrics as politics: the case of emerging disciplines. *Int Nurs Rev* 2011; **58**: 26–7
6 **Ball P.** Index aims for fair ranking of scientists. *Nature* 2005; **436**: 900
7 **Narin F, Olivastro D, Stevens KA.** Bibliometrics/Theory, Practice and Problems. *Eval Rev* 1994; **18**: 65–76
8 **Durieux V, Gevenois PA.** Bibliometric indicators: quality measurements of scientific publication. *Radiology* 2010; **255**: 342–51
9 **Alberts B, Fulton KR.** Election to the National Academy of Sciences: pathways to membership. *Proc Natl Acad Sci U S A* 2005; **102**: 7405–6
10 **Moppett IK, Hardman JG.** Bibliometrics of anaesthesia researchers in the UK. *Br J Anaesth* 2011; **107**: 351–6
11 **Hirsch JE.** An index to quantify an individual's scientific research output. *Proc Natl Acad Sci US A* 2005; **102**: 16569–72
12 **Bornmann L, Daniel H-D.** Are there better indices for evaluation purposes than the h Index? A comparison of nine different variants of the h Index using data from biomedicine. *J Am Soc Inf Sci* 2008; **59**: 830–7
13 **Thompson DF, Callen EC, Nahata MC.** New indices in scholarship assessment. *Am J Pharm Educ* 2009; **73**: 111
14 **Patel VM, Ashrafian H, Ahmed K *et al*.** How has healthcare research performance been assessed? A systematic review. *J R Soc Med* 2011; **104**: 251–61
15 **Anderson TR, Hankin RKS, Killworth PD.** Beyond the Durfee square: enhancing the h-index to score total publication output. *Scientometrics* 2008; **76**: 577–88
16 **Cronin B.** Bibliometrics and beyond: some thoughts on web-based citation analysis. *J Inform Sci* 2001; **27**: 1–7
17 **Bornmann L *et al*.** Citation counts for research evaluation: standards of good practice for analyzing bibliometric data and presenting and interpreting results. *Ethics Sci Environ Politics* 2008; **8**: 93–102
18 **Phelan TJ.** Bibliometrics and the evaluation of Australian sociology. *J Sociol* 2000; **36**: 345–63
19 **Giles J.** Thumbs up for fresh formula to gauge university funding. *Nature* 2004; **427**: 667–667
20 **Ball P.** Achievement index climbs the ranks. *Nature* 2007; **448**: 737
21 **Schulman CC.** What you have always wanted to know about the impact factor and did not dare to ask. *Eur Urol* 2005; **48**: 179–81
22 **Lundberg J.** Lifting the crown – citation z-score. *J Inform* 2007; **1**: 145–54
23 **Lundberg J *et al*.** Is it better or just the same? Article identification strategies impact bibliometric assessments. *Scientometrics* 2006; **66**: 183–97
24 **Lundberg J *et al*.** Measuring the validity of early health technology assessment: bibliometrics as a tool to indicate its scientific basis. *Int J Technol Assess Health Care* 2008; **24**: 70–5
25 **Devos P.** Research and bibliometrics: a long history. *Clin Res Hepatol Gastroenterol* 2011; **35**: 336–7
26 **Vinkler P.** The π-index: a new indicator

for assessing scientific impact. *J Inform Sci* 2009; **35**: 602–12
27 **Lazaridis T.** Ranking university departments using the mean h-index. *Scientometrics* 2010; **82**: 211
28 **Bergstrom C.** eigenFACTOR.org. 2009. Available at: http://www.eigenfactor.org/methods.htm. Accessed 9 August 2011
29 **Bergstrom CT, West JD, Wiseman MA.** The Eigenfactor™ Metrics. *J Neurosci* 2008; **28**: 11433–4
30 **Ortner HM.** The impact factor and other performance measures – much used with little knowledge about. *Int J Refract Metals Hard Mater* 2010; **28**: 559–66
31 **Bergstrom CT, West JD.** Assessing citations with the Eigenfactor™ Metrics. *Neurology* 2008; **71**: 1850–1
32 **Arendt J.** Are article influence scores comparable across scientific fields? *Issues Sci Technol Lib* 2010; **60**: 1–10
33 **Chia-Lin C, Michael M, Les O.** *Article Influence Score = 5YIF Divided by 2.* University of Canterbury, Department of Economics and Finance, 2010. Available at: http://repub.eur.nl/res/pub/20148/EI2010-43.pdf
34 **Egghe L.** Theory and practise of the g-index. *Scientometrics* 2006; **69**: 131–52
35 **Woeginger GJ.** An axiomatic analysis of Egghe's g-index. *J Informetr* 2008; **2**: 364–8
36 **Harzing A-W.** Publish or Perish. 2011. Available at: http://www.harzing.com/pop.htm. Accessed 9 August 2011
37 **Jin B *et al*.** The R – and AR -indices: complementing the h -index. *Chin Phys Lett* 2007; **52**: 855–63
38 **Alonso S *et al*.** h-Index: a review focused in its variants, computation and standardization for different scientific fields. *J Informetr* 2009; **3**: 273–89
39 **Benway BM *et al*.** Does citation analysis reveal association between h-index and academic rank in urology? *Urology* 2009; **74**: 30–3
40 **Labanaris AP *et al*.** Impact factors in urology. How well do our journals stack up against other medical and surgical journals? *Urol Int* 2007; **78**: 299–304

Correspondence: Kamran Ahmed, Department of Urology, Guy's Hospital, St Thomas Street, London SE1 9RT, UK.
e-mail: k.ahmed@imperial.ac.uk

Abbreviations: **ISI**, Institute of Scientific Information; **AWCR**, age-weighted citation ratio; **IF**, impact factor; **JCR**, Journal Citation Reports.