10-2013

Best Evidence in Emergency Medicine (BEEM) rater scores correlate with publications' future citations

Christopher R. Carpenter
Washington University School of Medicine in St. Louis

Cathy C. Sarli
Washington University School of Medicine in St. Louis

Susan A. Fowler
Washington University School of Medicine in St. Louis

Kulamakan Kulasegaram
McMaster University

Teresa Vallera
McMaster University School of Medicine

See next page for additional authors

Authors
Christopher R. Carpenter, Cathy C. Sarli, Susan A. Fowler, Kulamakan Kulasegaram, Teresa Vallera, Pierre Lapaine, Grant Schalet, and Andrew Worster

BEEM Rater Scores Correlate with Publications' Future Citations. Christopher R. Carpenter, Cathy C. Sarli, Susan A. Fowler, Kulamakan Kulasegaram, Teresa Vallera, Pierre Lapaine, Grant Schalet, Andrew Worster. Academic Emergency Medicine. Volume 20, Issue 10, pages 1004–1012, October 2013. © 2013 by the Society for Academic Emergency Medicine.

BEEM Rater Scores Correlate with Publications' Future Citations

Christopher R. Carpenter, Cathy C. Sarli, Susan A. Fowler, Kulamakan Kulasegaram, Teresa Vallera, Pierre Lapaine, Grant Schalet, Andrew Worster

Word Count: 3233

Reprints are not available from the authors.

Corresponding author contact information:

Christopher R. Carpenter, MD, MS
Campus Box 8072
660 S. Euclid Avenue
St. Louis MO 63110
Telephone: (314) 362-7979
FAX: (314) 362-0419
E-mail: carpenterc@wusm.wustl.edu

Category: Special contributions/original research

Key words: bibliometrics, knowledge translation, evidence-based emergency medicine, validation

BACKGROUND

Evidence-based clinical care is informed by relevant studies with statistically valid methodology and pertinent conclusions.1 Knowledge translation, otherwise known as implementation science is the methodology of delivering minimally biased, maximally effective clinical research-based care to the patient bedside expeditiously.2 The first two barriers to knowledge translation are awareness and acceptance of the knowledge imparted by relevant studies.3 Unfortunately, awareness alone is a formidable barrier to efficient knowledge translation since the volume of biomedical publications continues to expand at an increasing pace. In fiscal year 2012 alone, 760,903 citations from over 5,600 journals were added to MEDLINE/PubMed®, representing a five percent increase from 2011.4 According to Ulrich's Global Serials Directory as of February 2013, there were 14,212 active, scholarly peer-reviewed, English-language journals in the category of Medicine and Health. The term "emergency" is included in the titles of 83 of these journals.5 Many of the manuscripts contained in these professional journals are not clinically relevant. This reality adds to the barriers clinicians face in keeping current with the literature.6,7 Within Annals of Emergency Medicine, a leading emergency medicine (EM) journal, the "number needed to read" to identify one clinically relevant manuscript was estimated to be 26.8 To compound the problem of awareness, the specialty of EM crosses many other specialties so relevant publications often appear in non-EM journals that are not read by emergency physicians (EPs).9

Ideally, busy clinicians would have access to all relevant information pertinent to a particular question, but a one-stop resource does not yet exist.10,11 A variety of costly online sources are available, but these are often outdated with variable quality.12 One solution for clinicians to maintain a contemporary awareness of an ever-evolving research landscape exists in a variety of "secondary peer-reviewed" proprietary products.13-15 However, the selection processes for secondary peer-reviewed sources often neglect to include the clinician perspective, which is essential since the individual clinician-patient interaction constitutes the "end-user" of the research product. In response, research and educational leaders advocate a focused and deliberate effort to yield consistent, patient-outcome oriented science into the overly expensive and inefficient mainstream healthcare system.16 Another potential remedy for information overload proposed almost a decade ago by the American Board of Emergency Medicine (ABEM) was the Lifelong Learning and Self Assessment (LLSA) series, although the overall quality of these manuscripts has been challenged.17,18 Additional challenges to awareness and acceptance include physician biases and insufficient training for devising electronic search strategies to find pertinent evidence, to attain proficiency in critical appraisal, as well as an underdeveloped graduate medical education Evidence Based Medicine (EBM) curriculum.19-22 Finally, the path to enlightenment for all stakeholders is often shadowed by the belief that much of the published research is conceptually misleading, inapplicably complex, or wrongly interpreted.23

To overcome these barriers, EP researchers and educators at McMaster University and Washington University developed "Best Evidence in Emergency Medicine" (BEEM) to search for, identify, appraise, and translate potentially practice-changing studies for EPs. The criteria used to select articles for review by BEEM raters has been previously described.24 The BEEM rater scale (Table 1) is a medical literature-rating tool for EPs to collectively evaluate the relative clinical relevance of EM-related studies found in any medical journal based only on an article's title and conclusion. The scale is a 7-item instrument that serves as a clinical relevance filter to identify those studies with the greatest potential to affect EM practice. The BEEM ratings indicate those studies identified by raters as having the highest clinical relevance. These studies are subsequently selected for critical appraisal and, if found methodologically sound, are promoted as the best evidence in EM. A prospective randomized study demonstrated the BEEM rater scale to be a highly reliable single-question tool for a minimum of 12 EPs to collectively rate the relative clinical EM relevance of any published study from a variety of medical journals based on the study's title and conclusion alone.24

Because there is no other existing validated measure of clinical relevance, there is no reference standard by which to validate the BEEM rater scale. Ballard et al. proposed a model for measuring clinical relevance that contains four parameters using subjective assessment and circulation figures.25 However, this model is intended for all medical specialties with the caveat that the model requires further investigation. As an alternative, we hypothesized that the BEEM rater scale, known to be predictive of physician-perceived clinical relevance, also correlates highly with future citation of manuscripts chosen for review by BEEM reviewers. Future citations would thus serve as a bibliometric-based construct of clinical relevance of articles chosen for review by BEEM reviewers, providing an additional type of evidence of the clinical relevance of the articles selected for BEEM review. To that end, we sought to identify a bibliometric-based construct of clinical relevance specific to EM by which to correlate and indirectly validate the BEEM rater scale.

METHODS

Study Design and Population

The study population and BEEM rater survey methods have been previously described.24 From January 2007 to April 2012, BEEM raters completed an online survey every 2-3 months. BEEM raters consisted of approximately 200 practicing, English-speaking EPs from the United States, Canada, Great Britain, and Australia. We recruited these participants from BEEM continuing medical education courses during the years 2005-2011, as well as through non-peer-reviewed publications, advertisements, professional conferences, and via volunteers who contacted BEEM directly. During three years of the BEEM rater survey period, we awarded one randomly selected participant a complimentary 1-year subscription to PEPID (LLC, Evanston IL) or free registration to a BEEM course with each survey. Participants were not paid or otherwise compensated for their efforts. The e-mails containing the Survey Monkey (http://www.surveymonkey.com) hyperlink to the BEEM rater survey were sent to the entire list of volunteers each month.

Survey Content and Administration

Each month every BEEM rater received an e-mail containing the title and conclusions of 10 to 20 recently published manuscripts (Figure). The BEEM raters were also provided a hyperlink to the PUBMED abstract for those wishing to review additional details of any study. Each BEEM rater was asked to rate every manuscript using the 7-item Likert scale BEEM rater scale. The manuscripts selected for evaluation by the BEEM raters were identified through the McMaster Health Information Research Unit based upon methodological quality as previously described.24 The McMaster University/Hamilton Health Sciences Research Ethics Board approved this study.

Outcome Measures

Lacking any previously validated criterion standard for "clinical relevance", the BEEM Research team collaborated with medical librarians from Washington University in St. Louis Becker Medical Library to derive an EM-specific bibliometric construct of clinical relevance with acceptable face and content validity. We based our criterion for validity assessment on a two-fold approach incorporating bibliometric indices and study characteristics. The use of bibliometric-based measures for validation purposes was based on the rationale that future performance of a publication based on document-level, author-level, or journal-level patterns could be associated with clinical relevance. We describe the bibliometric construct measures selected for analysis below.

The term "bibliometrics" was coined by Alan Pritchard in 1962 and refers to studies seeking to quantify the processes of written communication.26 Today bibliometrics refers to quantitative analyses of publication data using document, author, or source (e.g. journal) level data elements to uncover characteristics, patterns, and relationships to demonstrate productivity, quality, or impact. The usefulness of bibliometrics has been described across a variety of scientific settings.27-30 One level of bibliometric-based analysis that represents a document-level characteristic is citation counts, which examine how frequently one publication is cited by subsequent publications. The inherent assumption with citation analyses is that significant or quality publications garner higher citation counts.31 Citation analysis is recognized as an accepted bibliometric tool to assess the merits of individual publications.32,33

For author-level analysis, we selected two measures. One was the h index. Hirsch derived the h index using number of publications and citations to provide "an estimate of the importance, significance, and broad impact of a scientist's cumulative research contributions" and asserted that the h index may be a more robust indicator than total citation count to predict future individual scientific achievement. 34-36 First and last author h indices were captured as these authors are traditionally considered to be the primary contributors to the research and intellectual content of the manuscript. Because the h index does not factor in number of authors to a publication, we included an additional author-level analysis based on the number of authors and the number of unique institutions represented by each manuscript. Publications involving collaboration among multiple authors from different institutions and countries sometimes demonstrate increased citation rates.37,38

For the journal-level analysis, we selected the Journal Citation Reports (JCR) journal impact factor score for journals. This journal score was developed in 1955 as a journal selection tool for inclusion in the Science Citation Index, the precursor to Thomson Reuter's Web of Science (WoS) database. The JCR score has gradually evolved to serve as an impact or quality proxy for journals and published manuscripts, albeit unintentionally.39,40 Study characteristics were included in the bibliometric-construct because design, category, and study type can also affect the subsequent impact of a manuscript due to methodological rigor, as well as standing of the research on the EBM hierarchy of evidence.41-43

After contemplating these various bibliometric constructs of "clinical relevance" including Thomson Reuters JCR score, article citation counts from Thomson Reuters WoS, and the first and last authors' h indices from Elsevier SciVal Scopus, 25,44 and author composition characteristics, the authors agreed upon a consensus criterion standard for "clinical relevance" defined by the WoS total citation count. Since citation counts increase as time passes post-publication, a time-adjusted version of the WoS total citation count was computed by dividing the citation count by the number of years in publication. This quantity was labeled the citation rate and was our criterion standard.

Two investigators (PL, GS), each blinded to the BEEM rater scores, independently abstracted the bibliometric indices for the BEEM rated articles. Each of these investigators abstracted the bibliometric indices for approximately 330 articles with 10% of the total number of articles selected to overlap between the two reviewers in order to assess reliability of the data abstraction process. The overlapping articles were not randomly selected. Instead, these manuscripts represented a portion of consecutive manuscripts abstracted during the year 2009, specifically BEEM Rater articles 353 through 387 using our administrative numbering system. In addition to the WoS total citation count from date of publication to 2011, these two investigators also abstracted the number of authors including corporate and group authors, JCR score, as well as first and last authors' h indices. In addition to these bibliometric indices, these two investigators also noted the study type (diagnostic, prognostic, therapeutic) and study design (systematic review (SR), randomized controlled trial (RCT), or observational) and sample size for each BEEM rated manuscript using MEDLINE/PubMed®. Self-citations were not removed as part of the study.

Data Analysis

We assessed mean and median citation rates stratified by year of publication. The validity of the BEEM rater score for each article was assessed using negative binomial regression with the citation rate as the criterion standard while adjusting for additional independent variables including the JCR score, first and last authors' h indices, and total number of authors. Over-dispersion exists when variance exceeds the mean of a sample and is adjusted for using negative binomial regression.45 We constructed this model using a theoretical framework and estimated parameters through logistic functions yielding odds ratios.

In order to assess the construct validity of our criterion standard for "clinical relevance", three negative binomial regression models were assessed. The first model assessed BEEM rater scores against the citation rate only. The second model assessed impact factors followed by the BEEM rater score. The third model assessed impact factors, study type, study design, and BEEM Rater score. Model fit was assessed using Likelihood ratio chi-square deviance and Bayesian information criterion.46,47 Analyses were conducted using SPSS 20 (Chicago IL, IBM Corporation).

We calculated inter-rater agreement by calculating the proportion of duplicate articles that had the same bibliometric data extracted by the investigators. Inter-rater agreement was calculated by proportions instead of kappa or intra-class correlation coefficients as the latter are reliability calculations most appropriate for discriminating or differentiating a series of objects of measurement.48 BEEM Rater scores were correlated with citation rate using Spearman's Rho.49

RESULTS

BEEM rating physicians reviewed 605 unique articles over 5-years, yielding a mean and median BEEM rater score of 3.84 and 3.85, respectively. The BEEM Rater scores displayed a normal distribution. Therapeutic articles predominated and almost half were published in 2008 (Table 2). The WoS total citation counts were available for 594 manuscripts with a highly skewed distribution (mean 23, median 9, standard deviation 59). Over 75% of the distribution had citation counts less than the median with the range of citation counts extending from 0 to over 1000. Average citation counts with standard deviations for 2007, 2008, 2009, 2010, and 2011 were 8.64 (SD 20.6), 8.12 (SD 20.2), 7.00 (SD 17.9), 2.97 (SD 9.1), and 0.14 (SD 0.74), respectively. Data abstraction had high inter-rater agreement (98%).

Analysis with Spearman's rho showed that the BEEM rater score had a low but significant correlation with citation count ($r=0.144$, $p<0.0001$). Greater association was seen with number of authors as well as Scopus first and last authors' h indices, all of which had correlation coefficients greater than 0.35 ($p<0.00001$) against citation count. Lastly, the JCR score was correlated with citation count at 0.52 ($p<0.00001$). The correlations are summarized in Table 3.

In the first model assessing BEEM rater score alone against the citation rate, the BEEM rater score significantly predicted the citation rate with an odds ratio of 1.24 (95% CI 1.11-1.40, Wald chi square 13.2 (df 1), $p<0.0001$). In other words, for every one unit score increase on the BEEM rater scale, the odds of being cited increased by 1.24. The estimates of model fit were deviance 1510 and Bayesian Information Criterion (BIC) 8870. Deviance indicates the extent of a statistical model or equation fitting the data when using maximum likelihood procedures such as logistic and binomial regression. The BIC compares multiple different models for extent of fit based on the deviance as well as the complexity of the model with lower numbers indicating better fit. Just as linear regression r2 compares different models, deviance and BIC can be used to compare the validity and explanatory power of maximum likelihood derived regression equations.46,47

The second model assessed the first and last authors' h indices, JRC score, number of authors, and the BEEM rater score. In this model, BEEM rater score was no longer significant with an odds ratio of 1.11 (95% CI 0.98-1.24). All of the other variables were significant in this model (Table 4).

The third model assessed study type, number of authors, first- and last-authors' h indices, JCR score, and BEEM rater scores. The BEEM rater scores were not significant with an odds ratio of 1.11 (95% CI 0.99-1.25, $p=0.71$). The only study type variable in this model that was significant was RCT compared with observational studies (odds ratio 2.29, 95% CI 1.75-3.00) and RCT compared with SR's (odds ratio 2.54, 95% CI 2.09-3.10). However, this model had only marginally lower BIC (8006) suggesting that study design assessed parameters similar to that of the JCR score and h-indices.

DISCUSSION

Distinguishing practice-changing or practice-enhancing research from the large volume of clinical publications is a long-standing challenge for clinicians and educators. Published research that is neither practice-enhancing nor practice-changing is not inconsequential since these papers often form the preliminary building blocks upon which future medical management is based. In fact, the conceptual and philosophical concept of "best evidence" is debatable. However, in an era when clinical care is heterogeneous, increasingly expensive, and often not evidence-based, healthcare providers need pragmatic solutions to close the gap between what we know and what we do.50-52 The ideal "best evidence" instrument would be simple to understand, administered briefly, and accurately distinguish practice-enhancing research from other publications. This instrument could be utilized to prioritize manuscripts for secondary peer review in graduate medical education or in continuing medical education such as annual scientific assemblies and the ABEM LLSA.18

Although the BEEM rater score correlates with future citations for EM-relevant publications, it is not an independent predictor of citation rates. Specifically, the BEEM rater score is not independently associated with citation rate when adjusted for first-and last-authors' h-indices, JCR score, number of authors, and study type.

There are several potential explanations as to why the BEEM rater score is not an independent predictor of citation rates. One is over-simplicity. The BEEM rater instrument was derived to provide a brief overview of approximately 20 potentially practice-enhancing research publications every month by busy clinicians. The 7-item Likert scale may fail to identify complexities in the interpretation, appraisal for bias, or application of medical research publications from the perspective of clinicians. Although future medical informatics research should explore more complex scoring systems to encapsulate this complexity, the trade off is that these alternative scoring systems may become counterintuitive, unreliable, or unwieldy for non-research clinicians who are busy with daily patient care responsibilities. In addition, citation rate alone is an imperfect and indirect indicator of clinical relevance.53 For example, publishing guidelines like the Standards for Reporting of Diagnostic Accuracy criteria are cited in multiple journals whenever these journals publish a diagnostic manuscript.54 Citation count skewing is also noted by self-citation, as well as citations for improper or incorrect findings.53,55 The distribution model of the journal is another variable since open access journals generally have higher citation counts.56,57 Citation counts also differ between different services such as Web of Science, Scopus, and Google Scholar.58 Another explanation is that the sampling of manuscripts rated was not representative of the totality of EM publications routinely available to clinicians. This is unlikely since our filtering process used the McMaster Health Information Research Unit, which has been relied upon as a source of secondary peer-reviewed data for over a decade.59,60

LIMITATIONS

Our choice of bibliometric indices upon which to model clinical relevance is imperfect. The JCR score is a metric applied to journals that have become surrogates for the publication impact of individual researchers in assessing merit for promotion.53,61 The JCR score may not even be an adequate measure of the relative strength of a publication, so it is debatable whether it is reasonable to include as a measure of "clinical relevance".25,62 In addition, citation rates for a publication are often easily manipulated by self-citations and subject to gratuitous citations among colleagues, neither of which reflect assessment of methodological rigor or overall study quality.63 Nonetheless, no validated gold standard for "clinical relevance" currently exists and our review of the medical and librarian literature identified no prior attempts to define one beyond theory.64 An alternative interpretation of our results is that the statistically significant correlations serve as a validation of the BEEM filtering process by which we identify the articles to rate rather than as an independent correlation of the BEEM rater score.

One approach to differentiate the correlation of the BEEM filtering process versus the BEEM rater score with bibliometric indices would be to randomly assign manuscripts to raters that did and did not pass through the initial BEEM filtering process.

In addition, our findings might be at risk for a selection bias in either the BEEM rater physicians who served as the subjects for this project or the pre-filtered manuscripts identified by the McMaster Health Information Research Unit. Our sampling was limited to English-speaking emergency physicians and manuscripts, although we have no compelling reason to believe that the BEEM rater scale would yield different operating characteristics amongst non-English speaking physicians. Finally, the sample of manuscripts for the reliability assessment was not randomly selected and therefore may not be representative of the entire collection of articles abstracted in this study which could bias our reliability assessment.

CONCLUSION

The BEEM rater scale correlates with future citations, but is not an independent predictor of citation rates. Future projects should assess this instrument, and more complex instruments that incorporate various forms of research critical appraisal, against alternative constructs of "clinical relevance".

REFERENCES

1. Lauritsen J, Moller AM. Publications in anesthesia journals: quality and clinical relevance. Anesth Analg 2004;99:1486-9.

2. Lang ES, Wyer PC, Haynes RB. Knowledge translation: closing the evidence-to-practice gap. Ann Emerg Med 2007;49:355-63.

3. Diner BM, Carpenter CR, O'Connell T, et al. Graduate medical education and Knowledge Translation: Role models, information pipelines, and practice change thresholds. Acad Emerg Med 2007;14:1008-14.

4. Key MEDLINE® Indicators. U.S. National Library of Medicine National Institutes of Health, 2012. (Accessed January 22, 2013, at http://www.nlm.nih.gov/bsd/bsd_key.html.)

5. Ulrich's Web Global Serials Directory [Internet]. ProQuest, 2013. (Accessed February 24, 2013, at http://ulrichsweb.com.)

6. Choudhry NK, Fletcher RH, Soumerai SB. Systematic review: the relationship between clinical experience and quality of health care. Ann Intern Med 2005;142:260-73.

7. Grimshaw JM, Eccles MP. Is evidence-based implementation of evidence-based care possible? Med J Aust 2004;180:S50-S1.

8. McKibbon KA, Wilczynski NL, Haynes RB. What do evidence-based secondary journals tell us about the publication of clinically important articles in primary healthcare journals? BMC Med 2004;2:33.

9. Reynolds JC, Menegazzi JJ, Yealy DM. Emergency medicine journal impact factor and change compared to other medical and surgical specialties. Acad Emerg Med 2012;19:1248-54.

10. Graber MA, Randles BD, Ely JW, Monnahan J. Answering clinical questions in the ED. Am J Emerg Med 2008;26:144-7.

11. Worster A, Haynes RB. How do I find a point-of-care answer to my clinical question? CJEM 2012;14:31-5.

12. Jeffery R, Navarro T, Lokker C, Haynes RB, Wilczynski NL, Farjou G. How current are leading evidence-based medical textbooks? An analytic survey of four online textbooks. J Med Internet Res 2012;14:e175.

13. Emergency Medical Abstracts. The Center for Medical Education, 2012. (Accessed March 1, 2013, at http://ccme.org/ema/.)

14. Emergency Medicine Journal Watch. Massachusetts Medical Society, 2013. (Accessed March 1, 2013, at http://emergency-medicine.jwatch.org/.)

15. Emergency Medicine. Oakstone Publishing, LLC, 2013. (Accessed March 1, 2013, at http://www.practicalreviews.com/Main/Home.aspx.)

16. Bonham AC, Rich EC, Davis DA, Longnecker DE, Heinig SJ. Commentary: Putting evidence to work: an expanded research agenda for academic medicine in the era of health care reform. Acad Med 2010;85:1551-3.

17. Miller SH. American Board of Medical Specialties and repositioning for excellence in lifelong learning: maintenance of certification. J Contin Educ Health Prof 2005;25:151-6.

18. Upadhye S, Carpenter CR, Diner B. Methodological quality of the American Board of Emergency Medicine Lifelong self-assessment reading lists. Acad Emerg Med 2008;15:S131.

19. Kuhn GJ, Wyer PC, Cordell WH, Rowe BH. A survey to determine the prevalence and characteristics of training in evidence-based medicine in emergency medicine residency programs. J Emerg Med 2005;28:353-9.

20. Carpenter CR, Kane BG, Carter M, Lucas R, Wilbur LG, Graffeo CS. Incorporating evidence-based medicine into resident education: a CORD survey of faculty and resident expectations. Acad Emerg Med 2010;17:S54-S61.

21. Krause R, Moscati R, Halpern S, Schwartz DG, Abbas J. Can emergency medicine residents reliably use the internet to answer clinical questions? Western J Emerg Med 2011;12:442-7.

22. Carpenter CR. Teaching Lifelong Learning Skills: Journal Club and Beyond. In: Rogers RL, Mattu A, Winter M, Martinez J, eds. Practical Teaching in Emergency Medicine 2nd ed. Oxford UK: Wiley-Blackwell; 2012:151-62.

23. Ioannidis JP. Why most published research findings are false. PLoS Med 2005;2:e124.

24. Worster A, Kulasegaram K, Carpenter CR, et al. Consensus conference follow-up: inter-rater reliability assessment of the Best Evidence in Emergency Medicine (BEEM) rater scale, a medical literature rating tool for emergency physicians. Acad Emerg Med 2011;18:1193-200.

25. Ballard J, Graf C, Young C. Measuring clinical relevance and impact in journal publishing: a proposed model and publisher's perspective. Clin Ther 2011;33:B11-B5.

26. Pritchard A. Statistical biography or bibliometrics? . J Documentation 1969;25:349.

27. White HD, McCain KW. The structure of medical informatics journal literature. J Am Med Inform Assoc 1998;5:448-66.

28. Smith DR. A 30-year citation analysis of bibliometric trends at the Archives Of Environmental Health, 1975-2004. Arch Environ Occup Health 2009;64:43-54.

29. Durieux V, Genenois P. Bibliometric indicators: quality measurements of scientific publication. Radiology 2010;255:343-51.

30. Spreckelsen C, Deserno TM, Spitzer K. Visibility of medical informatics regarding bibliometric indices and databases. BMC Med Inform Decis Mak 2011;11:24.

31. Lawani SM. Citation analysis and the quality of scientific productivity. Bioscience 1977;27:26-31.

32. MacRoberts MH, MacRoberts BR. Problems of citation analysis. Scientometrics 1996;36:435-44.

33. Phelan TJ. A compendium of issues for citation analysis. Scientometrics 1996;45:117-36.

34. Hirsch JE. An index to quantify an individual's scientific research output. Proc Natl Acad Sci U S A 2005;102:16569-72.

35. Hirsch JE. Does the H index have predictive power? Proc Natl Acad Sci U S A 2007;104:19193-8.

36. Tscharntke T, Hochberg ME, Rand RA, Resh VH, Krauss J. Author sequence and credit for contributions in multiauthored publication. PLoS Biol 2007;5:e18.

37. Katz J, Hicks D. How much is collaboration worth? A calibrated bibliometric model. Scientometrics 1997;40:541-54.

38. Figg WD, Dunn L, Liewehr DJ, et al. Scientific collaboration results in higher citation rates of published articles. Pharmacotherapy 2006;26:759-67.

39. Garfield E. Citation indexes for science; a new dimension in documentation through association of ideas. Science 1955;122:108-11.

40. Garfield E. The agony and the ecstasy: the history and meaning of the impact factor, 2005. In: International Congress on Peer Review and Biomedical Publication; Chicago IL (available at http://garfield.library.upenn.edu/papers/jifchicago2005.pdf).

41. Bhandari M, Montori VM, Devereaux PJ, Wilczynski NL, Morgan D, Haynes RB. Doubling the impact: publication of systematic review articles in orthopaedic journals. J Bone Joint Surg Am 2004;86A:1012-6.

42. Callaham M, Wears RL, Weber E. Journal prestige, publication bias, and other characteristics associated with citation of published studies in peer-reviewed journals. JAMA 2002;293:2362-6.

43. Patsopoulos NA, Analatos AA, Ioannidis JP. Relative citation impact of various study designs in the health sciences. JAMA 2005;293:2362-6.

44. Lokker C, Haynes RB, Chu R, McKibbon KA, Wilczynski NL, Walter SD. How well are journal and clinical article characteristics associated with the journal impact factor? a retrospective cohort study. J Med Libr Assoc 2012;100:28-33.

45. Hilbe JM. Negative Binomial Regression. Cambridge UK: Cambridge University Press; 2007.

46. Schwarz G. Estimating the dimension of a model. Ann Statistics 1978;6:461-4.

47. Vrieze SI. Model selection and psychological theory: a discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). Psychol Methods 2012;17:228-43.

48. Streiner DL, Norman GR. Health Measurement Scales: A Practical Guide to Their Development and Use. 4th ed. Oxford UK: Oxford University Press; 2008.

49. Norman GR, Streiner DL. Measurements of association for ranked data. In: Norman GR, Streiner DL, eds. Biostatistics: The Bare Essentials. Hamilton ON: B.C. Decker, Incorporated; 2000:230-5.

50. McGlynn EA, Asch SM, Adams J, et al. The quality of health care delivered to adults in the United States. N Engl J Med 2003;348:2635-45.

51. Brownlee S. Overtreated: Why Too Much Medicine is Making Us Sicker and Poorer. New York: Bloomsbury; 2007.

52. Welch HG, Schwartz L, Woloshin S. Overdiagnosed: Making People Sick in the Pursuit of Health. Boston: Beacon Press; 2011.

53. Sarli CC, Dubinsky EK, Holmes KL. Beyond citation analysis: a model for assessment of research impact. J Med Libr Assoc 2010;98:17-23.

54. Bossuyt PMM, Reitsma JB, Bruns DE, et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. Ann Intern Med 2003;138:W1-W12.

55. Kulkarni AV, Aziz B, Shams I, Busse JW. Author self-citation in the general medicine literature. PLoS One 2011;6:e20885.

56. Eysenbach G. Citation advantage of open access articles. PLoS Biol 2006;4:e157.

57. Davis PM, Lewenstein BV, Simon DH, Booth JG, Connolly MJ. Open access publishing, article downloads, and citations: randomised controlled trial. BMJ 2008;337:a568.

58. Kulkarni AV, Aziz B, Shams I, Busse JW. Comparisons of citations in Web of Science, Scopus, and Google Scholar for articles published in general medical journals. JAMA 2009;302:1092-6.

59. Haynes RB, Cotoi C, Holland J, et al. Second-order peer review of the medical literature for clinical practitioners. JAMA 2006;295:1801-8.

60. Haynes RB, Holland J, Cotoi C, et al. McMaster PLUS: a cluster randomized clinical trial of an intervention to accelerate clinical use of evidence-based information from digital libraries. J Am Med Inform Assoc 2006;13:593-600.

61. Garfield E. The history and meaning of the journal impact factor. JAMA 2006;295:90-3.

62. Cone DC. Measuring the measurable: a commentary on impact factor. Acad Emerg Med 2012;19:1297-9.

63. Berghmans T, Meert AP, Mascaux C, Paesmans M, Lafitte JJ, Sculier JP. Citation indexes do not reflect methodological quality in lung cancer randomised trials. Ann Oncol 2003;14:715-21.

64. Grad R, Pluye P. Physicians' assessment of the value of clinical information: operationalization of a theoretical model. J Am Soc Inform Sci Technology 2011;62:1884-91.

Table 1

The BEEM Rater Scale

"Assuming that the results of this article are valid, how much does this article impact on EM clinical practice?"

1. Useless information
2. Not really interesting, not really new, changes nothing...
3. Interesting and new but doesn't change practice
4. Interesting and new, has the potential to change practice

5. New and important: this would probably change practice for some Emergency Physicians.
6. New and important: this would change practice for most Emergency Physicians
7. This is a "must know" for Emergency Physicians

Table 2
Summary of Abstracts Reviewed by BEEM Raters

Characteristic	Frequency, %
Year of Publication	
2006	2
2007	9
2008	42
2009	19
2010	17
2011	11
Study Category	
Diagnostic	27
Prognostic	14
Therapy	59
Study Design	
Systematic Review	37
Randomized controlled trial	32
Observational	30

Table 3[1]
Correlations with Citation Rate

Variable	Spearman's Rho
BEEM rater score	0.144
Number of authors	0.363
SCOPUS first author *h*-index	0.358
SCOPUS last author *h*-index	0.277
JCR Journal Impact Factor	0.519

[1] P< 0.001 for all variables.

Table 4

Negative Binomial Regression Models for Citation Rate and Bibliometric Indices

Variable	Wald chi-square	p-value	Odds Ratio (95% CI)
Model #2			
SCOPUS first author *h*-index	18.5	<0.001	1.017 (1.009-1.025)
SCOPUS last author *h*-index	3.9	0.05	1.005 (1.000-1.009)
JCR Journal Impact Factor	110.4	<0.001	1.049 (1.040-1.059)
Number of Authors	7.9	0.005	1.028 (1.008-1.047)
BEEM Rater score	3.0	0.082	1.108 (0.987-1.243)
Model #3			
Observational trial	1.2	0.272	0.886 (0.714-1.100)
Randomized trial	0.8	0.386	0.909 (0.733-1.128)
Number of Authors	8.7	0.003	1.032 (1.011-1.054)
SCOPUS first author *h*-index	19.1	<0.001	1.017 (1.009-1.025)
SCOPUS last author *h*-index	4.0	0.046	1.005 (1.000-1.009)
JCR Journal Impact Factor	102.0	<0.001	1.048 (1.039-1.058)
BEEM Rater score	3.2	0.071	1.112 (0.991-1.249)

BEEM Rater Scores Correlate with Publications' Future Citations. Christopher R. Carpenter, Cathy C. Sarli, Susan A. Fowler, Kulamakan Kulasegaram, Teresa Vallera, Pierre Lapaine, Grant Schalet, Andrew Worster. Academic Emergency Medicine. Volume 20, Issue 10, pages 1004–1012, October 2013. © 2013 by the Society for Academic Emergency Medicine.

Figure 1

Exemplary BEEM Rater Surveys for Two Manuscripts Identified by the McMaster Health Information Research Unit