Received: 22 June 2020 | Accepted: 25 July 2020

DOI: 10.1002/lio2.444

ORIGINAL RESEARCH

Laryngoscope Investigative Otolaryngology

Does quality of research in otolaryngology correlate with academic impact?

Aasif A. Kazi PharmD, MD | Nima A. Vahidi MD | John Sinkovich BA | Daniel H. Coelho MD

Department of Otolaryngology—Head and Neck Surgery, Virginia Commonwealth University School of Medicine, Richmond, Virginia

Correspondence
Daniel H. Coelho, Department of Otolaryngology—Head and Neck Surgery, Virginia Commonwealth University Health System, PO BOX 980146, Richmond, VA 23298-0146, USA.
Email: daniel.coelho@vcuhealth.org

Abstract

Objective: To determine if the quality of otolaryngology-related journal articles correlates with traditional measures of article impact.

Methods: All articles published by Laryngoscope in 2011 were categorized according to level of evidence (LOE) according to the Oxford Center for Evidence Based Medicine rubric. Articles without a level of evidence assigned were alphabetically subcategorized type with letters A-D corresponding to Contemporary Reports, Case Reports, Basic Science or Animal Studies, and Other respectively. Citations per article were then recorded per article each year from 2012 to 2018.

Results: A total of 494 articles were included for analysis, 315 had numerical LOE and 179 had alphabetical LOE. There was a strong negative correlation between numerical LOE and median and interquartile number of citations (R = –.9014, P = .037). Overall, numerical LOE had a significantly higher median number of citations per article compared with the non-number/alphabetical group (14 vs 6, P < .001).

Conclusions: Higher quality research as determined by level of evidence is in fact being cited more frequently than lower quality articles. Although the scope of this study was relatively limited, these data suggest that better designed studies may exhibit greater impact by traditional measures. Such findings should serve as an impetus for (and validation of) continued pursuit of high LOE research.

Level of Evidence: NA.

KEYWORDS

bibliometrics, citation, impact, level of evidence, quality

1 | INTRODUCTION

Evidence based medicine (EBM) is the combination of evidence based research and clinical expertise.[1] Several rubrics for determining levels of EBM have been devised, though the most widely utilized is the "Level of Evidence" (LOE), published in 2011 by the Oxford Center for Evidence Based Medicine (OCEBM). Based largely on methodology, the OCEBM LOE is considered by most clinicians and researchers to be the authoritative standard for determining quality. In recognition of the importance of EBM, researchers and journal editorial staff have

increased efforts to publish articles of higher quality. By most metrics including LOE, these efforts appear to be working.[2] Even in smaller fields such as Otolaryngology, progress is being made demonstrating improvement in the quality of research output over the last several decades.[3-7]

There are several potential explanations for the observed increase in quality, but essentially distill down to two factors—authors submitting higher quality work and/or journals seeking/selecting higher quality work. Both are likely true. Yet irrespective of the relative influences each of these factors may contribute, the overriding goal of any publication is to maximize impact. Although intuitive that higher quality research is more likely to be read, cited, shared, taught, and ultimately incorporated into clinical practice, this is wholly based on assumption and not necessarily true. The aim of this study was to determine if quality of otolaryngology related articles (as measured by LOE) correlates with one traditional measure of impact—number of citations. Such findings could have substantial impact on how we should value EBM in our field.

2 | METHODS

2.1 | Data sources

Articles were selected from all issues and supplements published by *Laryngoscope* in 2011 (volume 121). The complete list was obtained using PubMed and confirmed directly from the publisher website (https://onlinelibrary.wiley.com/loi/15314995/year/2011). *Laryngoscope* was chosen for a variety of reasons. First, *Laryngoscope* had the highest (by nearly double) Eigenfactor score in 2011 of any Otolaryngology journals. Eigenfactor score, unlike Impact Factor, gives more weight to citations in journals with higher readership (circulation) and therefore may be a more accurate representation of journal's effective academic impact.[8] Additionally, *Laryngoscope* had the highest Eigenfactor and Impact Factor for any non-sub-specialty journal in 2011. The year 2011 was chosen as this was the first full year that *Laryngoscope* required authors to include Level of Evidence assignment to their manuscripts, the first such journal to do so. By including the earliest possible year, this allowed for the longest amount of time to elapse in which to accrue citations.

Level of Evidence was set by the editorial staff of *Laryngoscope* based on theOCEBM Rubric (https://www.cebm.net/wp-content/uploads/2014/06/CEBM-Levels-of-Evidence-2.1.pdf). Authors were asked to assign a LOE to their paper, and the editorial staff wound amend as appropriate prior to final publication. Publications that did not meet the numerical criteria set by the OCEBM (and were assigned "not applicable" by the journal), were further broken down into alphabetical subcategories defined as (a) Contemporary Reports, (b) Case Reports, (c) Basic Science or Animal Studies, or (d) Other to determine if there were any particular types of articles that accounted for the citation number while not meeting the numerical criteria (Table 1). Note that these alphabetical classes were devised for this particular study, are not ordinal, and are not part of the *Laryngoscope* LOE system.

TABLE 1 Definitions of levels of evidence

Level of evidence	Definition
Numeric	
1	Randomized control trial
2	Lower quality randomized trials/cohort studies
3	Case-control studies
4	Case series/low quality cohort or case-control studies
5	Expert recommendation
Non-numeric	
A	Contemporary reports
B	Case reports
C	Basic science or animal studies
D	Other

Citations per article were obtained from Web of Science Cited Reference Search, accessed through the Virginia Commonwealth University Medical Library. Citations were all recorded on November February 3, 2019. Only data from complete years were included (2012-2018).

2.2 | Statistical analysis

The primary outcome measure was total number of citations per LOE. Mean and median were calculated. In the case that outcomes were heavily skewed (due to outliers leading to large standard deviations from the mean or from having LOEs with a small number of observations) the data would be summarized by Median and Inter-Quartile range by numeric and non-numeric LOE overall and by each LOE. Correlation between Median LOE (numeric only) and number of citations was performed by calculating the Pearson coefficient (r). Differences between means for each numerical and alphabetical LOE was performed using independent two-tailed t tests. All calculations were performed using Microsoft Excel (Bellevue, Washington).

This study was considered exempt from Institutional Review Board approval.

3 | RESULTS

A total of 497 articles were published in calendar year 2011. Three articles were removed because they contained sub-studies that were assigned multiple LOEs. This resulted in 494 articles included for analysis, of which 315 were numerical and 179 were alphabetical. The distribution of the numerical and alphabetical LOEs are listed in Table 2. The most common numerical LOE was "Level 2" (n = 140) and the most common alphabetical LOE was "Level D (Other)" (n = 72).

TABLE 2 Citations by level of evidence

Level of evidence	n (%)	Total citations median (IQR)
Numeric overall	315 (100)	14.0 (8.00, 24.00)
1	50 (15.9)	17.0 (8.25, 27.50)
2	140 (44.4)	15.0 (9.00, 24.30)
3	39 (12.4)	12.0 (8.00, 21.00)
4	72 (22.9)	13.0 (6.75, 19.00)
5	14 (4.4)	11.5 (6.00, 20.75)
Non-numeric overall	179 (100)	6.00 (2.00, 14.00)
A	27 (15.1)	8.00 (6.00, 16.00)
B	53 (29.6)	5.00 (2.00, 8.00)
C	27 (15.1)	13.0 (7.00, 20.00)
D	72 (40.2)	4.00 (0.00, 12.00)

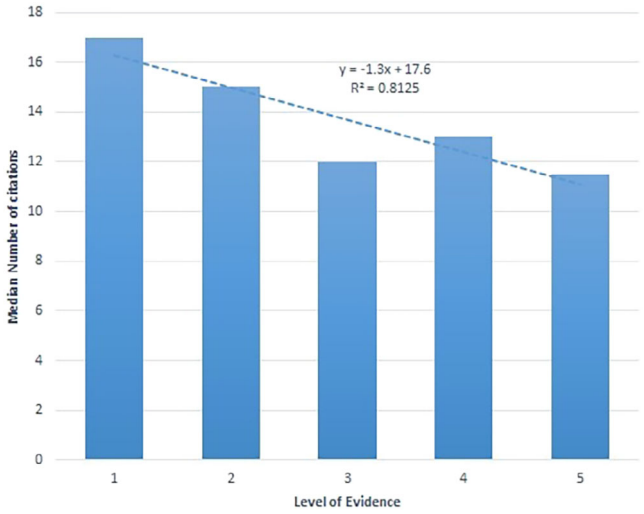

FIGURE 1 Median number of citations/article by level of evidence

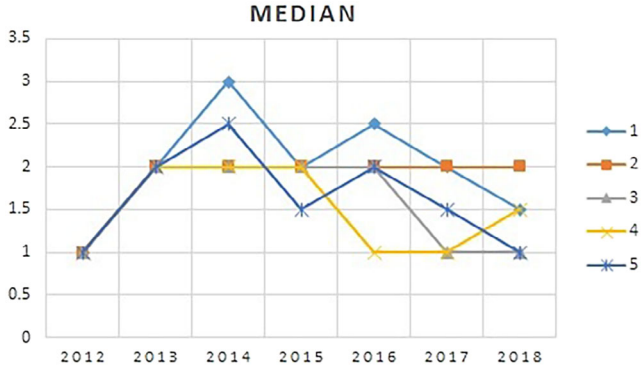

FIGURE 2 Yearly distribution of median citations per article per year by level of evidence

Mean citations per article by LOE did not demonstrate any statistically significant correlation. However, some LOEs demonstrated exceptionally high SDs from the mean. For example, level 5 articles had a mean citations/article of 39.7 over the study period, but with SD of ±83.6. When excluding LOE 5 from analysis, there was a strong negative correlation between LOE 1-4 and mean number of citations ($R = -.9635$, $P = .037$).

To include all LOEs, median values were analyzed and reported here (as per the *a priori* methodology established above). There was a strong negative correlation between numerical LOE and median number of citations ($R = -.9014$, $P = .037$) (Figure 1). Overall, numerical LOE had a significantly higher median number of citations per article compared with alphabetical LOE (14 vs 6, $P < .001$). For numerical LOE, statistical differences in median citations existed between LOE 2/5, (15 vs 11.5, $P = .009$), LOE 3/5 (12 vs 11.5, $P = .049$) and LOE 4/5 (13 vs 11.5, $P = .017$). For alphabetical LOE, statistical differences in median citations between LOE A/B (8 vs 5, $P = .033$), LOE B/C (5 vs 13, $P < .001$), and LOE C/D (13 vs 4, $P = .008$). Yearly distribution of median citations per article per year is presented in Figure 2.

4 | DISCUSSION

The aim of our study was to better understand the relationship between quality and impact—two terms that have historically been used interchangeably. Yet, they are quite different. While citation frequency has traditionally been used to evaluate the importance of a particular research article, quality is multidimensional and is difficult to assess. Current thinking places a higher emphasis on randomized controlled trials or meta-analyses as those are thought to be of the best quality. However even among randomized controlled trials or meta-analyses, "quality" is not explicitly defined. In a recent article by Fenton et al., the quality of the article was found to be related to the research topic, the research methods involved, the importance and the overall written quality of the study. Moreover, it should include the impact of the article on clinical practice.[9] However, as with "quality" there is no consensus definition of "impact."[10] In fact, impact can also be thought of as a multifactorial concept that changes depending on the point of view. For example, on a larger scale impact can be directly related to the appropriate dissemination of research thereby effecting policy making. However, this is rarely due to a single article making it difficult to determine the impact an individual article has had. Similarly, while the use of citation-based metrics seems like a natural way to quantify the impact of a particular article on other researchers, that would more aptly be categorized as a surrogate marker for utility rather than impact. It can also be influenced by citation bias and self-citations resulting in an inaccurate view of the evidence.[9,11] More recently, the use of Altmetrics as a way to assess impact has become a trend. (Altmetrics is defined as the quantification of the digital footprint of a particular article that is not confined to the publishers and scientific community that is, social media, blogs and news outlets.)[10]

This study is the first to examine the relationship between quality and impact in Otolaryngology. Similar assessments have been performed in other medical specialties including orthopedic surgery, psychiatry and oral and maxillofacial surgery. Sochacki et al. found no correlation between the level of evidence and the number of citations

in the top 50 articles related to rotator cuff repair surgery.[12] Nieminen et al. also did not find a significant relationship between the quality of psychiatric research among four journals and the total number of citations per article over 9 years. However, they only used half the articles and the quality of the research was evaluated by the primary author using their own rubric.[13] While Lau et al. attempted to correlate level of evidence to impact, they chose journal impact factor as their surrogate marker for impact which does not correspond to individual articles.[14]

The results of this study suggest that when using LOE as a surrogate for quality and citations as a surrogate for impact, quality does indeed correlate with impact. The correlation, while strong, is not perfect. This is due to the high number of outliers, especially in the lowest quality (LOE 5) group. In fact, LOE 5, which had the fewest number of articles overall, was found to have the highest number of mean citations per article during our study period, due to just a few very highly cited publications. This was accounted for in the statistical analysis and medians were used to provide a more realistic measure of the true value. Furthermore, these data support the notion that research for which LOE can be assigned according to the OCEBM rubric generally has more impact than those for which the OCEBM LOE are not-applicable (the "alphabetical" group in this study). When taken in total, it would appear that conventional wisdom holds true—higher quality work gets more traction. This is a rewarding validation for author and publisher alike.

This study is not without limitations. First, although chosen for specific reasons stated above, only articles from 1 year of one journal were included for analysis as it would provide the most consistently reproducible data points for analysis. As such, applicability of these findings may not necessarily be applied to the entire published field of Otolaryngology. As more time passes and more journals recognize the important of including LOE (or any metric of quality) in their publications, this will facilitate further analysis of the wider field by bibliometric researchers. In addition, for reasons outlined earlier, no perfect surrogate exists to measure either quality or impact. By using citations to represent impact this study focuses more on the impact of an individual on other researchers/authors, and not on the true intended audience—clinicians. Social media platforms and Altmetrics can help us to better understand how knowledge is shared *after* publication and may more accurately reflect impact on practice. This has been demonstrated in some early data from the Urology field.[15] Ongoing studies by our group are investigating the relationship between Altmetrics and impact. Finally, the use of alphabetical subcategories of LOE was completely arbitrary and designed for solely for this study so as to not limit our analysis to only articles with assigned LOE.

5 | CONCLUSION

This study found that higher quality research publications were cited more frequently than lower quality studies. This is encouraging as it not only reinforces the notion that publication platforms such as *Laryngoscope* have transitioned towards higher quality research content, but also that higher quality publications have a broader scope and impact. We believe that publishers should continue to encourage and disseminate higher quality publications, serving to reinforce the motivation authors have in pursing systematic research. We hope our findings continue to illustrate the importance of improving research quality, and further reaffirm that there is value in publishing higher quality content.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

ORCID

Aasif A. Kazi https://orcid.org/0000-0003-4755-2681
Daniel H. Coelho https://orcid.org/0000-0001-6485-2075

REFERENCES

1. Masic I, Miokovic M, Muhamedagic B. Evidence based medicine – new approaches and challenges. *Acta Inform Med*. 2008;16(4):219-225.
2. Hung BT, Long NP, Hung le P, et al. Research trends in evidence-based medicine: a joinpoint regression analysis of more than 50 years of publication data. *PLoS One*. 2015;10(4):e0121054.
3. Ah-See KW, Molony NC, Maran AG. Trends in randomized controlled trials in ENT: a 30-year review. *J Laryngol Otol*. 1997;111(7):611-613.
4. Gouveia CJ, Zaghi S, Awad M, et al. Publication trends and levels of evidence in obstructive sleep apnea literature. *Laryngoscope*. 2018;128(9):2193-2199.
5. Wasserman JM, Wynn R, Bash TS, Rosenfeld RM. Levels of evidence in otolaryngology journals. *Otolaryngol Head Neck Surg*. 2006;134(5):717-723.
6. Bell-Allen Nicholas, Cervin Anders, Vasani Sarju. Trends in the quality & quantity of otorhinolaryngology research in Australia: 1989–2018. *Australian Journal of Otolaryngology*. 2019;2:12-12. http://dx.doi.org/10.21037/ajo.2019.03.02.
7. Socolovsky Leandro D., Krishnan Pavan, Ramraj Rahul, Retnam Reuben, Sima Adam, Bauschard Michael J., Coelho Daniel H.. Quality of Otology and Neurotology Research in Otolaryngology Journals. *Otology & Neurotology*. 2020; Publish Ahead of Print. http://dx.doi.org/10.1097/mao.0000000000002731.
8. Bergstrom Carl. Eigenfactor: Measuring the value and prestige of scholarly journals. *College & Research Libraries News*. 2007;68(5):314-316. http://dx.doi.org/10.5860/crln.68.5.7804.
9. Fenton JE, O'Connor A, Ullah I, Ahmed I, Shaikh M. Do citation classics in rhinology reflect utility rather than quality? *Rhinology*. 2005;43(3):221-224.
10. Mackinnon S, Drozdowska BA, Hamilton M, Noel-Storr AH, McShane R, Quinn T. Are methodological quality and completeness of reporting associated with citation-based measures of publication impact? A secondary analysis of a systematic review of dementia biomarker studies. *BMJ Open*. 2018;8(3):e020331.
11. Nieminen P, Rucker G, Miettunen J, Carpenter J, Schumacher M. Statistically significant papers in psychiatry were cited more often than others. *J Clin Epidemiol*. 2007;60(9):939-946.
12. Sochacki KR, Jack RA 2nd, Nauert R, Harris JD. Correlation between quality of evidence and number of citations in top 50 cited articles in rotator cuff repair surgery. *Orthop J Sports Med*. 2018;6(6):2325967118776635.

13. Nieminen P, Carpenter J, Rucker G, Schumacher M. The relationship between quality of research and citation frequency. *BMC Med Res Methodol*. 2006;6:42.
14. Lau SL, Samman N. Levels of evidence and journal impact factor in oral and maxillofacial surgery. *Int J Oral Maxillofac Surg*. 2007;36(1):1-5.
15. O'Connor EM, Nason GJ, O'Kelly F, Manecksha RP, Loeb S. Newsworthiness vs scientific impact: are the most highly cited urology papers the most widely disseminated in the media? *BJU Int*. 2017;120(3):441-454.

How to cite this article: Kazi AA, Vahidi NA, Sinkovich J, Coelho DH. Does quality of research in otolaryngology correlate with academic impact? *Laryngoscope Investigative Otolaryngology*. 2020;5:819–823. https://doi.org/10.1002/lio2.444