Evidence-Based Medicine in Plastic Surgery

Are We There Yet?

Daniel G. Leal, MD, Maria Amelia Rodrigues,† Ana Carolina B. Tedesco,†*
*Fabio Xerfan Nahas, MD, PhD, MBA,† Lydia Masako Ferreira, MD, PhD, MBA,† Ana C.W. Roxo, MD, PhD,**
Claudio Cardoso de Castro, MD, PhD, and Jose Horacio Aboudib, MD, PhD**

Background: The practice of evidence-based medicine in plastic surgery is no longer a trend but a reality, with a growing number of studies published in recent years using evidence-based medicine as an assessment tool.

Objective: The aim of this study was to verify whether the number of citations to articles with a high level of evidence is greater than articles with low level of evidence.

Methods: A search was conducted in the 4 main international journals of plastic surgery. All original articles published in 2011 were analyzed, selected, and classified based on the study design. The articles were then divided into 2 groups: group 1, high level of evidence; and group 2, low level of evidence. Next, Scopus was searched for the number of citations of each article in the 2 subsequent years. The proportion of the number of citations received by articles in groups 1 and 2 was statistically compared.

Results: The articles with the highest level of evidence were the most cited among original articles, with 48.6% of them being cited more than 10 times over 2 years, whereas only 18.4% of articles in group 2 were cited with the same frequency. The mean number of citations was 12.6 citations per article in group 1 and 6.56 citations in group 2, with a significant difference between groups ($P < 0.0001$).

Conclusions: The articles with a higher level of evidence are, on average, cited more often than those with low levels of evidence in the leading journals of plastic surgery.

Key Words: evidence-based medicine, citations, impact factor

(*Ann Plast Surg* 2018;80: 71–75)

The practice of evidence-based medicine (EBM) in plastic surgery is no longer a trend but a reality, with a growing number of studies published in recent years using EBM as an assessment tool.[1] Despite its wide acceptance, some plastic surgeons are still reluctant to adapt EBM to the detriment of experience-based medicine.[2] For this group, most criticism of EBM is their loss of autonomy as surgeons, having to follow strict protocols to treat all patients the same way.[3] By definition, EBM "is the conscientious, explicit, and judicious use of current best evidence, combined with clinical experience, values, and needs of each patient in making decisions about the care of individual patients."[4] The application of the EBM principles not only can determine the best treatment for the patient but also may reduce costs for health systems.[5]

To stimulate the practice of EBM, consensus meetings were held with editorial board members of major journals in plastic surgery.[6,7] Some action points proposed in these meetings have been successfully carried out, such as the use of a level-of-evidence grading system for articles,[8] which has been used by leading journals. Other action points including standardization of the level-of-evidence categories, the search for improvement in quality of evidence, and creation of EBM groups within and across societies were also recommended by these meetings.

Today, after the consolidation of EBM in the scientific literature, it is important to observe the impact of these articles in terms of number of citations. Thus, the aim of this study was to verify whether the number of citations to articles with a high level of evidence is greater than that of articles with low level of evidence in major plastic surgery journals.

MATERIALS AND METHODS

A search was conducted in the 4 main international journals of plastic surgery, and all original articles published between January 1 and December 31, 2011, were analyzed.

The inclusion criteria for the journals were as follows: be among the 4 highest ranked plastic surgery journals with the largest impact factors listed in the ISI Web of Knowledge Journal Citation Reports (JCR), to publish articles on esthetic and reconstructive surgery in general, and be listed in the JCR for at least 5 years. The journals that did not publish articles in all areas of plastic surgery or exclusively publish nonsystematic reviews of the literature and those not listed in the JCR for at least 5 years were excluded from the study.

The selected journals were *Plastic and Reconstructive Surgery* (*PRS*); *Journal of Plastic, Reconstructive, and Aesthetic Surgery* (*JPRAS*); *Aesthetic Plastic Surgery* (*APS*); and *Annals of Plastic Surgery* (*AnPS*).

The PubMed database was then searched for articles published in 2011 by the selected journals. Filters for systematic reviews,

Received November 13, 2016, and accepted for publication, after revision May 18, 2017.

From the *Division of Plastic Surgery, State University of Rio de Janeiro, Rio de Janeiro; and †Division of Plastic Surgery, Federal University of São Paulo, São Paulo, Brazil.

The article has been seen and approved by all authors, and the authors accept full responsibility for the design and conduct of the study, had access to the data, and controlled the decision to publish.

All authors made substantial contributions to the article.

D.G.L. helped in the conception and design of the study, acquisition of data, analysis and interpretation of data, drafting the article, revising it critically for important intellectual content, and final approval of the version to be submitted.

A.C.B.T. contributed to the conception and design of the study, analysis and interpretation of data, revising it critically for important intellectual content, and final approval of the version to be submitted.

M.A.R. was partially responsible for conception and design of the study, analysis and interpretation of data, revising it critically for important intellectual content, and final approval of the version to be submitted.

A.C.W. R. assisted in the conception and design of the study, analysis and interpretation of data, revising it critically for important intellectual content, and final approval of the version to be submitted.

C.C.D.C participated in the design of the study, revising it critically for important intellectual content, and final approval of the version to be submitted.

J.H.A. was partially responsible for the design of the study, revising it critically for important intellectual content, and final approval of the version to be submitted.

L.M.F. helped in the design of the study, revising it critically for important intellectual content, and final approval of the version to be submitted.

F.X.N. contributed to the conception and design of the study, analysis and interpretation of data, drafting the article, revising it critically for important intellectual content, and final approval of the version to be submitted.

Conflicts of interest and sources of funding: none declared.

Reprints: Daniel G. Leal, MD, Division of Plastic Surgery, State University of Rio de Janeiro, Blvd 28 de Setembro, 77 Vila Isabel, Rio de Janeiro 20551–030, Brazil. E-mail: danileal.rlk@terra.com.br.

ISSN: 0148-7043/18/8001–0071
DOI: 10.1097/SAP.0000000000001179

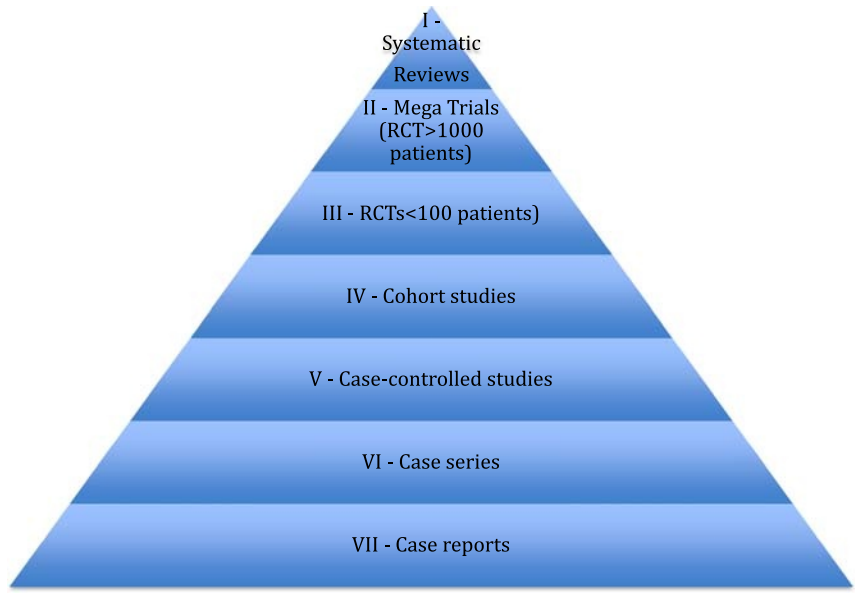

FIGURE 1. Cook et al's[2] level-of-evidence classification system based on the Cochrane criteria.

meta-analyzes, clinical trials, and randomized controlled trials (RCTs) were applied separately to each journal, using the "Advanced Search" function available at the PubMed site. Letters, editorials, communications, and nonoriginal articles were excluded. The retrieved articles were classified into 7 categories, according to the level of evidence classification system proposed by Cook et al[9] based on the Cochrane criteria (Fig. 1).

Two investigators with similar backgrounds and who completed the online course on systematic review provided by Cochrane Training independently selected and classified the articles based on the study design and patient selection.

The following criteria were established for evaluating randomized clinical trials: (1) description of the randomization procedure and (2) comparison of treatments.

The criteria for systematic reviews evaluated whether (1) the clinical question was defined, 2) the search strategy for retrieval of clinical studies was described, and if possible, (3) the statistical analysis (meta-analysis) was performed.

In cases of differing interpretation between investigators, a new analysis was performed together until a consensus was met on which studies would be considered adequate.

The articles were then divided into 2 groups: group 1, studies with levels of evidence I, II, and III; and group 2, studies with levels of evidence IV, V, VI, and VII.

Next, Scopus was searched for the number of citations of each original article in the 2 subsequent years. The proportion of the number of citations receive by articles in groups 1 and 2 was statistically compared for each journal of publication and for the total sample.

Statistical Analysis

The κ test was performed to evaluate interrater reliability (κ) for the level-of-evidence classification of articles. Fisher exact test was carried out to identify the journal with the greatest number of adequate studies. Altman's Kappa Benchmark scale[10] was used to characterize the extent of agreement among raters based on the reliability coefficient (κ). The Shapiro-Wilk test was applied to test the normality of the data, and the differences in proportions between the groups were tested with the χ^2 test. Student's *t* test was carried out to compare the mean number of citations between subgroups for each journal.

All statistical tests were performed at a significance level of 0.05 ($P < 0.05$).

RESULTS

A total of 1223 original articles were selected after excluding letters, notes, communications, errata, and nonoriginal articles; 116 of them were levels I, II, and III; 71 of which were systematic reviews, and 45 were RCTs. However, only 96 articles were considered adequate by rater consensus and therefore were included in the study.

A κ coefficient indicated moderate to excellent interrater reliability, according to the Altman's Kappa Benchmark scale, which classifies agreement among raters as poor (κ ≤ 0.20), fair (κ = 0.21−0.40), moderate (κ = 0.41−0.60), good (κ = 0.61−0.80), and excellent (κ = 0.81−1.00). Interrater reliability was moderate for the *PRS* ($P < 0.001$) and *JPRAS* ($P = 0.058$), good for the *APS* ($P = 0.018$), and excellent for the *AnPS* ($P = 0.003$).

The percentage of articles considered as adequate by rater consensus was similar among the journals (Table 1). The articles in group 1 (levels of evidence I, II, and III) that were considered inadequate by rater consensus with regard to their level of evidence were transferred to group 2 (levels of evidence IV, V, VI, and VII). Thus, the final sample (n = 1223) had 35 articles in group 1 and 1188 articles in group 2 (Fig. 2).

The *PRS* published 21 (60%) of the included articles; *JPRAS* published 7 (20%); *APS* had 4 (12%); *AnPS* published 3 (8%) articles (Table 1).

TABLE 1. Number of Articles Indexed in PubMed as Systematic Reviews and RCTs Considered Adequate or Inadequate by Rater Consensus in Each Journal

Journal	Adequate	Adequate, %	Inadequate	Inadequate, %	Total
PRS	21	34.4	40	65.6	61
JPRAS	7	41.2	10	58.8	17
APS	4	44.4	5	55.6	9
AnPS	3	33.3	6	66.7	9
Total	35	36.4	61	63.6	96

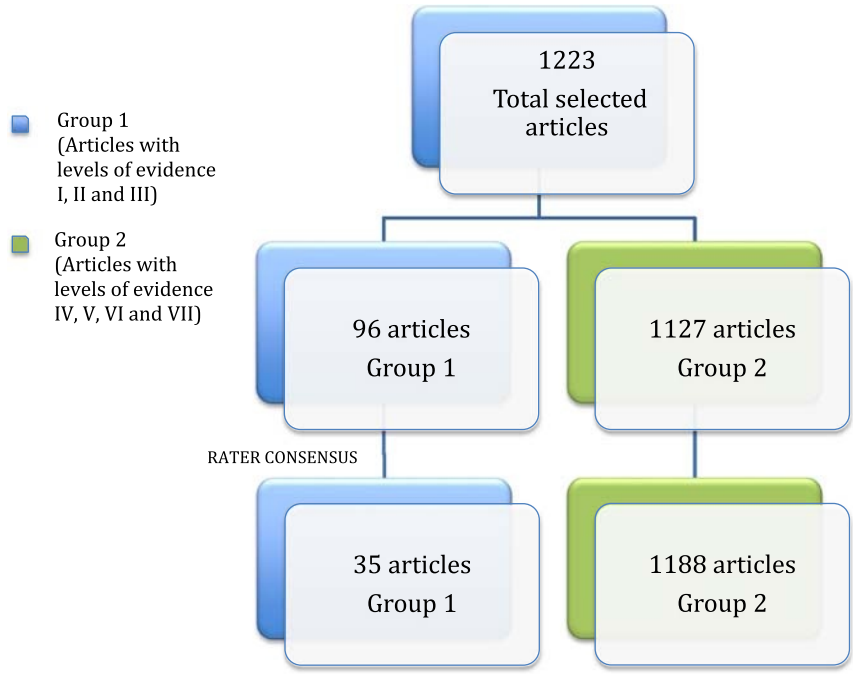

FIGURE 2. Flowchart of the article selection process, according to levels of evidence.

The articles with the highest level of evidence (group 1) were the most cited among original articles, with 48.6% of them being cited more than 10 times over 2 years, whereas only 18.4% of articles in group 2 were cited with the same frequency. These percentages are similar to those found for individual journals (Table 2 and Figs. 3 – 7). The mean number of citations was 12.6 citations per article (range, 0–43) in group 1 and 6.56 citations per article (range, 0–76) in group 2, with a significant difference between the groups ($P < 0.0001$). However, no article in group 1 was listed among the 10 most cited articles in 2013 (Table 3).

DISCUSSION

The practice of EBM in plastic surgery has been encouraged by different plastic surgery and related specialty societies. However, despite the growing number of articles published using this methodology, the use of EBM is still very limited. In 2009, only 32 articles with level of evidence I (according to the pyramid used by plastic surgery journals) were published in the 5 major plastic surgery journals.[10] It is important to differentiate the pyramid used in this article from the one that most plastic surgery journals are using. Levels of evidence I to III used in this study correspond to level I in the plastic surgery journals pyramid.

Despite the fact that recommendations to the practice of EBM lead to the generation of data with greater methodological strength in the form of meta-analysis and systematic reviews, theoretically providing a more reliable scientific basis,[11] there still remains a certain amount of criticism.[3] Some surgeons believe that EBM would serve as a "cookbook" for the medical practice, that it should be applied only to nonsurgical specialties, and that not always the best evidence is the same as best practices.[2] Other authors view systematic reviews in surgery as anachronistic publications that may compare techniques from different periods when some materials did not even exist. They consider it impossible to compare the surgical to nonsurgical treatments when procedures are performed with different materials or technologies.[12] However, these same authors believe that EBM can help improve the scientific

TABLE 2. Frequency Distribution of the Number of Article Citations in Groups 1 and 2 for the Total Sample and Per Journal

	No. Citations, n (%)			
Group (Journal)	**0–5**	**6–10**	**>10**	***P***
Group 1 (total)	8 (22.9)	10 (28.6)	17 (48.6)	<0.001*
Group 2 (total)	727 (61.2)	243 (20.5)	218 (18.4)	
Group 1 (*AnPS*)	1 (33.3)	1 (33.3)	1 (33.3)	0.21
Group 2 (*AnPS*)	193 (74.5)	43 (16.6)	23 (8.9)	
Group 1 (*PRS*)	3 (14.3)	7 (33.3)	11 (52.4)	0.07
Group 2 (*PRS*)	176 (39.5)	119 (26.7)	151 (33.9)	
Group 1 (*JPRAS*)	2 (28.6)	1 (14.3)	4 (57.1)	<0.001*
Group 2 (*JPRAS*)	240 (72.9)	54 (16.4)	35 (10.6)	
Group 1 (*APS*)	2 (50.0)	1 (25.0)	1 (25.0)	0.25
Group 2 (*APS*)	118 (76.6)	27 (17.5)	9 (5.8)	

* Statistical significance (χ^2 test; $P < 0.05$).

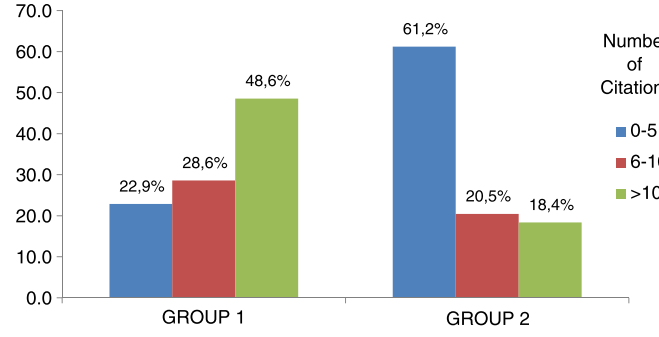

FIGURE 3. Frequency distribution of the number of articles in groups 1 and 2 in relation to their number of citations for the total sample (the 4 main journals in plastic surgery together).

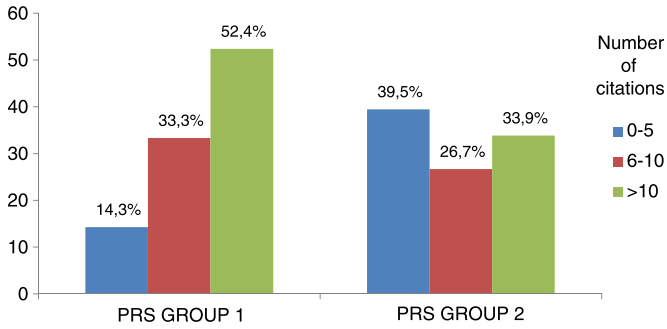

FIGURE 4. Frequency distribution of the number of articles in groups 1 and 2 in relation to their number of citations for the journal *PRS*.

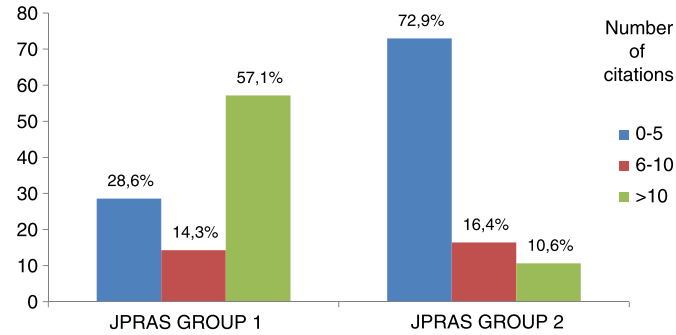

FIGURE 6. Frequency distribution of the number of articles in groups 1 and 2 in relation to their number of citations for the *JPRAS*.

quality by standardizing the methodology of future studies, especially in plastic surgery.[12]

This concern is even greater in cosmetic surgery where the results are usually measured by the presence of complications, number of revision surgeries, and patient satisfaction with the esthetic results.[13] Although complications are not frequent, the concept of beauty is subjective, and the surgeon's view of a good result may not accord with that of the patient. Double-blinded studies in surgery are difficult to design because surgeons can rarely be blinded. Most plastic surgery articles are self-controlled, which may be good for comparative studies but may not fit the criteria for a RCT. Thus, standardized and validated outcome assessment instruments, such as scales and questionnaires, should be used in the search for more objective and reliable results.

The classic concept of EBM[14] corresponds to the intersection of the best available evidence with clinical experience and the values and needs of each patient. By producing the best scientific evidence, researchers are, in fact, seeking the best for the patient. Clinical experience, surgical creativity, and common sense should be added to the best available treatment for a particular patient. This is a way of saying that "one size does not fit all."

With the rapid development of new technologies in medical practice, EBM can help improve the quality of care by reducing complications and lower costs through cost-effectiveness calculations[15] and, in a more comprehensive way, prioritize the allocation of resources, abandoning empiricism.[16]

At the same time, scientific articles sponsored by the biomedical and pharmaceutical industry should be carefully evaluated because this industry prioritizes study models with high statistical power. On the other hand, most publications based on industry-funded research tend to draw proindustry conclusions.[17] According to Fugh-Berman et al,[18] the pharmaceutical industry spent US$448 million in newspaper ads in 2003, and every dollar invested yielded a return of US$5 on average.

The researchers want to publish their articles in journals with a high impact factor, expecting to receive a large number of citations. However, the journal impact factor alone is not a good evaluation factor of the scientific relevance of individual articles, because a third of all articles published in leading journals of plastic surgery are cited only once or not at all.[19] Thus, the analysis of citation rates may be the best bibliometric measure to assess the impact of an individual article or author in the scientific community. For this reason, the number of citations to the selected articles was used in this study as a way to assess the impact of individual articles.

In recent years, an increase in the number of articles with high levels of evidence has been observed. However, no studies in plastic surgery showing a positive correlation between level of evidence and number of citations were found. This study was designed to evaluate whether articles showing the greatest scientific rigor are also the most cited by the scientific community.

The year 2011 was chosen for the selection of articles based on the method for calculating the impact factor of journals, which is obtained by the average number of citations received per article published in that journal during the 2 preceding years. It is also known that the number of citations received by an individual article in the first 2 years is correlated to the number of citations that it will receive in the following years.[20]

After rater consensus, only 2.86% (n = 35) of the selected original articles were considered as having levels of evidence I to III. Although the percentage of adequate articles was low, it is consistent with previous studies,[21] and historical analyses show an increasing trend.[1,10] Medical studies designed to provide evidence-based

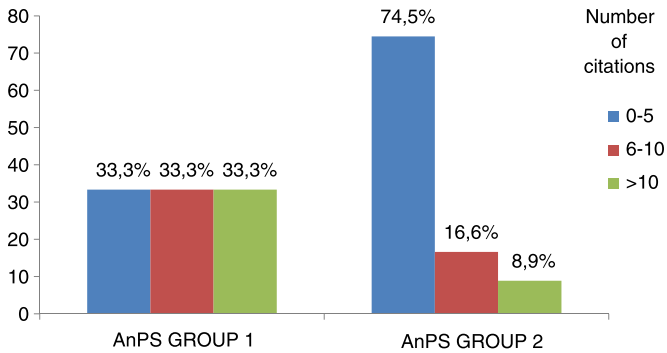

FIGURE 5. Frequency distribution of the number of articles in groups 1 and 2 in relation to their number of citations for the journal *AnPS*.

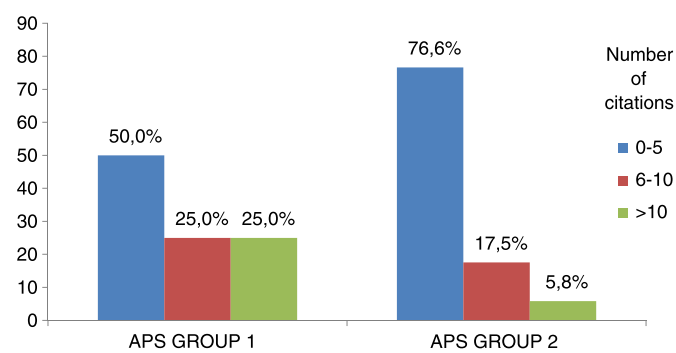

FIGURE 7. Frequency distribution of the number of articles in groups 1 and 2 in relation to their number of citations for the journal *APS*.

TABLE 3. The 10 Most Cited Articles, Published in 2011, as in 2013

Article	Journal	No. Citations
1. An 8-year experience of direct-to-implant immediate breast reconstruction using human acellular dermal matrix (AlloDerm)	*PRS*	76
2. Hyperbaric oxygen: its mechanisms and efficacy	*PRS*	59
3. Postoperative complications in prosthesis-based breast reconstruction using acellular dermal matrix	*PRS*	53
4. Expression of components of the renin-angiotensin system in proliferating infantile haemangioma may account for the propranolol-induced accelerated involution	*JPRAS*	51
5. Low-dose propranolol for infantile haemangioma	*JPRAS*	51
6. Retrospective review of 331 consecutive immediate single-stage implant reconstructions with acellular dermal matrix: indications, complications, trends, and costs	*PRS*	49
7. Propranolol as first-line treatment for rapidly proliferating infantile haemangiomas	*JPRAS*	48
8. Characteristic indocyanine green lymphography findings in lower extremity lymphedema: the generation of a novel lymphedema severity staging system using dermal backflow patterns	*PRS*	47
9. Acellular dermal matrix for the treatment and prevention of implant-associated breast deformities	*PRS*	45
10. Update on negative-pressure wound therapy	*PRS*	45

information are more expensive and time demanding, and therefore, some researchers may feel discouraged to perform them.[21] In addition, even studies of this nature may lack the necessary scientific rigor, especially regarding the description of the randomization procedure.[22] The difficulty of reaching an interexaminer consensus on attributing levels of evidence based on methodological quality highlights the difficulties met by authors in the selection of articles for systematic reviews on plastic surgery. In this study, only 35 of the 96 selected articles encompassing systematic reviews and RCTs were considered adequate.

Authors should look for journals that provide best evidence articles and not be limited to the impact factor. Rodrigues et al[23] proposed the journal evidence index, which correlates level of evidence with the journal impact factor and is calculated by dividing the number of RCTs by the total number of articles published in a specific journal over 1 year and multiplying by 100. The authors found that the *PRS* had the highest journal evidence index among the 4 plastic surgery journals studied, which justifies the largest number of articles with highest levels of evidence found in this journal in the present study.

When compared with group 2, articles in group 1 (with levels of evidence I to III) not only received the highest number of citations but also had the highest mean number of citations per article (Table 2). This shows that articles developed with greater scientific rigor tend to be more cited. However, despite statistical significance, there was a remarkable difference in the number of articles between the 2 groups. The articles with a large number of citations in a small group may result in an overestimate of the mean rate.

The individual analysis of the articles and their number of citations showed that no article in group 1 was among the top 10 most cited articles of the year. Thus, even though articles in group 1 had the highest mean number of citations, these were not necessarily the most cited articles individually. This could be explained by the low number of articles or the lack of hot topics research articles in this group.

The Scientometrics is the study of the quantitative aspects of science and scientific production. It is a very broad and important field in the analysis of the degree of relevance of EBM in plastic surgery. Despite the low number of articles in the literature on this topic, which may preclude a more accurate analysis, it is undeniable that there is a trend to pursue research with a greater scientific rigor, as evidenced by the higher number of citations received by articles providing high levels of evidence.

CONCLUSION

Despite the low the number of articles in plastic surgery produced annually with high levels of evidence, these articles are, on average, cited more often than those with low levels of evidence in the leading journals of plastic surgery.

REFERENCES

1. Becker A, Blümle A, Momeni A. Evidence-based Plastic and Reconstructive Surgery: developments over two decades. *Plast Reconstr Surg*. 2013;132:657e–663e.
2. Tebbetts JB. Evidence-based medicine in breast augmentation: is the goal to promote average or offer the most state-of-the-art treatment for patients? *Plast Reconstr Surg*. 2011;128:596–597. discussion 597–598.
3. Rhee JS, Daramola OO. No need to fear evidence-based medicine. *Arch Facial Plast Surg*. 2012;14:89–92.
4. Sackett DL, Rosenberg WM, Gray JA, et al. Evidence based medicine: what it is and what it isn't. *BMJ*. 1996;312:71–72.
5. Shauver MJ, Chung KC. Applying economic principles to outcomes analysis. *Clin Plast Surg*. 2013;40:281–285.
6. Rohrich RJ, Eaves FF 3rd. So you want to be an evidence-based plastic surgeon? A lifelong journey. *Plast Reconstr Surg*. 2011;127:467–472.
7. Eaves FF 3rd, Rohrich RJ, Sykes JM. Taking evidence-based plastic surgery to the next level: report of the second summit on evidence-based plastic surgery. *Aesthet Surg J*. 2013;33:735–743.
8. Sullivan D, Chung KC, Eaves FF 3rd, et al. The level of evidence pyramid: indicating levels of evidence in plastic and reconstructive surgery articles. *Plast Reconstr Surg*. 2011;128:311–314.
9. Cook DJ, Guyatt GH, Laupacis A, et al. Clinical recommendations using levels of evidence for antithrombotic agents. *Chest*. 1995;108(suppl 4):227S–230S.
10. McCarthy JE, Chatterjee A, McKelvey TG, et al. A detailed analysis of level I evidence (randomized controlled trials and meta-analyses) in five plastic surgery journals to date: 1978 to 2009. *Plast Reconstr Surg*. 2010;126:1774–1778.
11. Maier RV. What the surgeon of tomorrow needs to know about evidence-based surgery. *Arch Surg*. 2006;141:317–323.
12. Nahas FX, Ferreira LM. Evidence-based medicine: the past and the future of plastic surgery. *Plast Reconstr Surg*. 2014;134:499e–500e.
13. Alderman A, Chung KC. Measuring outcomes in aesthetic surgery. *Clin Plast Surg*. 2013;40:297–304.
14. Swanson JA, Schmitz D, Chung KC. How to practice evidence-based medicine. *Plast Reconstr Surg*. 2010;126:286–294.
15. Brown GC, Brown MM, Sharma S. Health care in the 21st century: evidence-based medicine, patient preference-based quality, and cost effectiveness. *Qual Manag Health Care*. 2000;9:23–31.
16. Niessen LW, Grijseels EW, Rutten FF. The evidence-based approach in health policy and health care delivery. *Soc Sci Med*. 2000;51:859–869.
17. Momeni A, Becker A, Bannasch H, et al. Association between research sponsorship and study outcome in plastic surgery literature. *Ann Plast Surg*. 2009;63:661–664.
18. Fugh-Berman A, Alladin K, Chow J. Advertising in medical journals: should current practices change? *PLoS Med*. 2006;3:e130.
19. Loonen MP, Hage JJ, Kon M. Value of citation numbers and impact factors for analysis of plastic surgery research. *Plast Reconstr Surg*. 2007;120:2082–2091. discussion 2092–2094.
20. Momeni A, Becker A, Antes G, et al. Evidence-based plastic surgery: controlled trials in three plastic surgical journals (1990 to 2005). *Ann Plast Surg*. 2009;62:293–296.
21. Chung KC, Ram AN. Evidence-based medicine: the fourth revolution in American medicine? *Plast Reconstr Surg*. 2009;123:389–398.
22. Ahmed Ali U, van der Sluis PC, Issa Y, et al. Trends in worldwide volume and methodological quality of surgical randomized controlled trials. *Ann Surg*. 2013;258:199–207.
23. Rodrigues MA, Tedesco AC, Nahas FX, et al. Journal impact factor versus the evidence level of articles published in plastic surgery journals. *Plast Reconstr Surg*. 2014;133:1502–1507.