



Level of Evidence in Facial Plastic Surgery Research: A Procedure-Level Analysis

Michael Eggerstedt¹ · Hannah J. Brown² · Aryan D. Shay² · Jennifer Westrick³ · Ethan M. Ritz⁴ · Ryan M. Smith¹ · Peter C. Revenaugh¹



Received: 20 March 2020 / Accepted: 8 April 2020 / Published online: 23 April 2020
© Springer Science+Business Media, LLC, part of Springer Nature and International Society of Aesthetic Plastic Surgery 2020

Abstract

Background As evidence-based medicine has taken hold across medical specialties, the level of evidence within the facial plastic surgery literature has risen, but remains weak in comparison. There has not yet been a systematic, critical appraisal of the relative strength of evidence among subsets of the practice of facial plastic surgery.

Methods The current study is a systematic review, designed to evaluate the level of evidence observed in the facial plastic surgery literature. Five journals were queried using facial plastic surgery terms for four selected years over a 10-year period. Following screening, articles were assigned to a category regarding subject matter, assessed for the presence of various methodological traits, and evaluated for overall level of evidence. Comparisons were made in regard to level of evidence across the breadth of facial plastic surgery subject matter.

Results A total of 826 articles were included for final review. Studies on operative facial rejuvenation and rhinoplasty had significantly fewer authors on average than studies on cancer reconstruction or craniofacial topics.

Craniofacial studies demonstrated higher levels of evidence relative to all other categories, with the exception of facial paralysis and facial trauma studies, from which there was no significant difference. In general, reconstructive studies had significantly more authors and higher levels of evidence than did articles with an aesthetic focus.

Conclusion Level of evidence in facial plastic surgery remains relatively weak overall. Reconstructive and particularly craniofacial studies demonstrate higher mean level of evidence, relative to other subsets of facial plastic surgery.

Level of Evidence III This journal requires that authors assign a level of evidence to each article. For a full description of these Evidence-Based Medicine ratings, please refer to the Table of Contents or the online Instructions to Authors www.springer.com/00266

Keywords Level of evidence · Facial plastic · Randomized controlled · RCT · Aesthetic · Reconstructive

Introduction

Dr. David Sackett, a pioneer of evidence-based medicine (EBM), writes that “the first sin committed by experts consists in adding their prestige and their position to their opinions, which give the latter far greater persuasive power than they deserve on scientific grounds alone” [1]. In an effort to combat this dogma, evidence-based medicine is the practice of applying the best and most appropriate evidence to varying clinical scenarios [2]. Critical to the foundation of EBM is the use of levels of evidence (LOE), which is a hierarchal rating system used to grade the methodological strength of any given study. EBM aids researchers in creating more robust analyses, allows

✉ Michael Eggerstedt
Michael_eggerstedt@rush.edu;
Michael.eggerstedt@gmail.com

¹ Section of Facial Plastic and Reconstructive Surgery, Department of Otorhinolaryngology-Head and Neck Surgery, Rush University Medical Center, 1611 W. Harrison St., Suite 550, Chicago, IL 60612, USA

² Rush Medical College, Rush University Medical Center, Chicago, IL, USA

³ Library of Rush University Medical Center, Rush University, Chicago, IL, USA

⁴ Bioinformatics and Biostatistics Core, Rush University, Chicago, IL, USA

clinicians to critically appraise external evidence, and optimizes diagnosis and treatment protocols, ultimately enabling thoughtful improvements in patient care.

Despite this, EBM has been less widely applied within the surgical community relative to medical counterparts. While 30–50% of decisions in medicine are based on findings from randomized controlled trials (RCT), only 10–20% of all surgical decisions are [3]. Skeptics of EBM's application in surgery suggest that wide adoption would lead to an algorithmic “cook-book” approach for the field and that best evidence does not always translate to best practice [4].

Both facial and full-body plastic surgery have been slow to pivot toward evidence-based medicine; however, recent years have shown a trend toward higher levels of evidence [5]. Though this overarching direction is encouraging, the trend among subsets of facial plastic and reconstructive surgery has not been described. In the current study, we aim to further examine the LOE within the field of facial plastic and reconstructive surgery by focusing on seven areas: office-based facial rejuvenation, surgical facial rejuvenation, rhinoplasty, facial paralysis, trauma, cancer reconstruction, and craniofacial surgery.

Materials and Methods

The current study is a systematic review, designed to evaluate the LOE observed in various subsets of the facial plastic surgery literature. This study is formatted in accordance with PRISMA guidelines. In conjunction with a medical research librarian, we systematically searched the literature for studies that met our criteria. Inclusion criteria consisted of articles matching facial plastic surgery-related keywords from the years 2008, 2011, 2014, and 2017. Articles were limited to those published in one of five journals: JAMA Facial Plastic Surgery, Facial Plastic Surgery, The Laryngoscope, Otolaryngology—Head and Neck Surgery, and Plastic and Reconstructive Surgery. Exclusion criteria included non-English-language articles, articles unrelated to facial plastic surgery, responses, recurring journal features, editorials, letters to the editor, and meeting/symposium reports. PubMed and Scopus were queried for terms relating to both surgical and nonsurgical treatment of aesthetic and reconstructive pathologies. Searches were completed in March 2019. There were no limitations on geographic area or age of study participants. Following de-duplication, 1840 articles were imported into commercial systematic review management software (Covidence Inc., Melbourne, Australia). The final search string can be found in supplementary materials.

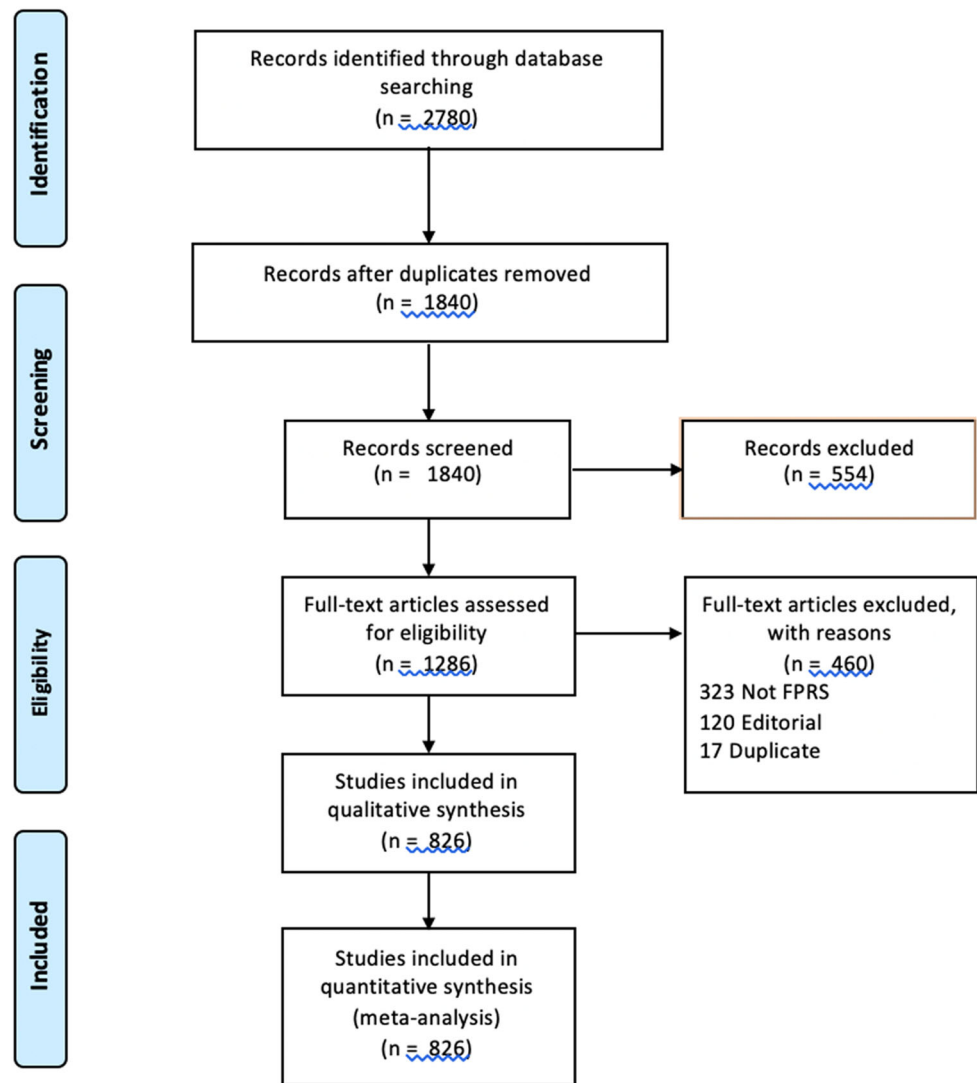
The title and abstract of each article were screened and resulted in exclusion of 554 articles based on the criteria

above. The full text of the remaining 1286 articles was assessed by one surgeon reviewer (ME), and any concerns were settled by the senior author (PCR). Following these screens, 826 articles were included in the final dataset for extraction (Fig. 1). Each article was classified as either aesthetic or reconstructive and was further classified into one of eight subcategories (office-based facial rejuvenation, peels, lasers, fillers, radiofrequency ablation, neurotoxins; facelift, blepharoplasty, facial implants; rhinoplasty, septal perforation, nasal obstruction; facial paralysis; trauma, fractures; cancer reconstruction, Mohs reconstruction, free flaps; cleft and craniofacial surgery; other). Additional variables collected included study design, subject matter, number of authors, year published, population studied, and presence of p values or confidence intervals. Based on these factors, a determination of overall LOE was assigned to each study according to the Oxford Centre for Evidence-based Medicine—Levels of Evidence [6]. As outlined by the Centre for Evidence-Based Medicine document, nonclinical, basic science research or those studies that are based on physiology and “first principles” are considered level 5 evidence.

Differences in continuous variables between nominal categories were assessed using analysis of variance (ANOVA), followed by post hoc pairwise tests using the Tukey's method. χ^2 tests or Fisher's exact tests were used for categorical data depending on cell frequencies. Post hoc testing on categorical data was performed by running pairwise Fisher's exact tests followed by controlling for multiple comparisons using the false discovery rate method. Changes in the percent of studies reporting specific LOE over time were analyzed using the Cochran–Armitage test for trend. All statistical analyses were performed using SPSS statistical software, version 24.0.0.0 (IBM Corp SPSS Statistics for Macintosh, Version 24.0. Armonk, NY). Statistical significance was considered at $P < 0.05$.

Results

A total of 826 articles met inclusion criteria and were included for analysis. These studies were classified as aesthetic or reconstructive and were divided further into eight categories based on content (Table 1, Fig. 2). Overall mean number of authors per study was 4.81 ± 2.91 . Studies classified as office-based facial rejuvenation or rhinoplasty had significantly fewer authors on average than studies discussing cancer reconstruction, craniofacial surgery, or other ($P < 0.0001$). Studies on facial paralysis and facial trauma did not demonstrate statistically significant differences in number of authors relative to other categories (Table 2, Fig. 3).

Fig. 1 PRISMA diagram**Table 1** Articles by subject

Subject	Number of articles	Mean number of authors per article (\pm SD)	Mean LOE (\pm SD)
1. Office-based facial rejuvenation	82	4.10 \pm 2.51	4.18 \pm 1.22
2. Surgical facial rejuvenation	91	3.54 \pm 2.01	4.23 \pm 0.92
3. Rhinoplasty	216	3.65 \pm 2.11	4.09 \pm 1.11
4. Facial paralysis	77	4.48 \pm 2.98	4.00 \pm 1.14
5. Facial trauma	56	4.30 \pm 2.14	3.86 \pm 1.14
6. Cancer reconstruction	162	4.72 \pm 2.43	4.08 \pm 0.91
7. Cleft and craniofacial	94	4.82 \pm 2.29	3.51 \pm 1.18
8. Other	47	5.17 \pm 3.05	4.36 \pm 0.79

SD standard deviation, LOE level of evidence

The overall mean LOE across all categories of facial plastic surgery was 4.04 ± 1.08 . Less than 3% of studies were considered level I evidence, while over 80% of studies were of level IV or V evidence (Table 2). The mean

LOE was then calculated for each category of subject matter (Table 1). Craniofacial surgery demonstrated higher LOE relative to all other study categories ($P < 0.0001$), with the exception of those regarding facial paralysis and

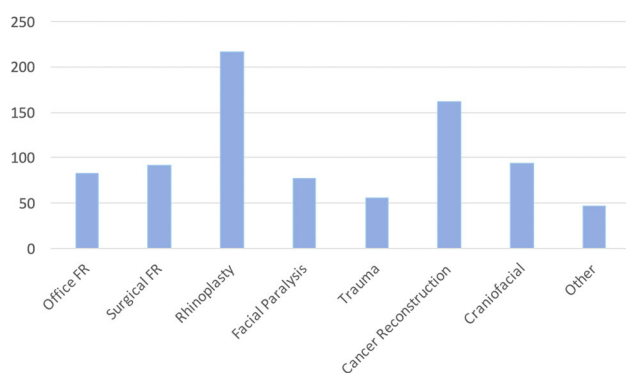


Fig. 2 Number of articles by subject. *FR* facial rejuvenation

Table 2 Overall level of evidence

LOE	Articles	% articles
1a, 1b, 1c	23	2.78
2a, 2b, 2c	101	12.22
3a, 3b, 3c	26	3.15
4	347	42.01
5	329	39.83
Total	826	100

LOE level of evidence

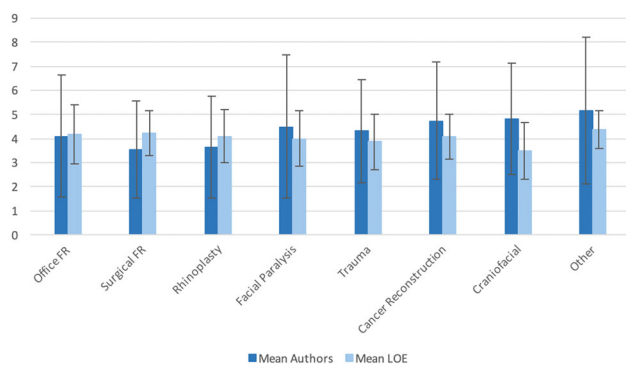


Fig. 3 Mean authors and LOE. *LOE* level of evidence, *FR* facial rejuvenation

facial trauma, from which there was no significant difference. With regard to the presence of RCTs, office-based rejuvenation had a higher proportion of RCTs (8.54%) than did head and neck cancer reconstruction ($P = 0.0004$). No other categories demonstrated a discrepancy in the relative proportion of RCTs. There were no significant differences among categories with regard to the presence of p values or confidence intervals ($P = 0.091$, $P = 0.083$).

Of the 826 articles analyzed, 502 were broadly identified as reconstructive and 324 as aesthetic studies. Reconstructive studies had significantly more authors than aesthetic studies (4.55 vs. 3.75, $P < 0.0001$). Additionally,

reconstructive studies exhibited significantly higher LOE overall than did articles with an aesthetic focus (3.93 vs. 4.21, $P = 0.0002$).

Discussion

The current study demonstrates trends consistent with those noted in prior studies, including a relatively low LOE and high number of authors across facial plastic surgery articles. With regard to the number of authors per study, our results are consistent with prior evidence finding a high number of authors per publication. Xu et al. [5, 6] found a significant increase in mean authorship from 1999 to 2008, with 4.13 authors per paper in the 2008 cohort. Similarly, Wasserman's group found an increase in mean otolaryngology authorship between 1993 and 2003, with the 2003 cohort having 4.2 authors per article [7]. Though the current study was not designed or powered to assess for change in authorship over time, our later cohort anecdotally demonstrates a higher mean authorship than either of these previous studies. The current study also found higher mean authorship for cancer reconstruction and craniofacial topics than for surgical rejuvenation and rhinoplasty topics. To our knowledge, this is the first study to note this discrepancy; however, it is expected, in light of the team-based approach to cancer and craniofacial care, in comparison with the relatively autonomous care of aesthetic facial plastic surgery. Within facial rejuvenation, rhinoplasty, and other aesthetic topics, a significant number of articles consist of the surgeon alone, whereas head and neck cancer reconstruction and craniofacial topics often included an ablative surgeon, oromaxillofacial surgeons, speech and language pathologists, medical/radiation oncologists, trainees, and other team members. Whether this can be interpreted as to aesthetic manuscripts being more often written by more senior authors and surgeons may be an avenue for further research. In general, a lower mean authorship should not be taken as a weakness of the literature, but perhaps a strength in that it better approximates the understanding of more senior providers.

When assessing the LOE across facial plastic surgery as a whole, our results were again consistent with prior studies that demonstrate an overall low LOE. The majority of included studies were of level IV evidence, as is consistent with multiple prior authors [8–12]. Chang et al. systematically isolated the 50 most cited papers in facial plastic surgery and found that even among these high-impact studies, over half were of level IV or level V evidence, and only one was of level I evidence [13]. Analogously, a similar study was performed for the full-body plastic surgery literature, demonstrating that 42 of the top 50 cited papers were of level IV or level V evidence [8]. This led

the authors to conclude that there was no discernible positive correlation between number of citations and LOE within the field. When aesthetic surgery is compared to otolaryngology as a whole, we see sixfold less prevalence of level I evidence within the aesthetic surgery literature [9]. However, despite these harrowing statistics, LOE within the facial plastic surgery literature has demonstrated steady gains over the past 20 years [14].

When LOE was considered in regard to subject matter, the cleft and craniofacial segment demonstrated significantly higher LOE than nearly all other categories, with the exception of facial paralysis and facial trauma. There is potential for bias in the current study regarding this comparison, as the journal with the highest impact factor is also the only journal with a dedicated craniofacial content section. Another possible explanation for craniofacial surgery's higher LOE is the widespread adoption of objective cephalometric data into the specialty. Regardless, the trend is consistent with the relative strength of reconstructive procedures observed in the current study.

Aside from the "other" category, surgical facial rejuvenation demonstrated the overall lowest LOE within the current review, followed closely by office-based facial rejuvenation. This speaks to the inherent difficulties in performing high-level randomized or controlled investigations into procedures that are elective and often via direct compensation models. By comparison, the topic of facial trauma demonstrated a relatively high level of evidence. This was a surprising finding, due to the inherent difficulties with consent and capacity within the trauma setting. It may be that the delayed, often semi-elective aspect of facial fractures and trauma is responsible for this fact, in contrast to the acuity of general trauma surgery.

Though mean LOE did not demonstrate significantly higher evidence for nonoperative facial rejuvenation in the current study, this category showed higher proportions of RCTs when compared to head and neck cancer reconstruction. This finding is unsurprising, in that nonoperative facial rejuvenation techniques such as peels and injectables are more amenable to randomization than are surgical reconstructive techniques. It is worth noting that the proportion of RCTs in the current study is consistent with prior authors, demonstrating that 1–4% of studies within the plastic and aesthetic surgery literature qualify as RCTs [9, 15]. Similarly, when otolaryngology journals are considered, facial plastic and reconstructive surgery demonstrates a lower proportion of RCTs relative to head and neck cancer, otology, rhinology, and others [16].

Despite a higher proportion of RCTs within the facial rejuvenation literature, when studies were reclassified as simply aesthetic or reconstructive, reconstructive procedures demonstrate a higher mean LOE. This may be partly due to the relative strength of the craniofacial literature,

which was considered as reconstructive during this analysis. Prior authors have failed to demonstrate a significant difference between cosmetic and reconstructive literature within the full-body plastic surgery literature [12]. In contrast, of the top 50 most cited (though not necessarily highest LOE) articles within the facial plastic surgery literature, free flap reconstruction and nasal reconstruction are the most common topics [13]. Finally, with regard to individual study design, there were no significant differences among subject matter categories in reference to the presence of confidence intervals or *P* values. Encouragingly, prior studies have demonstrated an increased use of all of these measures over the last 20 years [5, 7, 11].

The current study suffers from multiple limitations that warrant consideration. The selection of journals included in the current study was based on a similar article published prior; however, the discrepant impact factors associated with these journals bias the types and strength of articles submitted to, and accepted by, them [5]. Additionally, the initial article screen was completed by multiple authors, which introduces bias depending on the authors' individual assessment of relevance. Likewise, some articles were not easily classified as solely aesthetic or reconstructive, and it was the judgment of the primary author (ME) to make this determination. Finally, as all articles from the years selected were included and screened, no formal power analysis was performed to determine whether there was sufficient power to detect subtle differences in LOE among subject matter categories. Despite these shortcomings, the current systematic review represents a comprehensive examination of the last decade of the LOE within various subsets of facial plastic surgery literature, as well as the literature as a whole.

Conclusion

While multiple authors have demonstrated that the LOE in facial plastic surgery continues to improve, it remains relatively low at the current time. Craniofacial literature appears to offer a higher mean LOE, relative to nearly all other subsets. Office-based facial rejuvenation techniques appear most amenable to the RCT format for future studies. Finally, reconstructive publications demonstrate significantly more authors per study and higher mean LOE, relative to aesthetic publications. Clinicians within the aesthetic realm stand to benefit from high-level evidence to guide clinical decision making. Office-based rejuvenation offers a significant advantage by way of blinding and randomization. Prospective collaborations between aesthetic surgeons and industry are likely to be paramount in elevating the literature in this segment. Aesthetic surgical topics demonstrate the difficulty in performing randomized

or controlled studies for elective surgical cases. Higher levels of evidence can be achieved through direct cohort comparisons, retrospective if necessary. As the field moves away from case series as the bedrock of its literature, it is likely that cohort studies can fill this space where controlled trials are not feasible. Similarly, the increased adoption of objective cephalometric data and validated patient-reported outcome measures can bolster the quality of evidence within the facial plastic surgery field. Whatever the method, the field can only stand to benefit from the solid footing of an evidence-based approach.

Acknowledgements The authors wish to acknowledge Ashwin Ganti, BA and Eleni Varelas, BA for their contributions to this piece.

Funding None.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest to disclose.

Ethical Approval This study is considered exempt from institutional review board approval due to no involvement of patients or patient data and has been performed in accordance with the ethical standards as laid down in the 1964 Declaration of Helsinki and its later amendments.

Informed Consent For this type of study informed consent is not required.

References

1. Deaton A (2006) Making aid work: evidence-based aid must not become the latest in a long string of development fads. *Boston Rev* 31(4):13
2. Dedhia R, Hsieh TY, Tollefson TT, Ishii LE (2016) Evidence-based medicine in facial plastic surgery: current state and future directions. *Facial Plast Surg Clin N Am* 24:265–274
3. Rohrich RJ, Cho MJ (2020e) Evidence-based medicine in aesthetic surgery: the significance of level to aesthetic surgery. *Plast Reconstr Surg* 139:1195e–1202e
4. Leal DG, Rodrigues MA, Tedesco ACB et al (2018) Evidence-based medicine in plastic surgery: are we there yet? *Ann Plast Surg* 80:71–75
5. Xu CC, Côté DW, Chowdhury RH, Morrissey AT, Ansari K (2011) Trends in level of evidence in facial plastic surgery research. *Plast Reconstr Surg* 127:1499–1504
6. Howick J, Chalmers I, Glasziou P et al (2011) Explanation of the 2011 Oxford Centre for evidence-based medicine (OCEBM) levels of evidence (background document). Oxford Centre for Evidence-Based Medicine. <https://www.cebmnet/index.aspx?o=5653>. Accessed 2 Feb 2020
7. Wasserman J, Wynn R, Bash T, Rosenfeld R (2005) Levels of evidence in otolaryngology journals. *Otolaryngol Head Neck Surg*. <https://doi.org/10.1016/j.otohns.2005.05.190>
8. Joyce KM, Joyce CW, Kelly JC, Kelly JL, Carroll SM (2015) Levels of evidence in the plastic surgery literature: a citation analysis of the top 50 classic papers. *Arch Plast Surg* 42(4):411. <https://doi.org/10.5999/aps.2015.42.4.411>
9. Chang EY, Pannucci CJ, Wilkins EG (2009) Quality of clinical studies in aesthetic surgery journals: a 10-year review. *Aesthet Surg J* 29(2):144–147
10. Chuback JE, Yarascavitch BA, Eaves F, Thoma A, Bhandari M (2012) Evidence in the aesthetic surgical literature over the past decade: how far have we come? *Plast Reconstr Surg* 129(1):126e–134e
11. Loiselle F, Mahabir RC, Harrop AR (2008) Levels of evidence in plastic surgery research over 20 years. *Plast Reconstr Surg* 121(4):207e–e211
12. Nguyen A, Mahabir RC (2016) An update on the level of evidence for plastic surgery research published in plastic and reconstructive surgery. *Plast Reconstr Surg Glob Open* 4(7):e798
13. Chang MT, Schwam ZG, Schutt CA, Kamen EM, Paskhover B (2017) The 50 most cited articles in facial plastic surgery. *Aesthet Plast Surg* 41(5):1202–1207
14. Rifkin WJ, Yang JH, Demitchell-rodriguez E, Kantar RS, Diaz-siso JR, Rodriguez ED (2019) Levels of evidence in plastic surgery research: a 10-year bibliometric analysis of 18,889 publications from four major journals. *Aesthet Surg J* 40:220–227
15. Momeni A, Becker A, Antes G et al (2009) Evidence-based plastic surgery: controlled trials in three plastic surgical journals (1990 to 2005). *Ann Plast Surg* 62(3):293–296
16. Yao FB, Singer M, Rosenfeld RM (2006) R153: randomized controlled trials in otolaryngology journals. *Otolaryngol Head Neck Surg*. <https://doi.org/10.1016/j.otohns.2006.06.908>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.