Published in final edited form as:
J Surg Res. 2019 April ; 236: 22–29. doi:10.1016/j.jss.2018.10.015.

A Systematic Review of Gender-Based Differences in Hirsch Index Among Academic Surgeons

Sara P. Myers, MD, MA, MS[a,*], **Katherine M. Reitz, MD**[a], **Charles B. Wessel, MLS**[b], **Matthew D. Neal, MD**[a], **Jennifer A. Corbelli, MD, MS**[c], **Leslie R.M. Hausmann, PhD**[d,e], and **Matthew R. Rosengart, MD, MPH**[a]

[a]Department of General Surgery, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania

[b]University of Pittsburgh Health Sciences Library System, Pittsburgh, Pennsylvania

[c]Department of Internal Medicine, University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania

[d]Department of Medicine, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania

[e]Veterans Affairs Pittsburgh Healthcare System, Center for Health Equity Research and Promotion, Pittsburgh, Pennsylvania

Abstract

Background—The *h-index* is a commonly used bibliometric in academic medicine which enumerates the number of publications (*h*)that have beencited h times. Recent investigations have suggested that gender-based differences in *h-index* may exist among academic physicians. We systematically reviewed studies of academic surgeons' *h-index*, hypothesizing that a significant difference would exist between the *h-index* of men and women at all academic ranks.

Methods—Peer-reviewed journal articles authored by academic surgeons of any subspecialization in the United States between January 1, 2006, and November 20, 2017, were reviewed. We excluded studies of trainees or gender-based differences in funding without mention of *h-index*. Two reviewers assessed article quality using the Newcastle-Ottawa criteria. Pooled estimates of standard mean differences (SMD) in *h-index* between genders were calculated using randomeffects meta-analyses. A subgroup analysis based on the academic rank was performed. Heterogeneity was assessed using the I^2 statistic. Sensitivity analyses determined the effect of study on *h-index*. Meta-regression identified whether surgical specialty contributed to heterogeneity.

Results—Twelve articles comparing *h-index* between genders were selected from 7950. Men possessed higher *h-indices* than women (SMD, 0.547; $P < 0.001$; $I^2 = 89.5\%$). Men exhibited higher *h-indices* at the assistant rank (SMD, 0.12; 95% confidence interval [CI], 0.01–0.24; $P =$

[*]*Corresponding author.* Department of General Surgery, University of Pittsburgh Medical Center, Pittsburgh, PA. Tel.: (917) 602-9028; fax: +1 412 647 1999. myerssp@upmc.edu (S.P. Myers).

Disclosure
The authors reported no proprietary or commercial interest in any product mentioned or concept discussed in this article.

0.039) but not at the associate (SMD, 0.14; 95% CI, −0.06 to 0.33; $P = 0.165$) or full professor (SMD, 0.12; 95% CI, −0.08 to −0.31; P = 0.25) ranks.

Conclusions—The *h-index* is higher for men than that for women in academic surgery overall but not at individual ranks. Further investigations are necessary to address limitations in h- index and to further characterize the relationship between *h-index*, gender, and promotion.

Keywords

Hirsch index; Gender disparities; Bibliometrics

Introduction

Despite an increasing number of female physicians, surgical specialties continue to be dominated by men.[1] Academic surgery, in particular, has fewer women in leadership positions.[2] Successful promotion in academia involves achievement in clinical, educational, and scholarly domains.[3] Differences in scholarly productivity have been implicated as a contributing factor to gender imbalance in senior academic positions. Research productivity is multidimensional and involves volume and trajectory of publication, contribution of one's scholarship to a field,[4] funding, and academic recognition.[5] Existing metrics for research productivity focus on one or few of these parameters and are not comprehensive. The *h-index*[6] is a commonly used bibliometric in academic medicine[7] which enumerates the number of publications (*h*) that have been cited *h* times for an investigator. Previous studies have suggested gender differences in publication patterns: men have more publications than women early in their careers, but as women advance, the number of publications, and consequently *h-index*, increases over time such that gender-based differences in *h-index* may diminish for professionals late in their career.[8]

Uncertainty persists as to whether there are gender differences in *h-index* within the academic surgery domain. There is disagreement not only with respect to magnitude of difference in *h-index* but also about whether differences exist within each academic rank. Some studies have found that the *h-index* is similar among genders in senior academic positions,[7,9] whereas others have demonstrated that a disparity remains.[8,10] Insofar as *h-index* is used as a consideration for career advancement, we conducted a systematic review of prior studies that have examined gender differences in *h-index* among academic surgeons. A subgroup analysis was performed based on the academic rank. We hypothesized that a statistically significant difference exists between the *h-index* of men and women at all academic ranks.

Methods

Study design

We conducted a systematic review using the meta-analysis of observational studies in epidemiology (MOOSE) guidelines[11–13] to determine whether or not there is a difference in *h-index* among men and women academic surgeons overall and within each academic rank.

Inclusion criteria and outcomes

We included all studies of gender-based differences in *h-indices* of academic surgeons (i.e., surgeons who hold faculty positions at academic medical centers with affiliated surgery residency programs). Secondary outcomes included differences in *h-index* between men and women surgeons according to the academic rank (i.e., assistant professor, associate professor, or full professor) or subspecialization.

Studies that reviewed authorship trends within journals rather than the publications of surgeons themselves were excluded. Randomized control trials and observational trials published within the last 10 y were eligible for inclusion. The literature search was restricted to this time period given that the increase in the proportion of women surgeons by almost 10%[14] may reduce the effect of historical gender imbalance within the field. Investigations taking place at institutions outside of the United States were excluded as cultural differences may affect issues of gender equality within academia. Commentaries, reviews, studies not published in peer- reviewed journals, studies focusing on surgery residents/trainees or medical students as they are ineligible for promotion, and studies involving gender-based differences in funding without mention of *h-index* were also excluded.

Information sources and search strategies

A comprehensive search of PubMed and EMBASE electronic databases was performed on all citations from January 1, 2006, through November 20, 2017 (Appendix A).[15] Only English-language articles were selected. The search strategy was designed by an expert librarian and was reviewed by a coinvestigator (S.P.M.) using a Peer Review of Electronic Search Strategy checklist.[16] Search results from the two databases were merged after duplicate references were discarded. Database searches were conducted on November 20, 2017.

Study selection, data extraction, and data items

Two reviewers (S.P.M. and K.M.R.) screened titles and abstracts of all studies that resulted from the search to determine eligibility for review of full text. Full-text articles were then reviewed independently by both the investigators using data abstraction forms developed to facilitate standardized collection of variables of interest (e.g., author, sample size, study design, statistical analysis, outcomes). All discrepancies and disagreements were resolved through discussion.

Assessment of study quality

Reviewers independently assessed the quality of each study using the modified Newcastle-Ottawa criteria categories of selection, exposure/outcome, and comparability as a guide to detect study limitations.[17] A quality score ranging from low to high was assigned based on sample selection, comparability between academic ranks and gender, attempts to address confounding, and measurement of outcomes. Disagreements with regard to quality scores were resolved by group discussion. Publication bias was assessed by a funnel plot.

Data synthesis and analysis

As *h-index* is a continuous variable, the means and standard deviations for male and female surgeons were abstracted from each study. The reviewed literature determined *h-index* from SCOPUS or Google Scholar. In circumstances where outcomes were reported in alternate form (e.g., median, standard error, interquartile range, and so forth), standardization was performed using statistical methods in accordance with the Cochrane handbook.[18] These values were then used to determine the mean difference between male and female surgeons. Pooled, estimated standardized mean differences in *h-index* between genders were calculated using random-effects meta-analyses.[19]

Potential sources of heterogeneity include type of surgical subspecialty, sample size, and differing rates of surgeons at each academic rank. Heterogeneity of each type was assessed using the I^2 statistic to determine the percentage of variation across studies which could not be explained by chance.[20] Low, moderate, and high heterogeneity was categorized as 25%, 50%, and 75%, respectively.[21] Magnitude of the standard mean difference (SMD) was classified as small, medium, and large at levels of 0.2, 0.5, and 0.8, respectively.[22] A sensitivity analysis was performed by removing studies in a sequential fashion to determine the effect of each study on *h-index*. Meta-regression was performed to determine whether surgical specialty may have contributed to heterogeneity. All statistical analyses were performed using STATA version 14 (StataCorp LP). A *P*-value of less than 0.05 was considered statistically significant.

Results

Included studies

Our search strategy produced 7950 articles, of which 27 articles reported *h-index* among academic surgeons (Fig. 1). After review, 12 articles comparing *h-index* between genders were included.[7,9,23–32] Although seven of these articles specifically cited data on subgroup analysis-based academic rank, only five included *h-index* for faculty at all three levels.

All studies were cross-sectional and included academic surgeons from general surgery and/or orthopedic, plastic surgery, ophthalmology, neurosurgery, colorectal surgery, urology, and gynecologic oncology subspecialties (Table 1). A funnel plot demonstrated symmetry indicating an absence of publication bias (Fig. 2). Assessment of study quality (Table 2) classified three studies as having low quality (1–2 stars for selection and assessment of outcome and no stars for comparability), seven studies to be of moderate quality (2–3 stars) due to ambiguity in selection criteria and adjusting for confounding, and two studies as high quality (4 stars).

Differences in h-index

Eleven studies had sufficient data to report overall mean *h-index* and standard deviation by gender. In our data, SMD values greater than 0 indicated men had higher *h-index* than women. The pooled SMD between genders in overall *h-index* was 0.547 (95% confidence interval [CI], 0.360–0.734; $P < 0.001$; Fig. 3A) with a significant difference between study variation ($I^2 = 89.5\%$). Five studies specifically addressed gender differences in *h-index* at all academic levels. Meta-analysis at the level of assistant professor yielded an SMD of 0.12 (95% CI, 0.01–0.24; $P = 0.039$; Fig. 3B) with 1.6% I^2 ($P = 0.397$). The SMD in *h-index* between genders at the associate professor rank was 0.14 (95% CI, −0.06 to 0.33; $P = 0.165$; Fig. 3C) with ($I^2 = 14.1\%$, $P = 0.324$). There remained no significant difference in *h-index* between men and women at the level of full professor (SMD = 0.12, 95% CI, −0.08 to −0.31; $P = 0.25$; Fig. 3D) with ($I^2 = 0\%$, $P = 0.979$). Summary of findings are presented in Table 2.

Sensitivity analyses

Sensitivity analyses demonstrated that no single study changed the SMD in *h-index* between men and women (Fig. 4). Surgical specialty did not significantly contribute to heterogeneity of the primary outcome based on meta-regression analysis ($P = 0.37$).

Discussion

Research productivity is an important factor in promotion within academic surgery. Although several methods of assessing productivity exist, *h-index* is widely accepted as an objective, easily obtainable metric that incorporates the number of publications and frequency of citation. Our systematic review demonstrates that there is a gender-based difference in overall *h-index* but also that there is no significant difference at the level of associate or full professor.

Data regarding gender disparities in *h-index* among academic surgeons are inconsistent. In two of the studies included in this review, the difference in *h-index* observed between genders overall was no longer present after subgroup analysis based on the academic rank.[7,9] Other investigations indicate that differences in *h-index* are less at higher academic levels, noting that men had higher *h-index* only at the assistant professor level.[8,10] Accordingly, it has been suggested that women may prioritize familial obligations early in their careers (i.e., at a period that coincides with child-bearing) but also that, subsequently, publication rates and *h-index* between men and women converge. The larger sample size afforded by this review supports this interpretation as there was no difference in *h-indices* observed for women relative to men at associate or full professor levels. Gender-based differences in attrition at the level of junior faculty may lead to inaccurate interpretations of trends in *h-index* as academics advance in their careers; we cannot assume that female surgeons who remain to advance to more senior positions did not have publication rates and/or *h-indices* similar to their male colleagues even early in their careers.

Although our study calls attention to gender-based differences in *h-index*, further investigations are needed to determine the etiology and impact of these findings and how *h-index* at senior academic ranks reflects the type of academic contributions (i.e., scholarly productivity, clinical performance, and education) made by men and women. Gender-based differences in other areas of an academic's tenure portfolio may account for professional advancement despite a lower *h-index*. For example, recent evidence suggests that female surgeons may perform better on clinical metrics (e.g., postoperative outcomes) than male surgeons.[33,34] Scholarly productivity may be universally important in career advancement, but institutions vary with respect to their tenure criteria and the degree to which they value other parameters such as clinical performance and pedagogy. In the studies reviewed, there was a wide range in the reported values for *h-index*, highlighting the ambiguity in determining an average *h-index* that might be expected at each faculty rank or a value that might be required for promotion.

The studies considered in this review illustrate disadvantages in using *h-index* as a sole surrogate for scholarly productivity. Methods for resolving inconsistencies in abstraction, gender identification, or academic title were rarely discussed in the reviewed manuscripts. A change in surname secondary to marriage or divorce was frequently cited as a limitation that could underestimate *h-index*. The h- index does not reflect the quality of the publications themselves or the degree to which an individual contributed to each publication. Studies did not systematically adjust for the fact that *h-index* fails to distinguish based on the impact-factor of the journal within which articles are published, classification or quality of article type (e.g., case-review, original study, and so forth), or the position of a particular author within author sequence. Although there does not appear to be a difference between genders in co-first authorship in basic science journals, there is evidence to suggest that women co-first authors of clinical journal articles are less likely to be listed as first authors in the byline.[35] In addition, men may have an inflated *h-index* because their self-citation rates are higher.[36] Although *h-index* is associated with academic standing,[37,38] these issues reflect challenges that arise by using this surrogate measure rather than a direct assessment of quality of scholarly endeavors.

Limitations

This systematic review should be considered in the context of several limitations. The considerable heterogeneity with respect to overall differences between men and women challenges the accuracy of these findings. Several surgical subspecialties were not included in the articles reviewed in this study. Published articles that did not include data in a form that could be converted to SMD for statistical evaluation were excluded. Finally, as the studies considered were all performed in the United States, these data cannot be generalized to academic surgeons practicing in other countries.

Conclusion

Despite its drawbacks, *h-index* remains an instrument for assessing research productivity. Although this review cannot establish causality, it does indicate that differences in *h-index* exist between genders overall but also that this difference does not persist at higher academic ranks. Further investigations are necessary to address limitations in *h-index* and to further characterize the relationship between *h-index*, gender, and promotion.

Acknowledgment

Author contributions: S.P.M. was primarily responsible for conceptualization of this study, data analysis, and drafting of this manuscript. K.M.R. aided in data abstraction, analysis, and manuscript review. C.B.W. was responsible for the literature search and assisted in manuscript development and review. J.A.C. and M.D.N. provided assistance with data analysis and acquisition, as well as with critical review of this manuscript. M.R.R. and L.R.M.H. aided in study design, data interpretation, drafting of the manuscript, and critical review.

REFERENCES

1. Ahmed AA, Egleston B, Holliday E, Eastwick G, Takita C, Jagsi R. Gender trends in radiation oncology in the United States: a 30-year analysis. Int J Radiat Oncol Biol Phys. 2014;88:33–38. [PubMed: 24189127]
2. Weiss A, Lee KC, Tapia V, et al. Equity in surgical leadership for women: more work to do. Am J Surg. 2014;208:494–498. [PubMed: 24933664]
3. Baldwin C, Chandran L, Gusic M. Guidelines for evaluating the educational performance of medical school faculty: priming a national conversation. Teach Learn Med. 2011;23:285–297. [PubMed: 21745065]
4. Beasley BW, Wright SM, Cofrancesco J Jr, Babbott SF, Thomas PA, Bass EB. Promotion criteria for clinician- educators in the United States and Canada. A survey of promotion committee chairpersons. JAMA. 1997;278:723–728. [PubMed: 9286831]
5. Atasoylu AA, Wright SM, Beasley BW, et al. Promotion criteria for clinican-educators. J Gen Intern Med. 2003;18:711–716. [PubMed: 12950479]
6. Hirsch JE. An index to quantify an individual's scientific research output. Proc Natl Acad Sci USA. 2005;102:16569–16572. [PubMed: 16275915]
7. Tomei KL, Nahass MM, Husain Q, et al. A gender-based comparison of academic rank and scholarly productivity in academic neurological surgery. J Clin Neurosci. 2014;21:1102–1105. [PubMed: 24411320]
8. Reed DA, Enders F, Lindor R, et al. Gender differences in academic productivity and leadership appointments of physicians throughout academic careers. Acad Med. 2011;86:43–47. [PubMed: 21099390]
9. Mayer EN, Lenherr SM, Hanson HA, Jessop TC, Lowrance WT. Gender differences in publication productivity among academic urologists in the United States. Urology. 2017;103:39–46. [PubMed: 28232174]
10. Eloy JA, Svider P, Chandresekhar SS, et al. Gender disparities in scholarly productivity within academic otolaryngology departments. Otolaryngol Head Neck Surg. 2013;148:215–222. [PubMed: 23161882]
11. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis of Observational Studies in Epidemiology (MOOSE) group. JAMA. 2000;283:2008–2012. [PubMed: 10789670]
12. Lau J, Ioannidis JP, Schmid CH. Summing up evi-dence. Lancet. 1998;351:123–127. [PubMed: 9439507]
13. Huston P Health services research. CMAJ. 1996;155:1697–1702. [PubMed: 8976336]
14. Wolfe L Statistics on the number of women surgeons in the United States. Available at: https://www.thebalance.com/number-of-women-surgeons-in-the-us-3972900. Accessed February 18, 2018.
15. Cocanour CS, Chang P, Huston Jared M, et al. Management and novel adjuncts of necrotizing soft tissue infections. Surg Infect (Larchmt). 2017;18:250–272. [PubMed: 28375805]
16. Sampson M, McGowan J, Cogo E, Grimshaw J, Moher D, Lefebvre C. An evidence-based practice guideline for the peer review of electronic search strategies. J Clin Epidemiol. 2009;62:944–952. [PubMed: 19230612]
17. Wells GA, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of non-randomised studies in meta-analyses. Available at: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp. Accessed November 20, 2017.
18. Cumpston M Analysing continuous outcomes. Available at: http://training.cochrane.org/resource/analysing-continuous-outcomes. Accessed December 1, 2017.
19. Khoushhal Z, Hussain MA, Greco E, et al. Prevalence and causes of attrition among surgical ResidentsA systematic review and meta-analysis. JAMA Surg. 2017;152:265–272. [PubMed: 27973673]
20. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002;21:1539–1558. [PubMed: 12111919]

21. Melson WG, Bootsma MCJ, Rovers MM, Bonten MJM. The effects of clinical and statistical heterogeneity on the predictive values of results from meta-analyses. Clin Microbiol Infect. 2014;20:123–129. [PubMed: 24320992]
22. Faraone SV. Interpreting estimates of treatment effects: implications for managed care. Pharm Ther. 2008;33:700–703.
23. Mueller C, Wright R, Girod S. The publication gender gap in US academic surgery. BMC Surg. 2017;17:16. [PubMed: 28193221]
24. Paik AM, Mad IJ, Villaneuva NL, et al. Research productivity and gender disparities: a look at academic plastic surgery. J Surg Educ. 2014;71:593–600. [PubMed: 24776868]
25. Martinez M, Lopez S, Beebe K. Gender comparison of scholarly production in musculoskeletal tumor society using the Hirsch index. J Surg Educ. 2015;72:1172–1178. [PubMed: 26232946]
26. Lopez SA, Svider PF, Misra P, Bhagat N, Langer PD, Eloy JA. Gender differences in promotion and scholarly impact: an analysis of 1460 academic opthalmologists. J Surg Educ. 2014;71:851–859. [PubMed: 24852601]
27. Hill EK, Blake RA, Emerson JB, et al. Gender differences in scholarly productivity within academic gynecologic oncology departments. Obstet Gynecol. 2015;126: 1279–1284. [PubMed: 26551177]
28. Ence AK, Cope SR, Holliday EB, Somerson JS. Publication productivity and experience: factors associated with academic rank among orthopaedic surgery faculty in the United States. J Bone Joint Surg Am. 2016;98:e41. [PubMed: 27194503]
29. Theratti PJ, Hoppe IC, Granick MS, Lee ES. Application of the h-index in academic plastic surgery. Ann Plast Surg. 2016;76:545–549. [PubMed: 25422983]
30. Khan NR, Thompson CJ, Taylor DR, et al. An analysis of publication productivity for 1225 academic neurosurgeons and 99 departments in the United States. J Neurosurg. 2014;120:746–755. [PubMed: 24359012]
31. Bastian S, Ippolito JA, Lopez SA, Eloy JA, Beebe KS. The use of h-index in academic orthopedic surgery. J Bone Joint Surg Am. 2017;99:e14. [PubMed: 28196042]
32. Geltzeiler CB, Kelley KA, Srikanth P, et al. Does sex influence publication productivity among colorectal surgeons participating in fellowship training programs? Dis Colon Rectum. 2017;60:537–543. [PubMed: 28383454]
33. Wallis CJD, Ravi B, Coburn N, Nam RK, Detsky AS, Satkunasivam R. Comparison of postoperative outcomes among patients treated by male and female surgeons: a population based matched cohort study. BMJ. 2017;359:j4366. [PubMed: 29018008]
34. Toubassi D Re: comparison of postoperative outcomes among patients treated by male and female surgeons: a population based matched cohort study. BMJ. 2017;359:j4366. [PubMed: 29018008]
35. Aakhus E, Mitra N, Lautenbach E, et al. Gender and byline placement of co-first authors in clinical and basic science journals with high impact factors. JAMA. 2018;319:610–611. [PubMed: 29450515]
36. Chawla DS. Self-citation rates higher for men. Nature. 2016;535:212.
37. Pagel PS, Hudetz JA. H-index is a sensitive indicator of academic activity in highly productive anaesthesiologists: results of a bibliometric analysis. Acta Anaesthesiol Scand. 2011;55:1085–1089. [PubMed: 22092205]
38. Hirsch JE. Does the H index have predictive power? Proc Natl Acad Sci USA. 2007;204:19193–19198.

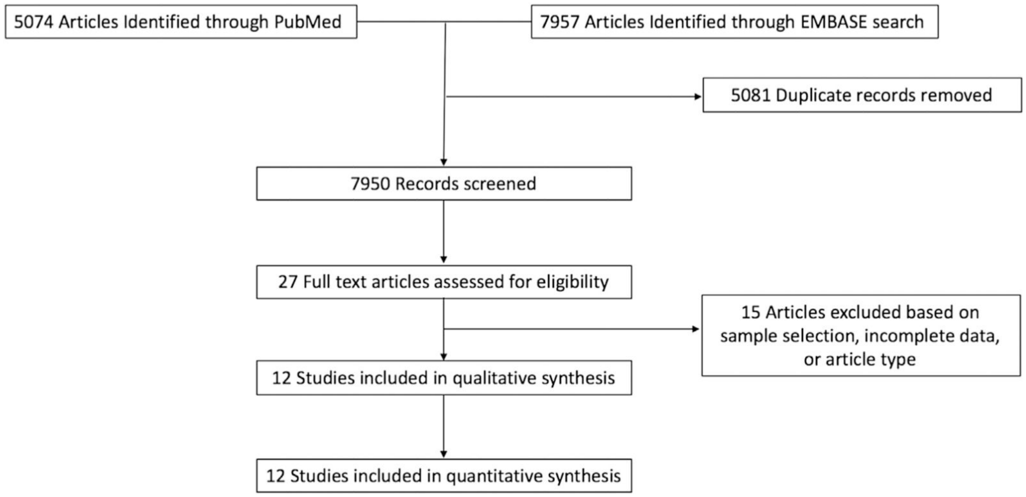

Fig. 1 –.

CONSORT flow diagram of study inclusion.

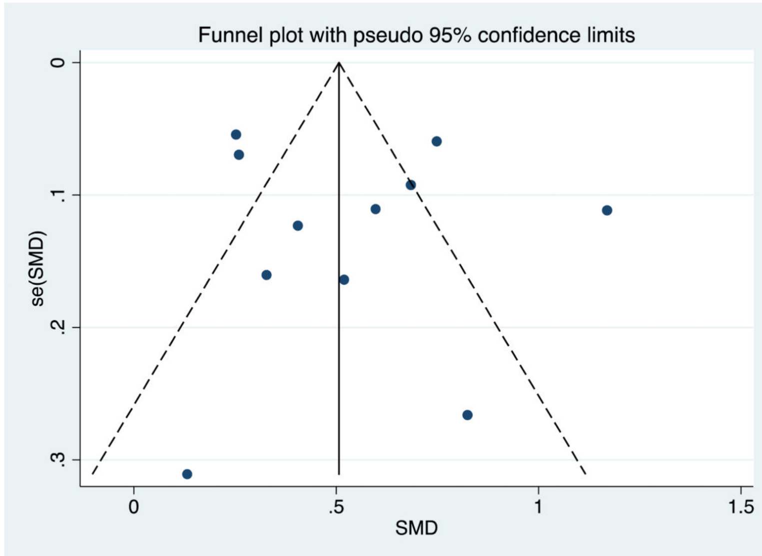

Fig. 2 –.
Assessment of publication bias. (Color version of figure is available online.)

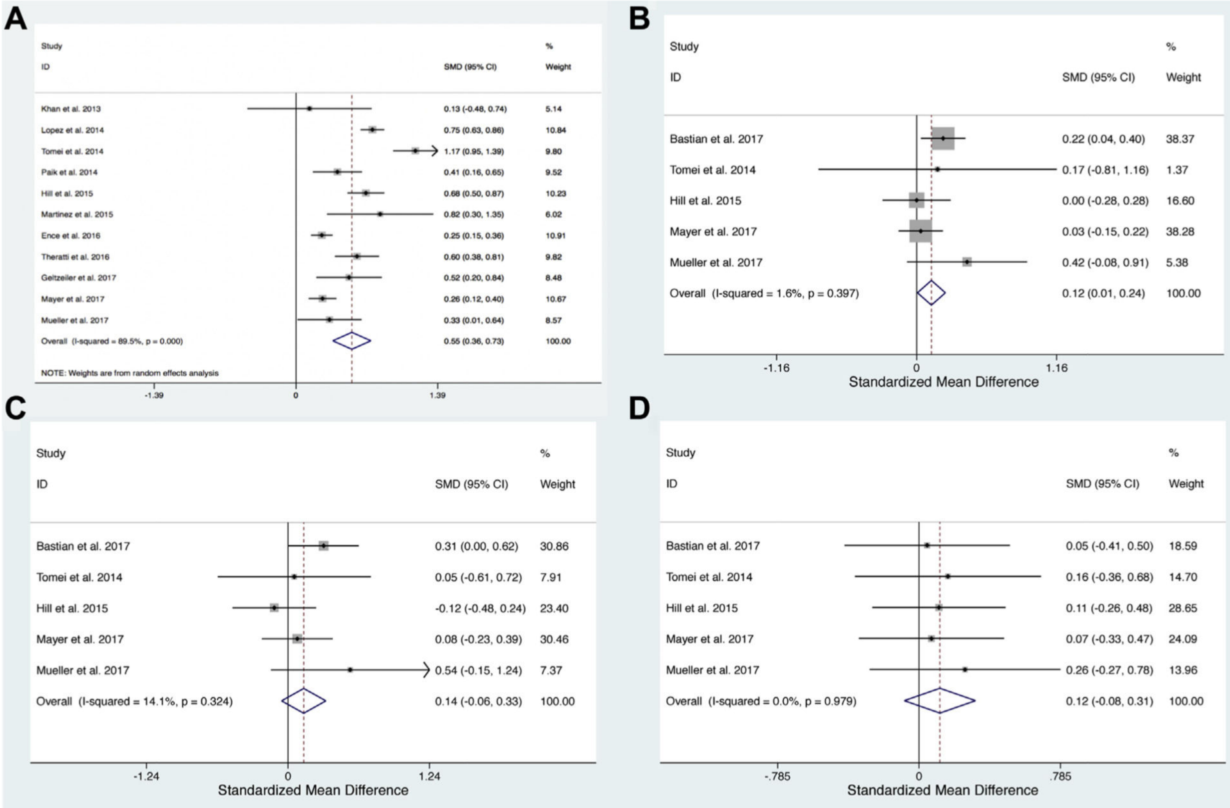

Fig. 3 –.

Differences in *h-index* between men and women academic surgeons (A) overall, at the (B) assistant professor rank, (C) associate professor rank, and (D) full professor rank. SMD greater than 0 indicates women with lower *h-index*. (Color version of figure is available online.)

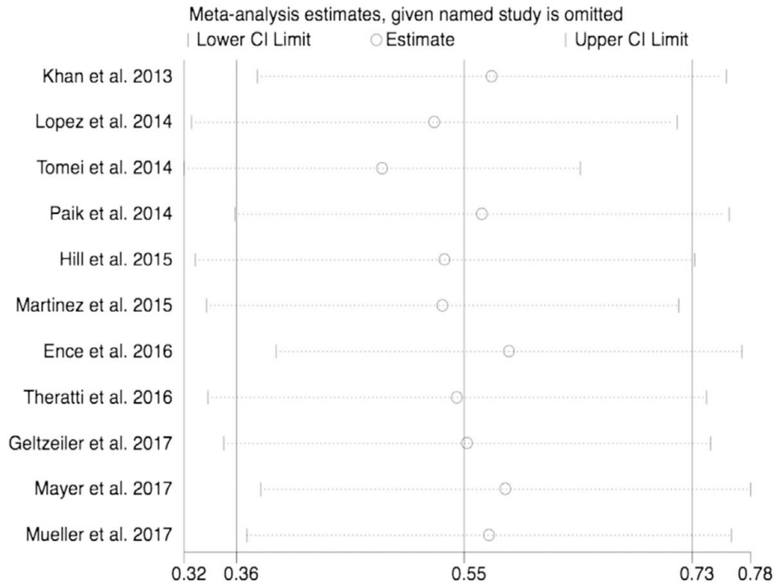

Fig. 4 –.

Sensitivity analysis for overall *h-index*.

Table 1 –

Standard mean difference (SMD) in *h-index* between men (M) and women (F) in included studies.

Study	*h-index* by academic rank			
	Overall SMD (95% CI), *n* (M; F)	**Assistant professor SMD (95% CI), *n* (M; F)**	**Associate professor SMD (95% CI) *n* (M; F)**	**Full professor SMD (95% CI) *n* (M; F)**
Khan *et al.* 2013 (neurosurgery)	0.13 (0.48–0.75), 188 (177; 11)	—	—	—
Lopez *et al.* 2014 (opthalmology)	0.75 (0.63–0.86), 1460 (1041; 419)	—	—	—
Tomei *et al.* 2014 (neurosurgery)	1.17 (0.95–1.39), 1052 (959; 93)	0.39 (0.09 to 0.58), 415 (366; 49)	0.137 (−0.253 to 0.526), 233 (204; 29)	0.160 (−0.355 to 0.676), 404 (389; 15)
Paik *et al.* 2014 (plastic surgery)	0.41 (0.16–0.65), 505 (426; 79)	—	—	0.625 (−0.203 to 1.453), 107 (101; 6)
Hill *et al.* 2015 (gynecologic oncology)	0.68 (0.50–0.87), 507 (292; 215)	0.42 (−0.13 to 0.13), 208 (77; 131)	−0.120 (−0.485 to 0.244), 120 (71; 49)	0.111 (−0.259 to 0.480), 179 (144; 35)
Martinez *et al.* 2015 (orthopedic surgery)	0.82 (0.30–1.35), 125 (108; 17)	—	—	—
Ence *et al.* 2016 (orthopedic surgery)	0.25 (0.15–0.36), 3509 (3128; 381)	0.00 (−0.13 to 0.13), 2036 (1773; 263)	—	—
Theratti *et al.* 2016 (plastic surgery)	0.60 (0.38–0.81), 592 (491; 101)	—	—	—
Geltzeiler *et al.* 2017 (colorectal surgery)	0.52 (0.20–0.84), 220 (171; 49)	—	—	—
Mayer *et al.* 2017 (urology)	0.26 (0.12–0.40), 1922 1686; 236)	0.19 (0.01 to 0.38), 744 (603; 141)	0.369 (0.058 to 0.679), 361 (315; 46)	0.387 (−0.017 to 0.791), 66 (41; 25)
Mueller *et al.* 2017 (multiple)	0.30 (0.01–0.64), 212 (160; 52)	0.42 (−0.08 to 0.92), 74 (51; 23)	0.545 (−0.148 to 1.238), 55 (45; 10)	0.255 (−0.274 to 0.785), 77 (59; 18)
Bastian *et al.* 2017 (orthopedic surgery)	—	0.23 (0.08 to 0.37), 976 (843; 133)	0.311 (0.004 to 0.618), 504 (459; 45)	0.045 (−0.414 to 0.504), 461 (442; 19)

In the studies mentioned in this table, men have higher *h-index*, which is reflected in a positive SMD.

Table 2 –

Summary of findings.

h-index	Standard mean difference (95% CI)	Number of participants (number of studies)	Comments
Overall	0.547 (0.360 to 0.734)	8369 (11)	All academic ranks
Assistant professor	0.12 (0.01 to 0.24)	2372 (5)	N/A
Associate professor	0.14 (−0.06 to 0.33)	920 (5)	N/A
Full professor	0.12, (−0.08 to 0.31)	1606 (5)	N/A

Study	Selection	Assessment of outcome	Comparability	Overall quality assessment
Khan *et al.* 2013		*		Low
Lopez *et al.* 2014	*	*	**	Moderate
Tomei *et al.* 2014	*	*	**	High
Paik *et al.* 2014	*	*		Low
Hill *et al.* 2015		*	**	Moderate
Martinez *et al.* 2015	*	*	**	High
Ence *et al.* 2016		*	**	Moderate
Theratti *et al.* 2016	*	*	*	Moderate
Geltzeiler *et al.* 2017	*	*	*	Moderate
Mayer *et al.* 2017	*	*	*	Moderate
Mueller *et al.* 2017		*	*	Moderate
Bastian *et al.* 2017	*	*		Low

Selection, assessment of outcome, and comparability are parameters considered as part of the quality assessment. Each parameter can receive 0–2 asterisks depending on degree to which the parameter was addressed.

Population: Surgeons.

Intervention: Research productivity as measured by Hirsch index (h-index).

Comparison: Men and Women.

Setting: Academic institutions (i.e., affiliated with residency, fellowship, and medical schools).