OPEN ACCESS

Check for updates

Gender differences in how scientists present the importance of their research: observational study

Marc J Lerchenmueller,[1,2] Olav Sorenson,[2] Anupam B Jena[3]

[1]University of Mannheim, Mannheim, Germany
[2]Yale University, School of Management, New Haven, CT, USA
[3]Department of Health Care Policy, Harvard Medical School, Boston, MA, USA; Massachusetts General Hospital, Boston, MA, USA; and National Bureau of Economic Research, Cambridge, MA, USA

Correspondence to:
M J Lerchenmueller
marc.lerchenmueller@uni-mannheim.de
(or @MLerchenmueller on Twitter;
ORCID 0000-0001-8875-3602)

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2019;367:l6573
http://dx.doi.org/10.1136/bmj.l6573

Accepted: 11 November 2019

ABSTRACT

OBJECTIVES

Women remain underrepresented on faculties of medicine and the life sciences more broadly. Whether gender differences in self presentation of clinical research exist and may contribute to this gender gap has been challenging to explore empirically. The objective of this study was to analyze whether men and women differ in how positively they frame their research findings and to analyze whether the positive framing of research is associated with higher downstream citations.

DESIGN

Retrospective observational study.

DATA SOURCES

Titles and abstracts from 101 720 clinical research articles and approximately 6.2 million general life science articles indexed in PubMed and published between 2002 and 2017.

MAIN OUTCOME MEASURES

Analysis of article titles and abstracts to determine whether men and women differ in how positively they present their research through use of terms such as "novel" or "excellent." For a set of 25 positive terms, we estimated the relative probability of positive framing as a function of the gender composition of the first and last authors, adjusting for scientific journal, year of publication, journal impact, and scientific field.

RESULTS

Articles in which both the first and last author were women used at least one of the 25 positive terms in 10.9% of titles or abstracts versus 12.2% for articles involving a male first or last author, corresponding to a 12.3% relative difference (95% CI 5.7% to 18.9%). Gender differences in positive presentation were greatest in high impact clinical journals (impact factor >10), in which women were 21.4% less likely to present research positively. Across all clinical journals, positive presentation was associated with 9.4% (6.6% to 12.2%) higher subsequent citations, and in high impact clinical journals 13.0% (9.5% to 16.5%) higher citations. Results were similar when broadened to general life science articles published in journals indexed by PubMed, suggesting that gender differences in positive word use generalize to broader samples.

CONCLUSIONS

Clinical articles involving a male first or last author were more likely to present research findings positively in titles and abstracts compared with articles in which both the first and last author were women, particularly in the highest impact journals. Positive presentation of research findings was associated with higher downstream citations.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Women remain underrepresented on faculties of medicine and the life sciences. Women also earn lower salaries, receive fewer research grants, and receive fewer citations than their male colleagues

One mechanism that may contribute to these gender gaps is differences in the extent to which women promote their research accomplishments relative to men. Yet, systematic evidence of differences in how men and women present their research findings in the academic life sciences is lacking. Identifying gender differences in how research is self presented is potentially important given that visible research productivity is central to career progress in the academic life sciences and medicine, affecting hiring, promotion, pay, and funding decisions

WHAT THIS STUDY ADDS

This study investigates gender differences in the positive framing of research findings (that is, use of words such as "novel," "excellent," etc), analyzing 101 720 research articles published between 2002 and 2017 in clinical journals indexed in PubMed, as well as over 6.2 million general life science articles

Articles in which the first and last authors were both women were, on average, 12.3% less likely to use positive terms to describe research findings compared with articles in which the first and/or last author was male. The gender difference in positive presentation was greatest in high impact clinical journals, with women being 21.4% less likely to present research positively. Positive presentation was, on average, associated with 9.4% higher subsequent citations and 13.0% higher citations in high impact clinical journals. These results remained consistent even after accounting for the specific journal, journal impact factor, scientific area of study, and year of publication

Clinical articles involving at least one male first or last author were more likely to present research findings positively in titles and abstracts compared with articles in which both the first and last authors were women, particularly in the highest impact journals. Positive presentation of research findings was associated with higher downstream citations

Introduction

Women remain underrepresented in academic medicine and the life sciences more broadly.[1] Even the most recent surveys indicate that the proportion of women declines at every career step, including promotion to full professorship.[2] Women also earn lower salaries,[3 4] receive fewer research grants,[5] and receive fewer citations than their male colleagues.[6]

One mechanism that may contribute to these gender gaps is differences in the extent to which women promote their research accomplishments relative to men.[7] Yet, systematic evidence of differences in how men and women present their research findings in the academic life sciences is lacking. Identifying gender differences in how research is self-presented is potentially important given that visible research productivity is central to career progress in the academic life sciences and medicine, affecting hiring, promotion, pay, and funding decisions.[8 9]

BMJ: first published as 10.1136/bmj.l6573 on 16 December 2019. Downloaded from http://www.bmj.com/ on 17 December 2019 by guest. Protected by copyright.

We studied gender differences in the self presentation of scientific research—as identified by authors' use of terms such as "novel," "unique," "unprecedented," etc to describe their research in scientific titles and abstracts—among 101 720 clinical research articles published between 2002 and 2017 and indexed in PubMed. We examined how gender differences in positive presentation evolved over 15 years and varied with journal impact. We also examined the external validity of our findings in a dataset of roughly 6.2 million general life science articles published over the same time span in PubMed-indexed journals.

Methods

Data sources

We assembled data from several sources. First, we extracted information from the PubMed database, which represents the most comprehensive reference list of scientific publications in the life sciences. Second, we determined the probable gender of the authors through the Genderize database, an established approach[8 10 11] that allows gender assignment for a large number of authors based on their first names. Third, we used information from the *Journal Citation Report* (Clarivate Analytics) to determine the impact factor of the publishing journals as well as the MeSH (Medical Subject Headings) library to categorize articles into fields of research.

Assignment of author gender

We assigned author gender to the first and last (senior) authors of each article in our database. We focused on these authorship positions because in the life sciences these two authorship positions are typically most important. First authorship often goes to a junior scientist who executed and wrote up the research, while the senior investigator who may have conceived of and/or funded the research receives last, or senior, authorship. Interior authors generally serve in supporting roles.

We determined the gender of these individuals by associating the first names of the authors with the probability of the name being held by a man versus by a woman, using the Genderize database. Tests of the accuracy and comprehensiveness of four gender assignment algorithms, using a control sample of gender-matched forenames from a US government office, found that the Genderize algorithm provided the most accurate estimates of gender.[12] The Genderize algorithm allowed us to assign the probable gender of first and last authors for over 90% of the articles in our database (for additional details see appendix on bmj.com). Applying a 90% probability threshold to the algorithm's gender designation yielded an average probability of 99.5% with which gender could be assigned in our dataset for analysis.

Outcomes

To determine whether men and women differed in the positive presentation of their research, we focused on the use of a set of 25 words (see fig 2) that past research has identified as distinctively positive[13]—such as "unprecedented" and "novel"—in the titles and abstracts of clinical research articles. Our primary outcome was a binary variable for whether any of these words were used in a scientific abstract or title. In a secondary analysis, we explored gender differences in the use of each specific word. Titles and abstracts represent some of the most important text in research papers because readers often use this information to screen articles to determine which ones deserve further attention.

Covariates

To account for the possibility that men and women may differ in the positive presentation of research due to the journals or fields in which they tend to publish, we adjusted for these factors using a set of granular covariates. Covariates included indicator variables for the year of publication and the specific scientific journal (based on unique journal International Standard Serial Number (ISSN) numbers)—to account, for example, for the possibility that author gender and the positive presentation of research findings vary across journals or over time. The covariates also included features of the author byline (number of authors and percentage female authors).

We also adjusted for the frequency with which positive terms appear in titles and abstracts in certain research areas. Specifically, we used the article keywords (more than 26 000 MeSH terms) to adjust for the areas of research being reported by creating a centile ranking of these keywords in terms of their frequency of being associated with the 25 positive words. Since many articles are associated with more than one MeSH term, we calculated the median centile ranking across all MeSH terms associated with a given article and included this measure as a covariate in our analyses.

Statistical analyses

We estimated the probability that an article used one or more positive words in its title or abstract as a function of the gender composition of the first and last authors, using multivariable linear probability models. We categorized authorship of articles into two categories: articles with female first and last authors versus articles where a man served as first and/or last author, which served as the baseline. We estimated the probability that an article used one or more positive words in its title or abstract as a function of this indicator variable for gender composition, adjusting for journal indicator variables, indicator variables for year of publication, number of authors and percentage of female authors in the author byline, and the extent of use of positive words in the article's area of study (centile rank for MeSH category). The inclusion of journal fixed effects accounted for factors that do not vary within a journal over time and that might be correlated with both the prevalence of positive terms in abstracts or titles of a given journal and the gender composition of first and last authors. Our approach

BMJ: first published as 10.1136/bmj.l6573 on 16 December 2019. Downloaded from http://www.bmj.com/ on 17 December 2019 by guest. Protected by copyright.

doi: 10.1136/bmj.l6573 | *BMJ* 2019;367:l6573 | the bmj

therefore effectively compared gender differences in the positive presentation of research between articles published within the same journal, year, and field of research.

We next analyzed whether gender differences in the positive presentation of clinical research varied with the impact of the journal. We classified clinical journals with an impact factor of up to 10 versus those exceeding 10 (supplementary table S1 in appendix). The former category includes the journals ranked 25 and lower by journal impact factor in the 2017 Journal Citation Report (Clarivate Analytics); the latter category represents the top 24 clinical journals. To analyze the generalizability of our findings, we conducted a similar analysis extended to a broader set of life science journals indexed in PubMed. We also analyzed how the positive presentation of research findings varied over the 15-year sample window across journal impact categories. To further contextualize the use of the positive words, we determined the frequency with which each individual word occurred in different sections of the structured abstracts of clinical research articles, broadly classified as background, methods, results, and conclusions.

Finally, using article citation data, we estimated the relation between the positive presentation of research findings and the downstream citations received by articles. Gender differences in the positive presentation of research findings may be particularly important if the positive presentation of findings leads to more downstream citations. Citations are often used to gauge a researcher's influence and many organizations use cumulative citations explicitly in their decisions regarding recruitment, promotion, pay, and funding. We estimated a multivariable linear regression of the association between (logged) downstream article citations and whether an article positively presented its research findings, adjusting for the covariates described above.

Sensitivity analyses

We conducted several sensitivity analyses. First, we estimated Poisson regressions of the frequency with which positive words are used as a function of author gender composition and the previously described covariates. Second, we used Natural Language Processing to further characterize the context in which positive words appeared, extracting the most common nouns that immediately followed the 25 positive words (adjectives). We conducted this analysis to understand better whether the use of terms like "novel" might reflect subjective interpretations of research findings, such as the couplet "novel findings" would suggest. Third, to account for the possibility that men and women might write abstract of different lengths, we included a covariate for abstract length. Fourth, we replicated our analysis in samples beyond research articles, including, for example, letters and comments. Fifth, to test the accuracy of our gender designation method, we examined variation in the reliability with which Genderize designated gender for US-based authors of different ethnic backgrounds, including Asian, black, Hispanic, and white authors. Sixth, we replicated our results for different gender designation probability thresholds (50%, 80%, 90%, and 95%). Seventh, we estimated the probability of positive framing for different gender compositions of first and last authors (male-male, male-female, female-male, and female-female). Eighth, we estimated both multivariable least squares regressions and Poisson regressions to estimate the effect of positive presentation on downstream citations using the Scopus database as a secondary source for citation data. Ninth, we replicated the analyses with a set of negative words.[13] Tenth, we examined whether gender differences existed in the use of positive words in a sample of introductions of analyzed articles to gauge whether language use might differ between titles and abstracts and the main body. Additional details on these analyses appear in the appendix.

Patient and public involvement

No patients were involved in setting the research question or the outcome measures, nor were they involved in developing plans for design or implementation of the study. No patients were asked to advise on the interpretation or write up of the results. There are no plans to disseminate the results of the research to study participants or the relevant patient community.

Results

Overall, 101 720 clinical research articles and 6 246 112 general life sciences articles, published between 2002 and 2017, were analyzed. Among clinical articles, 12.1% (12 309/101 720) used at least one of the 25 positive words to frame research findings. In comparison, 11.7% (729 619/6 246 112) of all life science articles indexed in PubMed used at least one of the 25 positive words (table 1). Overall, 17% of clinical research articles involved a female first and last author, whereas 83% of articles involved a male first or last author. Positive presentation of research findings increased significantly over the study period (fig 1). For example, among clinical journals with an impact factor exceeding 10, the prevalence of positive framing rose by over 80% from 2002 to 2017 (from 9.0% to 16.9% of articles).

Gender differences in positive presentation

Figure 2 ranks individual positive words in descending order in terms of prevalence and compares the proportion of articles using each word in which the first and last author were both women with the proportion of articles in which the first and/or last author was a man. The most commonly used positive word was "novel" (used in 44.8% of positively framed articles). Articles with a male first and/or last author used the word 59.2% (95% confidence interval 48.8% to 69.6%) more often than articles with female first and last authors. For nine of the 10 most commonly used positive words, gender differences were statistically significant at the 5% level (for example, "unique"

BMJ: first published as 10.1136/bmj.l6573 on 16 December 2019. Downloaded from http://www.bmj.com/ on 17 December 2019 by guest. Protected by copyright.

Table 1 | Characteristics of analyzed articles

Characteristic	Clinical journals (n=101 720)	PubMed indexed journals (n=6 246 112)
Sex composition of lead authors:		
Female first and last author	17%	16%
Female first and male last author	16%	23%
Male first and female last author	9%	11%
Male first and last author	57%	51%
Percentage female authors on byline	37%	35%
Mean No of authors per paper	7.2	4.8
Median publication year	2010	2011
Journal impact factor (JIF) tier:		
Paper published in journal with JIF >10	61%	5%
Paper published in journal with JIF ≤10	39%	95%
Median centile for positively presented research topics based on MeSH terms per paper*	43.8	47.7
No (%) of papers using positive words	12 309 (12%)	729 619 (12%)

*Based on 96 543 clinical articles and 5 046 315 life science articles with associated MeSH terms. Percentages on the sex composition of lead authors do not sum to 100% due to rounding. Clinical journals included in the analysis and their associated journal impact factor (JIF) are listed in the appendix on bmj.com.

appearing in 13.3% of positively framed articles and being used 43.8% (25.9% to 61.7%) more often by teams with male authors, and "promising" appearing in 12.5% of positively framed articles and being used 72.3% (50.6% to 94.0%) more often by teams with a male first and/or last author. (See supplementary table S26 in appendix for further details on the 10 most common words.)

We next made use of the fact that most (81%) abstracts of clinical research articles are structured, broadly classifying content into categories of background, methods, results, and conclusions. Figure 3 depicts the frequency with which the top 10 positive words occurred across these structured categories, with deeper shading indicating higher frequency. The top 10 positive words were used in 73% of cases, on average, to describe results and conclusions. That use did not differ across men and women. Considering the descriptive results from figures 2 and 3 together, the data suggest that women and men use positive words in a similar fashion (fig 3) but that women use them less often (fig 2).

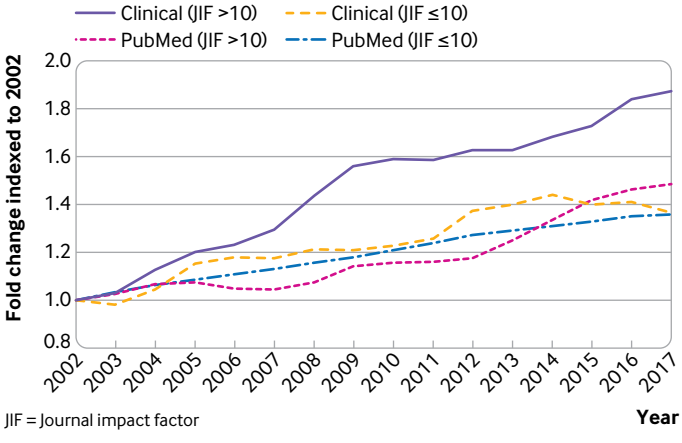

Fig 1 | Change in use of positive terms to present research findings in articles published between 2002 and 2017. Proportion of positively presented articles was calculated as a three-year rolling average and indexed to the proportion of positively presented articles in 2002 for the respective journal categories. Figure presents trends for research articles in clinical journals and broader life science articles indexed in PubMed.

After multivariable adjustment for the specific journal in which an article appeared, year of publication, journal impact factor, and extent of use of positive words in the article's area of study, clinical research articles in which the first and last author were both women were less likely to use positive terms to report research findings compared with articles in which the first and/or last author was a man (10.9% *v* 12.2% of articles, corresponding to a 12.3% relative difference (95% CI 5.7% to 18.9%); see table S11 in the appendix for full regression results). For broader life science research indexed in PubMed, positive modifiers were used to report research in 11.4% of articles in which both the first and last author were women compared with 11.5% of articles in which the first and/or last author was a man, corresponding to a non-significant relative difference of less than 1% (−0.3% to 1.5%). Importantly, however, the samples differed in the share of journals with an impact factor exceeding 10 (61% of articles in clinical research articles *v* 5% of articles in the broader life sciences (table 1)).

Gender differences in positive presentation were largest among articles published in journals with high impact factors (fig 4), for both clinical and broader life science research and even after multivariable adjustment. While the gender difference in positive presentation was small and not statistically significant in journals with an impact factor below 10, among articles published in clinical journals with an impact factor exceeding 10, those that involved a male first and/or last author were significantly more likely to present research positively compared with articles in which both first and last authors were women (absolute difference 10.7% *v* 12.9%, relative difference 21.4% (12.3% to 30.5%); see also table S11 in the appendix). We observed a similar pattern for high impact journals for the broader set of life science articles indexed in PubMed. Articles published in journals with an impact factor of more than 10 with a male first and/or last author included positive words to describe the research more frequently than those with female first and last authors (absolute difference 14.4% *v* 12.8%, relative difference 12.8% (8.2% to 17.4%)).

Positive presentation of research findings and subsequent article citations

After multivariable adjustment, the positive presentation of research findings was associated with greater downstream citations for both clinical research articles and for life sciences articles more broadly (fig 5). Among clinical journals with an impact factor exceeding 10, positive presentation was associated with a 13.0% (95% CI 9.5% to 16.5%) increase in downstream citations, versus a non-significant 3.3% (−0.6% to 7.2%) increase in clinical journals with an impact factor of 10 or less. Among general life science articles, the positive presentation of research findings was associated with a 12.0% (10.2% to 13.8%) relative increase in downstream citations for articles published

BMJ: first published as 10.1136/bmj.l6573 on 16 December 2019. Downloaded from http://www.bmj.com/ on 17 December 2019 by guest. Protected by copyright.

doi: 10.1136/bmj.l6573 | *BMJ* 2019;367:l6573 | the bmj

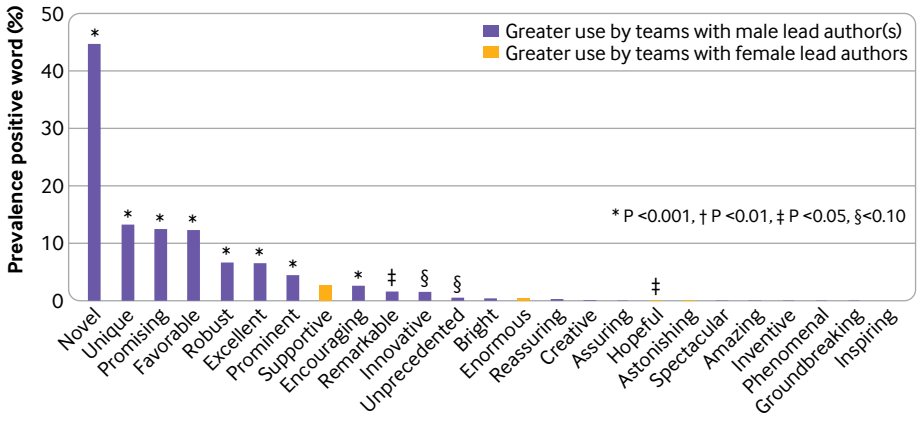

Fig 2 | Gender differences in positive presentation of clinical research articles according to specific positive words, based on 101 720 articles during 2002-17. Gender differences in positive presentation were estimated from a linear probability model with a binary dependent variable (variable equal to one if positive word was used in an article's title or abstract and zero otherwise) and a single independent variable that recorded the gender composition of the first and last authors (variable equal to one if both authors were female ("female lead authors") and zero otherwise). In the figure, positive words are ranked according to the proportion of all positively framed articles that used a given word (for example, "novel" was the most commonly used positive word with a prevalence of 45% among all positively-framed articles). The word "favorable" included US and UK spelling.

in journals with an impact factor exceeding 10, versus a 5.7% (5.5% to 5.9%) increase in journals with an impact factor of 10 or less.

Sensitivity analyses

Our findings were robust to analyses of positive word use as a count rather than binary variable; articles by female first and last authors used positive words approximately 40% less often compared with articles in which a man served in the first and/or last author position (supplementary table S17 in appendix). To understand better the context of the use of positive words, we used Natural Language Processing to identify the nouns that appeared immediately following the positive words (adjectives) (supplementary tables S8 to S10). The positive words commonly modified nouns such as "strategy," "approach," and "mechanism," suggesting that these words are often used to characterize the research findings (such as "novel mechanism" or "novel approach"). Positive word use did not differ between titles and abstracts (supplementary table S9) or by author gender (supplementary table S10). Adding abstract length as a covariate to our regressions only marginally affected the magnitude of our estimated effect sizes (roughly 1%; supplementary table S12). Extending the article

Gender compositions	Background	Methods	Results	Conclusions	Total occurences of 25 positive words
All gender pairings	26%	8%	24%	42%	11 701
Female first and last authors only	25%	8%	26%	41%	1239
Mixed and all male gender pairings	26%	8%	24%	43%	10 462
All gender pairings					
Top 10 individual words					**Occurence of positive word**
Novel	35%	8%	14%	43%	4756
Favorable	13%	7%	41%	39%	1722
Promising	36%	1%	2%	61%	1243
Unique	24%	14%	28%	33%	960
Excellent	9%	3%	49%	38%	825
Robust	15%	6%	43%	36%	719
Prominent	20%	5%	49%	25%	415
Supportive	15%	28%	30%	27%	373
Encouraging	15%	0%	5%	79%	272
Remarkable	16%	2%	45%	37%	122

0-10% 11-20% 21-30% 31-40% 41-50% >50%

Fig 3 | Location of most common positive words in structured abstracts of clinical research articles during 2002-17. Percentages may not sum to 100% due to rounding. Based on 62 877 structured abstracts (81% of available abstracts for clinical research articles).

BMJ: first published as 10.1136/bmj.l6573 on 16 December 2019. Downloaded from http://www.bmj.com/ on 17 December 2019 by guest. Protected by copyright.

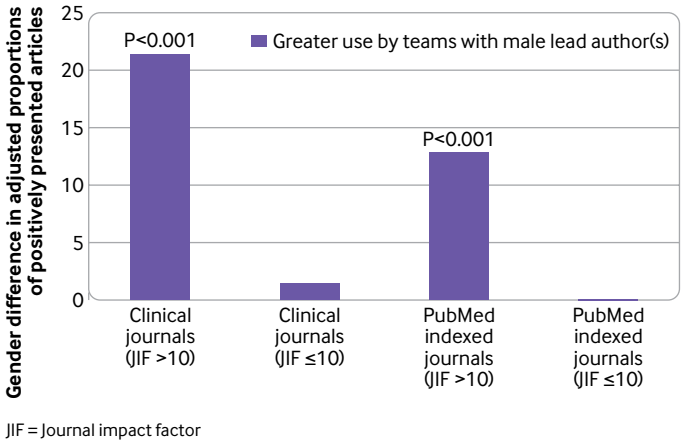

Fig 4 | Adjusted gender differences in positive presentation of articles, 2002-17. Figure shows relative percentage difference in positively framed articles between articles in which both first and last authors were female ("female lead authors") and articles in which first and/or last author was male ("male lead authors") based on 101 720 clinical research articles and approximately 6.2 million life science articles. Gender differences in positive presentation were estimated from a multivariable linear probability model with a binary dependent variable equal to one if any positive word was used in the articles' title or abstract and zero otherwise. Covariates included a binary variable equal to one if both first and last authors were female ("female lead authors") and zero otherwise ("male lead author(s)"), indicator variables for year of publication and publishing journal, a continuous measure of use of positive terms in the article's research field, number of co-authors, and percentage female co-authors in the article byline. Full regression results are provided in the appendix on bmj.com. Figure presents estimates for research articles in clinical journals and broader life science articles indexed in PubMed.

types to include non-research articles produced similar results to our baseline analysis (supplementary table S15). An examination of the accuracy with which Genderize designated gender across different ethnic backgrounds found that the algorithm could assign gender with average probabilities exceeding 90% across all ethnic categories (supplementary table S5). Moreover, varying the probability threshold for gender designation to 50%, 80%, and 95% produced results consistent with our main findings based on a 90% probability threshold (supplementary table S24). Examining the probability of positive framing by the four possible gender combinations of first and last authors found that female first and last author pairs used positive words the least often. Mixed gender pairs and teams of male first and last authors, by contrast, had relatively similar levels of positive word usage (supplementary figure S18). Next, sensitivity analyses on the relationship between positive framing and downstream citations, using Poisson regressions and Scopus as a secondary source for citations, yielded consistent results but effect sizes of roughly double those presented in the main analysis (supplementary table S19). Finally, we neither discerned corresponding gender differences in the use of negative words (supplementary figure S28) nor in the use of positive words in introductions (supplementary table S31), suggesting that gender differences in the use of positive words primarily relate to how research is presented in titles and abstracts of publications.

Discussion

Principal findings

Analyzing titles and abstracts from 101 720 clinical research articles and from approximately 6.2 million articles indexed in PubMed during 2002 to 2017, we found that articles in which the first and last author were both women were significantly less likely to use positive terms to describe research findings compared with articles in which the first and/or last author was a man. Gender differences in the positive presentation of research findings were largest in high impact journals. Positive presentation of research findings was also associated with higher downstream citations, suggesting that observed gender differences in the positive framing of research may have a number of important implications.

Study implications

Although the gender gap in academic medicine and the life sciences has been closing at the earliest stages in the career pipeline, it has remained at later stages, with female scientists being less likely to transition to junior and then senior faculty.[8] Prior research has mostly pointed to discrimination, biased resource allocation, or outright exclusion as causes for slower career progress.[8 14 15] Less empirical research has focused on whether gender differences may exist in how scientists promote themselves and their research. This lens seems important as it may inform possible interventions to reduce the gender gap in science. Although our findings suggest that men and women differ in how they "spin" research results, our approach cannot determine the optimal degree of positive framing for the dissemination of research (that is, "spin" may have disadvantages for the advancement of science).

A broad social science literature that examines gender disparities in professional labor markets has argued that men engage in more self promotion than women,[16 17] potentially to their advantage.[18-20] Gender stereotypes depicting women as more communal and less forceful and achievement oriented than men have been shown to influence women's careers,[21-23] for example, through muting women's voices in teams that include both men and women. In particular, some evidence suggests that women tend to underestimate their abilities relative to men,[24-26] especially in public settings, which may contribute to women negotiating less often and less forcefully[27-29] and being less likely to seek promotion than men.[30 31]

Self-promotion may take different forms: positively framing research findings, using social media to call attention to one's research, or presentation at scientific meetings. Compared with salary negotiations or hiring and promotion decisions, opportunities for self promotion occur more often and depend largely on the discretion of the individual. These features may render self promotion a powerful tool for gradually challenging gender stereotypes.[20] Moreover, although science is increasingly produced by teams,[32] perceptions of individual performance continue to be important

BMJ: first published as 10.1136/bmj.l6573 on 16 December 2019. Downloaded from http://www.bmj.com/ on 17 December 2019 by guest. Protected by copyright.

doi: 10.1136/bmj.l6573 | *BMJ* 2019;367:l6573 | the bmj

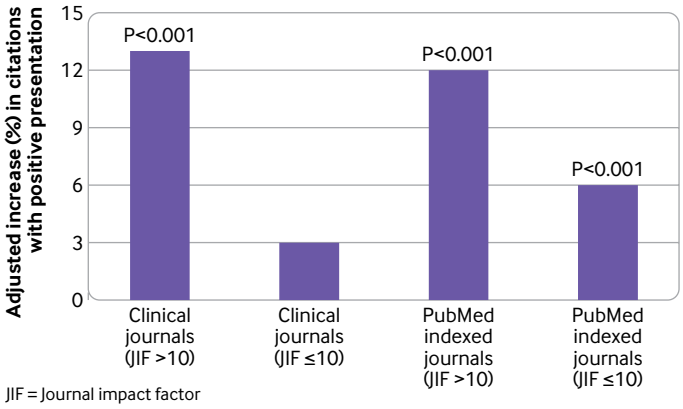

Fig 5 | Adjusted percent increase in downstream citations associated with positive presentation of articles, 2002-17. Adjusted relative percent increase in downstream citations was estimated from a multivariable linear regression in which the dependent variable was forward citations (log transformed) and covariates included a binary variable for whether an article was positively presented, indicator variables for year of publication and publishing journal, a continuous measure of use of positive terms in the article's research field, number of co-authors, percentage female co-authors in the article byline, and the gender composition of first and last authors. Full regression results are provided in the appendix on bmj.com. Figure presents estimates for research articles in clinical journals and broader life science articles indexed in PubMed.

determinants of career progress. Self promotion may therefore be critical to drawing attention to one's abilities and to pursuing careers more forcefully.[33 34]

Our study provides large scale evidence that men in academic medicine and the life sciences more broadly may present their own research more favorably than women, and that these differences may help to call attention to their research through higher downstream citations. These findings suggest that differences in the degree of self promotion may contribute to the well documented gender gaps in academic medicine and in science more broadly.[6 8 10] Moreover, the observed gender differences seem most pronounced in the highest impact journals, which may disadvantage women when it matters the most.

Strengths and limitations of study

Although this study is the first, to our knowledge, that analyzes gender differences in self promotion at a large scale, our approach has several limitations. First, our data do not allow us to assess whether men and women are held to different standards during the journal evaluation process. Recent research examining article submissions to a leading economics journal suggests that journal editors and reviewers opine on women's versus men's writing more frequently, in turn, tangibly altering the end product and prolonging peer review for women.[35] Whether this phenomenon may extend to the large body of clinical and life science research analyzed here remains unclear but could, in part, contribute to women being systematically less likely to use positive words to describe their research. Second, and relatedly, our analysis focused on published articles' titles and abstracts, both of which result from editorial curation. Larger gender differences may exist in how scientific findings are reported at the article submission stage, before selection and editorial modifications occur, in which case our estimates would understate gender differences in positive framing.

Of course, articles with greater scientific novelty and importance may appropriately use terms that positively frame research findings. If men more commonly author such articles, it could account for our observed findings. However, our analysis adjusted for the specific journal, for the area of research based on a granular assessment of articles' MeSH terms, and for journal impact factor, which should mitigate these concerns.

Our analysis uses a set of 25 words that were identified by prior research as distinctly positive and frequently used in life science articles. Expanding this set of positive words might have resulted in a larger percentage of articles being classified as positively framed. But our study focused on the relative gender difference in positive framing and, other than the frequency of using these words, we found no differences between men and women in the way that they used them. If the results from this sample of positive words extended to other positive words, then these gender differences might be large not only in relative terms but in absolute terms as well. Moreover, gender differences may also exist in other forms of self promotion, potentially contributing to gender differences in a variety of career outcomes.

Conclusions

In an analysis of titles and abstracts from over 100 000 clinical research articles and over six million articles in the life sciences, we found that articles in which the first and last author were both women were less likely to use positive terms to describe research findings compared with articles in which the first and/or last author was a man. As this gender difference was most pronounced in the highest impact journals and was associated with higher downstream citations, the potential propensity of women to present equivalent work less positively than men may influence career progress and deserves further attention.

Contributors: All authors contributed to the design and conduct of the study, data collection and management, analysis and interpretation of the data; and preparation, review, or approval of the manuscript. ABJ supervised the study and is the guarantor. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: external funding support from the Office of the Director, NIH (1DP5OD017897) to ABJ, from the German Research Foundation (DFG grant LE 3426/1-2) to MJL, and from Yale University's Initiative on Leadership and Organisation to MJL and OS; no financial relationships with any organizations that might have an interest in the submitted work in the previous 3 years; and no other relationships or activities that could appear to have influenced the submitted work. ABJ received consulting fees unrelated to this work from Pfizer, Hill Rom Services, Bristol Myers Squibb, Novartis, Amgen, Eli Lilly, Vertex Pharmaceuticals, AstraZeneca, Celgene, Tesaro, Sanofi Aventis, Biogen, Precision Health Economics, and Analysis Group. MJL is a co-founder and shareholder of AaviGen GmbH.

Funding: The research conducted was independent of any involvement from the sponsor of the study. The study sponsor was not

BMJ: first published as 10.1136/bmj.l6573 on 16 December 2019. Downloaded from http://www.bmj.com/ on 17 December 2019 by guest. Protected by copyright.

involved in study design, data interpretation, writing, or the decision to submit the article for publication. The authors acknowledge research assistance by Leo Schmallenbach as well as support by the German state of Baden-Württemberg through bwHPC—high performance cluster computing resources.

Ethical approval: Not required.

Transparency: ABJ affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies are disclosed.

Data sharing: The data and analytic methods will be available to other researchers for purposes of reproducing the results or replicating the procedure from Harvard Dataverse at https://doi.org/10.7910/DVN/YHJUCJ

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

1 Tesch BJ, Wood HM, Helwig AL, Nattinger AB. Promotion of women physicians in academic medicine. Glass ceiling or sticky floor?*JAMA* 1995;273:1022-5. doi:10.1001/jama.1995.03520370064038
2 Association of American Medical Colleges. The state of women in academic medicine: the pipeline and pathways to leadership. 2013-14. https://store.aamc.org/downloadable/download/sample/sample_id/228/.
3 Jagsi R, Griffith KA, Stewart A, Sambuco D, DeCastro R, Ubel PA. Gender differences in the salaries of physician researchers. *JAMA* 2012;307:2410-7. doi:10.1001/jama.2012.6183
4 Ly DP, Seabury SA, Jena AB. Differences in incomes of physicians in the United States by race and sex: observational study. *BMJ* 2016;353:i2923. doi:10.1136/bmj.i2923
5 Ley TJ, Hamilton BH. Sociology. The gender gap in NIH grant applications. *Science* 2008;322:1472-4. doi:10.1126/science.1165878
6 Larivière V, Ni C, Gingras Y, Cronin B, Sugimoto CR. Bibliometrics: global gender disparities in science. *Nature* 2013;504:211-3. doi:10.1038/504211a
7 Scharff C. Blowing your own trumpet: exploring the gendered dynamics of self-promotion in the classical music profession. *Sociol Rev* 2015;63:97-112. doi:10.1111/1467-954X.12243
8 Lerchenmueller MJ, Sorenson O. The gender gap in early career transitions in the life sciences. *Res Policy* 2018;47:1007-17. doi:10.1016/j.respol.2018.02.009
9 Jena AB, Khullar D, Ho O, Olenski AR, Blumenthal DM. Sex differences in academic rank in US medical schools in 2014. *JAMA* 2015;314:1149-58. doi:10.1001/jama.2015.10680
10 Lerchenmüller C, Lerchenmueller MJ, Sorenson O. Long-term analysis of sex differences in prestigious authorships in cardiovascular research supported by the National Institutes of Health. *Circulation* 2018;137:880-2. doi:10.1161/CIRCULATIONAHA.117.032325
11 Lerchenmueller MJ. Gender designation via Genderize API with unlimited request doi:106084/m9figshare4563814 2016.
12 CodingNews. Gender detection. 2015. http://codingnews.info/post/gender-detection.html.
13 Vinkers CH, Tijdink JK, Otte WM. Use of positive and negative words in scientific PubMed abstracts between 1974 and 2014: retrospective analysis. *BMJ* 2015;351:h6467. doi:10.1136/bmj.h6467
14 Ly DP, Seabury SA, Jena AB. Differences in incomes of physicians in the United States by race and sex: observational study. *BMJ* 2016;353:i2923. doi:10.1136/bmj.i2923
15 Moss-Racusin CA, Dovidio JF, Brescoll VL, Graham MJ, Handelsman J. Science faculty's subtle gender biases favor male students. *Proc Natl Acad Sci U S A* 2012;109:16474-9. doi:10.1073/pnas.1211286109
16 Cheryan S, Ziegler SA, Montoya AK, Jiang L. Why are some STEM fields more gender balanced than others?*Psychol Bull* 2017;143:1-35. doi:10.1037/bul0000052
17 Scharff C. Blowing your own trumpet: Exploring the gendered dynamics of self-promotion in the classical music profession. *Sociol Rev* 2015;63:97-112. doi:10.1111/1467-954X.12243
18 Rudman LA. Self-promotion as a risk factor for women: the costs and benefits of counterstereotypical impression management. *J Pers Soc Psychol* 1998;74:629-45. doi:10.1037/0022-3514.74.3.629
19 Moss-Racusin CA, Phelan JE, Rudman LA. When men break the gender rules: Status incongruity and backlash against modest men. *Psychol Men Masc* 2010;11:140. doi:10.1037/a0018093
20 Moss-Racusin CA, Rudman LA. Disruptions in women's self-promotion: the backlash avoidance model. *Psychol Women Q* 2010;34:186-202. doi:10.1111/j.1471-6402.2010.01561.x
21 Eagly AH, Karau SJ. Role congruity theory of prejudice toward female leaders. *Psychol Rev* 2002;109:573-98. doi:10.1037/0033-295X.109.3.573
22 Eagly AH, Wood W, Diekman AB. Social role theory of sex differences and similarities: a current appraisal. In: Eckes T, ed. *The Developmental Social Psychology of Gender*. Routledge, 2000: 123-74.
23 Fiske ST, Stevens LE. *What's so special about sex? Gender stereotyping and discrimination.* Sage Publications, 1993.
24 Daubman KA, Heatherington L, Ahn A. Gender and the self-presentation of academic achievement. *Sex Roles* 1992;27:187-204. doi:10.1007/BF00290017
25 Brown LB, Uebelacker L, Heatherington L. Men, women, and the self-presentation of achievement. *Sex Roles* 1998;38:253-68. doi:10.1023/A:1018737217307
26 Gneezy U, Niederle M, Rustichini A. Performance in competitive environments: Gender differences. *Q J Econ* 2003;118:1049-74. doi:10.1162/00335530360698496
27 Leibbrandt A, List JA. Do women avoid salary negotiations? Evidence from a large-scale natural field experiment. *Manage Sci* 2015;61:2016-24. doi:10.1287/mnsc.2014.1994
28 Blau FD, Kahn LM. The gender wage gap: Extent, trends, and explanations. *J Econ Lit* 2017;55:789-865. doi:10.1257/jel.20160995
29 Babcock L, Laschever S. *Women don't ask: Negotiation and the gender divide.* Princeton University Press, 2009.
30 Janoff-Bulman R, Wade MB. The dilemma of self-advocacy for women: Another case of blaming the victim?*J Soc Clin Psychol* 1996;15:143-52. doi:10.1521/jscp.1996.15.2.143
31 Blau FD, Ferber MA, Winkler AE. *The economics of women, men and work.* Pearson Higher Education, 2013.
32 Wuchty S, Jones BF, Uzzi B. The increasing dominance of teams in production of knowledge. *Science* 2007;316:1036-9. doi:10.1126/science.1136099
33 Williams CL, Muller C, Kilanski K. Gendered organizations in the new economy. *Gend Soc* 2012;26:549-73. doi:10.1177/0891243212445466
34 Shackelford S, Wood W, Worchel S. Behavioral styles and the influence of women in mixed-sex groups. *Soc Psychol Q* 1996;59:284-93. doi:10.2307/2787024
35 Hengel E. *Publishing while female: Technical report, Technical report.* University of Liverpool, 2016.

Appendix: Supplementary data and methods; supplementary tables and figures S1 to S31

BMJ: first published as 10.1136/bmj.l6573 on 16 December 2019. Downloaded from http://www.bmj.com/ on 17 December 2019 by guest. Protected by copyright.