Published in final edited form as:

Epidemiology. 2017 March ; 28(2): 159–168. doi:10.1097/EDE.0000000000000593.

The changing face of epidemiology: gender disparities in citations?

Enrique F Schisterman¹, Chandra W Swanson¹, Ya-Ling Lu^{1,2}, and Sunni L Mumford¹

¹Epidemiology Branch, Division of Intramural Population Health Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD

²NIH Library, National Institutes of Health, Bethesda, MD

Abstract

Background—Female biomedical scientists tend to publish fewer articles as last author than their male colleagues and accrue fewer citations per publication. We seek to understand whether epidemiology follows this pattern.

Methods—We gathered aggregate information on the current gender distribution of epidemiology departments (n=29 of 71 surveyed), societies (n=4 of 8), and journal editorial boards (n=6 of 6) using two online surveys and publicly available online information. Bibliometric data from 4149 articles published between 2008–2012 in six high-impact epidemiology journals were drawn from Web of Science and PubMed.

Results—We observed a higher prevalence of female than male doctoral students and epidemiology faculty, particularly at lower faculty ranks. 54% of society members were female. Among editorial boards, all current and emeritus editors-in-chief were male and board membership was largely male (64%). Females were more likely to be first authors, but less likely to be last authors. There were no differences in accrued citations at the 50th percentile by first or last author gender. However, articles with male first and last authors tend to accrue more citations (5.7 citations, 95% CI 2.1, 9.4), mostly driven by the most highly cited articles. This disparity is not fully explained by potential confounders, including seniority.

Conclusions—We found a greater number of female epidemiologists in early career positions and further evidence of potential gender disparity in publication metrics in epidemiology. If epidemiology continues to be practiced by a majority of women, it remains to be seen if these patterns will change over time.

Corresponding author: Enrique F. Schisterman, PhD, Epidemiology Branch, Division of Intramural Population Health Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda MD, 6100 Executive Blvd Room 7B03, Bethesda, MD 20892. schistee@mail.nih.gov; Ph: 301-435-6893.

Conflicts of interest: None declared.

Data and computing code: The data and code used for this analysis will be made available at the Eunice Kennedy Shriver National Institute of Child Health and Human Development Data and Specimen Hub (DASH) website: <https://dash.nichd.nih.gov/>.

eAppendix

The questionnaires were administered through SurveyMonkey and can be found online. The Department Survey can be accessed at the following link: <https://www.surveymonkey.com/r/disparitiesinepicareers-dept>. The Society Survey can be accessed at the following link: <https://www.surveymonkey.com/r/disparitiesinepicareers-society>.

INTRODUCTION

Although more women are entering science-based fields, women biomedical scientists on average are still paid less, promoted less, less likely to receive grant funding, and more likely to leave their careers than men.¹ Women scientists also tend to publish fewer last-author articles than men and accrue fewer citations per publication,^{2–4} though the pattern varies by scientific field and geographic region.⁵

The limited available data specific to epidemiology suggest that women played a major role in the field's development, constituting 35% of American Public Health Association members by 1984, 46% of members of the Society for Epidemiologic Research by 1996, and reaching 66% of attendees of the Congress of Epidemiology in 2006.^{6,7} However, a study evaluating the editorial board composition of top epidemiology journals from 1982 through 1994 found membership was highly skewed towards males.^{6,7} Since rigorous studies of the current demographic make-up and potential gender disparities in epidemiology have not been undertaken, it is unknown whether or not there has been a gender shift and, if so, whether it has translated into an equitable gender distribution of editorial board membership and publication metrics. Therefore, we have evaluated the current gender distribution of epidemiology faculty, students, societies, and editorial boards. We have further evaluated whether there are disparities in epidemiology publication and citation rates using bibliometric analyses, including analyses of factors that may explain the gender distribution in author position of published articles.

METHODS

Data sources

We used publicly available online information and two versions of surveys, one for academic epidemiology departments and the other for selected epidemiology societies, to obtain the current gender distribution of epidemiology faculty, students, and society members. We received institutional review board exemption from the NIH's Office of Human Subject Research (OHSRP #13079). We distributed surveys were distributed to academic departments in the United States and the societies via SurveyMonkey from January 14, 2016 to March 1, 2016. Aggregate de-identified data were collected to maintain confidentiality. We identified the gender of editorial board members of six major epidemiology journals (*American Journal of Epidemiology*, *Annals of Epidemiology*, *Epidemiology*, *European Journal of Epidemiology*, *International Journal of Epidemiology*, and *Journal of Clinical Epidemiology*) using the listing provided by each journal's website. When we were uncertain of a member's gender we completed an online search using the author's full name and institution, as available, to identify gender.

Data for bibliometric analysis were drawn from the Thompson Reuters' Web of Science and PubMed databases. Publications in the top six epidemiology journals (listed above) from 2008–2012 were included. We selected this period to give sufficient time for citations to accrue, while maintaining a recent time frame.^{8,9} In total, 4,149 articles and 1,323 reviews or editorials were available. We restricted this analysis to the 4,149 peer-reviewed articles, which included 119 single-authored articles. Our use of "research articles" followed the

definition of our data source, Web of Science.¹⁰ A cross-check with references in PubMed showed that our dataset included one guideline, and 1,648 cohort or longitudinal studies, among other types of research articles. For each article, Web of Science provided complete metadata, including full names of all authors, article title, journal, author institutions, citation count, and year published. We separately extracted Medical Subject Heading (MeSH) terms of these articles from PubMed.

Survey Instrument

The “Epidemiology Departments” survey consisted of 16 questions regarding the gender, age, and self-identified race/ethnicity (using U.S. Census categories) of students graduating between 2013 and 2015 and current faculty, as well as the tenure status and rank of faculty by gender (eAppendix 1). We first identified departments by filtering the listing on the Association of Schools and Programs of Public Health website for schools offering a Ph.D./Sc.D. in Epidemiology.¹¹ This list was confirmed and augmented using the Council on Education for Public Health website, as well as manual online searching, and included epidemiology programs in both Schools of Medicine and Schools of Public Health.¹² The “Epidemiology Departments” survey was sent to the department chairs of all identified U.S. academic programs (n=68) and three National Institutes of Health intramural epidemiology branches. We selected a random sample of 20% of the non-responding departments to assess potential non-response bias, and determined the gender of these faculty members using an explicit statement of gender, use of sex-specific pronouns on university websites, or, where needed, after locating captioned pictures of the faculty.

The “Epidemiology Societies” survey consisted of 12 questions regarding the gender, age, self-identified race/ethnicity, employment sector, and membership type (student, member, emeritus member) of society members (eAppendix 2). We sent the survey to the Chair or President of eight epidemiology societies chosen for their large membership: the Society for Epidemiologic Research, the Society for Pediatric and Perinatal Epidemiologic Research, the International Society for Environmental Epidemiology, International Epidemiological Association, American Epidemiological Society, American College of Epidemiology, American Public Health Association–Epi Section, and International Society of Pharmacoepidemiology.

Identification of Author Gender

Gender information of all authors was determined by the GenderChecker¹³ database that classified the authors’ first names as male (M), female (F), or unisex (B). We coded names that were not included in the database as unidentified (U). When first names were classified as unisex or unidentified, a manual online search was completed using the author’s full name and institution, as available. The gender was then attributed for authors after locating sex-specific pronouns, an explicit statement of gender, or by locating captioned pictures of the author. When we could find no information to identify the author’s gender, she or he remained classified as unidentified. Thus in the final dataset, all authors were classified as male (M), female (F), or unknown (U). 73% (18,482) of the names were identified by the database, and after the augmented manual name checking, 81% (20,307) were identified. This analysis focused on first and last author genders for the 4,149 articles; 90% of first-

author and 92% of last-author genders were identified. Excluding those that could not be cross-checked, the database's accuracy rate was 86% for female and 98% for male names (eTable 1).

The romanization of East Asian names did not provide sufficient gender information. To check the accuracy of GenderChecker in classifying Chinese, Korean, and Japanese names, we randomly selected 220 publications, resulting in approximately 1,200 author names, 110 of which appear to be Chinese, Korean, or Japanese names. We found that 40 (36%) were classified into certain gender, while the remaining 70 names were either unfound or classified as unisex in GenderChecker. We therefore decided that GenderChecker does not appropriately identify gender for Chinese, Korean, and Japanese names. When first or last authors appeared to be Chinese, Korean, or Japanese, we treated them as neutral and retrieved either images or statements to identify gender when possible.

Categorization of Research Field

Articles were categorized into research fields using assigned MeSH terms. MeSH is the National Library of Medicine's controlled vocabulary thesaurus, which consists of descriptors arranged in a hierarchical structure, and asterisks on MeSH terms represent key paper topics.¹³ Unlike author-generated keywords, MeSH staff subject specialists systematically assign MeSH terms from the set list, improving their validity.^{14–16} Within our dataset, 1,167 unique starred MeSH terms were identified and each article had between one and ten MeSH terms. We assigned each starred MeSH term to one of ten general epidemiology fields (chronic disease, environmental, genetics, infectious disease, methods, nutrition, reproductive, social, and policy) or other. For example, the “Methods” subject area included not only the MeSH term “Methods,” but also “Algorithms,” “Bayes Theorem,” and “Cohort Studies.” Each article was assigned to one subject area using the first starred MeSH term. When the first starred MeSH term was categorized as “other,” the second MeSH term was used, and so on. For some analyses, we used collapsed categories to improve model fit.

Categorization of Regions and Collaborations

Articles were grouped into the six WHO geographic regions according to the first and last authors' institutions: African Region, Region of the Americas, Southeast Asia Region, European Region, Eastern Mediterranean Region, and Western Pacific Region (93 countries contributed articles).¹⁷ We also categorized articles into two collaboration categories: non-international – having all authors from the same country, or international – having authors from at least two countries. During analysis, Africa, Southeast Asia, and Eastern Mediterranean were combined into an “Other” category to account for the low number of articles from these regions.

Categorization of Other Factors

We further evaluated the journal of publication (assigned to numbers one–six in order to deidentify them), the publication year (2008–2012), and the number of co-authors (fewer than four versus four or more). The distribution of number of co-authors was highly right-skewed. After exploring several thresholds, we found the analysis was not sensitive to choice of threshold (eTable 2).

Rank status

As seniority could have been a possible confounder of the association between gender of the first and last author and total citation count,^{18–20} we sought to take this factor into account. However, neither Web of Science nor PubMed metadata included any rank status or seniority information for the authors, so we used h-index²¹ as a proxy for seniority. Proposed by Hirsch in 2005, the h index characterizes the scientific output of a researcher: “A scientist has index h if h of [the number of articles he or she has published (N_p) over n years] have at least h citations each and the other ($N_p - h$) articles have $< h$ citations each”.¹⁹ Because a person’s h-index can never decrease, the higher the h-index, the more senior the researcher, on average.²² We randomly selected 150 articles, distributed evenly by first and last author gender, and used the Web of Science database to extract information regarding the first and last authors’ h-indices and lifetime citations (for the first 50 articles the h-index of the first author was determined, for the next 50 the h-index of the last author, and for the last 50 the h-indices of both the first and last authors were determined). Thus, we obtained information for 100 unique first authors and another 100 last authors; multiple imputations were used to account for the missing values by design. After testing for nonlinearity using splines, we found the relationship between h-index and citations to be linear. We then adjusted for the h-index of the first and last author as a proxy for seniority/rank in the models for total citations.

Bibliometric and Statistical Analyses

To evaluate the factors that predicted the gender of first and last authors, and whether there were gender disparities in citations in epidemiology journals, we conducted bibliometric and statistical analysis as detailed below. Descriptive statistics were used to compare the gender distribution of students, faculty, societies, and editorial boards.

We performed a sensitivity analysis to account for potential non-response bias for the percentage of female faculty overall, and by rank. Using the characteristics of the random sample of non-responding departments, we reweighted the percentage of female faculty to estimate the proportion of male and female faculty had all departments responded. We used two multinomial regression models to evaluate the factors associated with 1) the gender of the first author (M, F, U) and 2) the gender of the last author (M, F, U), including region, research field, journal, publication year, and number of co-authors. We evaluated potential interactions between covariates. Of particular interest was the interaction between first and last author gender, and as such stratified results are presented. Odds ratios and 95% confidence intervals (CI) are presented for the odds of a female first author compared to a male first author, and the odds of a female last author compared to a male last author.

The distribution of citations for articles in this dataset was examined by paired genders of the first and last authors. Articles were classified as male-male (MM), female-female (FF), male-female (MF), female-male (FM), and any unidentified (U). We used general linear models and quantile regressions at different percentiles to test which factors were associated with the number of times an article was cited in this time period. To explore co-author dynamics and collaboration patterns by gender, we used Science of Science Tool²¹ and Gephi²¹ to model networks of collaborations between the top 100 most prolific authors, who

were determined by the frequency of author names in any author position in the articles dataset. The collaboration network displays connections between the 100 authors by gender, weighted by number of publications.

RESULTS

Current Gender Distribution of Epidemiology Departments, Societies, and Editorial Boards

Of the departments queried (n=71), a total of 29 responded to the online survey (41% response rate). Within departments, we observed a higher prevalence of female faculty (Table 1: 55% versus 45%). The female faculty tended to be younger, of lower rank, and less likely to have tenure, compared with their male counterparts. Epidemiology doctoral students graduating between 2013 and 2015 were also more likely to be female (Table 1: 73% versus 27%).

Of the societies queried (n=8), a total of five responded to the online survey (63% response rate), though one society noted that they did not have the information necessary to complete the form. Overall, 54% of society members were female, compared to 46% male (Table 1). Notably, only one society collected enough detailed information to enable comparison of the gender distribution by age, race/ethnicity, and membership type. As such, we have limited the data presented to reflect the overall gender distribution.

Among the editorial boards of the six epidemiology journals studied, all current and emeritus editors-in-chief were found to be male, and editorial board membership was largely male (Table 1: 64%).

Of the departments invited to participate in the survey, 55% were chaired by males and 45% by females, with a non-response rate of 71% among male-chaired departments and 52% among female-chaired departments. Given the relatively low response rates, a random sample of 20% of the non-responding departments was selected, and demographic data were collected from faculty lists from university websites. The non-responding departments had a greater proportion of males in all rank categories (Figure 1: 53% female overall among responders versus 43% female overall among non-responders; 41% female full professors among responders versus 18% female full professors among non-responders), supporting the potential for non-response bias to affect our results. After reweighting the findings to account for the non-responders, we observed a lower percentage of female faculty overall (48%) and among full professors (32%). Though the higher percentage of female faculty was attenuated by weighting, it remained among assistant (61%) and associate (52%) professors. Thus, there could be potential overestimation of the prevalence of female epidemiologists in departments, particularly across the lower ranks; however, response bias did not fully account for the survey results.

Bibliometric Analyses

Of the 4,149 articles evaluated, first authors were more likely to be female (47% (95% CI: 45, 49) than male 42% (95% CI: 40, 44) or unidentified 10% (95% CI: 9, 11) (Table 2). Last authors were less likely to be female (30%; 95% CI: 29, 31) than male (63%; 95% CI: 62, 64) and unidentified 7.5% (95% CI: 7.2, 8.8). These patterns were consistent for both first

and last authors across publication years and no differences were observed by field of research. We also observed that female first and last authors were more likely to be from the Americas, Europe, and the West Pacific, but not other regions. There were differences in the proportion of articles published with female first authors by journal, with the percentages ranging from 39% to 53%, as well as with the proportion of articles published with male last authors by journal ranging from 58% to 72%.

Predictors of Gender in Key Authorship Positions—Table 3 summarizes predictors of gender in key authorship positions. Regardless of the gender of the last author, region was predictive of the gender of the first author, with authors from “Other” regions being less likely to be a female first author. Also, first authors from Europe were less likely to be female when the last authors were also female. Publication year and research field were not associated with the gender of the first or last author. There were no year trends observed (95% CI predicting first author gender when last author male: 0.96, 1.1; when last author female: 0.87, 1.0; predicting last author gender when first author male: 1.0, 1.2, when first author female: (1.0, 1.1)). Differences between journals were observed even after accounting for all other factors. In particular, among certain journals, first authors were less likely to be female when writing with a male last author, and last authors were less likely to be female when writing with a female first author.

Predictors of total citations by gender—The articles included in this analysis accrued an average of 25 ± 37 (standard deviation) citations, with female first author articles accruing 24 ± 26 on average compared with 28 ± 48 among male first authors and 21 ± 28 among unidentified. Male last author articles also tended to accrue more citations on average than female or unidentified last author articles (male: 26 ± 41 ; female 24 ± 28 ; unidentified 23 ± 28).

Female first and last author articles accrued the greatest number of median citations (female first: 16 (IQR:15, 17), female last: 16 (15, 17), male first: 15 (14, 16), male last: 15 (14, 16), unidentified first: 14 (13, 15), unidentified last: 15 (13, 17)). There were no substantial differences in the total citations accrued between the gender author dyads (FM, MF, FF, MM, any unknown) up to the 90th percentile of total citation count (Figure 2), though large differences existed among the top 10% of articles. Articles written with a male first author and male last author (MM) accrued approximately six citations more on average (5.7 citations, 95% CI 2.1, 9.4) than articles written by all other dyads after adjusting for region, publication year, journal, and number of co-authors (Table 4). Results using quantile regression at the 90th percentile showed similar results (eTable 3). As expected, articles published in 2008 accrued the greatest number of citations, given their longer period available for citation. Some differences in citation count were also observed by journal and number of co-authors. Finally, articles published by authors not from the Americas or Europe received fewer citations. No differences were observed in citation counts by field of research, specifically comparing methods articles with subject-specific fields. These results were consistent after adjusting for first and last authors' h-index as a proxy for seniority.

Collaboration Network—To try to explain the differences in citation counts by gender of the first and last authors, we evaluated the collaboration networks of the top 100 most prolific authors in our dataset in any authorship position (Figure 3). This collaboration

network displays connections between various authors by gender, weighted by number of publications, and demonstrates that females tend to work with other females and males tend to work with other males. Thus, the collaboration network parallels the patterns seen in the citations analysis, in which FF and MM collaborations are more likely to occur than FM or MF collaborations.

DISCUSSION

In what to our knowledge is the first study of gender and its relation to publication metrics in the field of epidemiology, we found a greater number of female epidemiologists in early-career positions, suggesting a possible demographic shift may be on the horizon, but has not already arrived. This is also highlighted by the fact that the articles in epidemiology journals are written primarily by female first authors, and were less likely to be written by female last authors. These changes may not yet have reached geographic regions outside of the Americas and Europe, and indeed also vary by journal. We also observed striking differences in citations by gender of the first and last author, especially in the most highly cited articles, providing evidence of a potential gender disparity in publication metrics in epidemiology. These differences were not explained by other factors, including seniority.

Our finding showing a greater number of early career female epidemiologists in the field today is consistent with previous descriptions of the field.^{7,23} This finding is also in line with several other scientific disciplines that have observed a similar demographic shift, with more women entering science-based fields.¹ A recent longitudinal study of vision scientists found that while the male:female ratio of predoctoral trainees approached 1:1, female trainees were more likely to drop out at the graduate level than males.²⁴ At the investigator level, the male:female ratio was well over 2:1, though the number of females appeared to be increasing over time. In the life sciences, the proportion of female graduate students has been increasing since 1975 to over 50% in 2014 in several disciplines.²⁵ However, a longitudinal study of patent rates, an important measure of success in the life sciences, found that though the gender gap has improved over time; women patent at about 40% of the rate of men.²⁶ Likewise, in clinical research fields such as psychiatry, gastroenterology, physical therapy and radiology where the percentage of U.S. female first authors has increased proportionally to the growth of females in their respective fields, the percentage of women in the senior author position has remained lower than expected.^{27–30} However, some signs of positive change have been observed. One study evaluated researchers within the first 3 years after receiving a social science doctorate and found that the number of female grant applicants had increased, that a similar number of articles had been published by male versus female researchers, and that female authors tended to have a greater number of citations than male authors.³¹

The determination of first author and last author of a published paper has long-term career implications.^{32,33} Given that we saw crude differences in the percentage of female first and last authors across the four publication years studied, we sought to determine what factors might predict those differences. Indeed, geographic region of the authors, as seen in prior studies,^{1,5,34} and the journal of publication may play a role. Moreover, several other unmeasured elements, including level of contribution, team dynamics, and cultural

differences may affect the initial determination of the order of authors on the byline. Once the authorship order is set, other external factors in the reviewing process may further influence whether a female- versus male-authored manuscript is accepted for publication. These factors likely reflect an implicit gender bias. It is not likely that co-authors would determine authorship order based on gender, and journal authorship conventions, particularly in this dataset, often list authors by first initial and last name, leaving the authors' genders ambiguous. Readers would be unable to choose which article to cite based on author gender only, unless authors were familiar to them. In contrast, selection of majority male editorial board members could reflect explicit bias, as gender of these individuals is more often known at the time of the invitation to join the board. While studies of the peer review process in *Nature*, *Science*, and *JAMA* suggest that gender bias does not explain the differences in author gender of published articles,^{35,36} we lack data on potential gender bias, implicit or explicit, in the review process of epidemiology journals. We were only able to evaluate differences among published articles in this time period and these journals, which may result in selection bias. However, given that we saw factors that explain the gender distribution in author position of published articles alone, it is not inconceivable that these and other factors may be even stronger predictors in unpublished articles.

We observed a difference in citations by first and last author gender, providing evidence of a potential gender disparity in publication metrics in epidemiology. This gender disparity in citations is consistent with several studies across disciplines,^{1-4,32,35-37} although reasons for such a disparity remain unclear^{2,4,38}. Unlike many studies in other fields that did not consider seniority, this study accounted for confounding due to seniority³⁷ and still found similar results. One potential limitation of using h-index of the first and last authors is that not all authors follow the convention that senior author is listed as the last author, and this cannot be determined. To explore this, we compared the h-index of male and female first and last authors, finding no difference between h-index by gender across the distribution of h-index values, which may suggest that the h-index may not be an adequate proxy of seniority. Moreover, residual confounding may still be a possibility as we relied on the h-index and this was only done in a small subset of authors due to the labor intensive manual search necessary to obtain the h-index. The difference in senior author rates may still reflect the lower seniority of women as they may have entered the field later than those males in leadership positions.

When we evaluated the collaboration patterns of the top 100 most prolific authors in this dataset, there were pronounced clusters indicating gender preferences that may coincide with gender bias in citations, with the most prolific men working together, and correspondingly having the highest citation counts. A preference for working with others of the same gender has been noted in other areas of scientific collaboration. For example, an analysis of two American Society for Microbiology meetings found that having at least one female member on the convening team correlated to a higher proportion of invited female speakers.³⁹ Several other studies have also suggested that collaboration patterns differ by gender, though with respect to propensity for collaboration and type of collaboration.^{3,38} Of note, the higher number of citations accrued by MM articles is not likely explained by self-citations, as self-citations make up only 2.5% of citations in this analysis based on the Web of Science metadata. This proportion is well within the generally accepted range of 10%.³

However, again, the sample of top 100 authors represents only a small percentage of the authors included in our dataset. Furthermore, the publication count includes all co-authored articles, in addition to the first- and senior-authored articles, which may inflate the publication counts of those authors involved in large cohort studies or guideline publications and may not reflect the true productivity and impact of the author themselves. This subset also only represents authors who have published extensively in the six journals included in this dataset. Other equally or more prolific authors may have been excluded if their work was spread across a greater variety of journals, or if they did not target these journals. Still, the clustering of same-gender collaborations within highly cited authors may contribute to the overall greater number of citations accrued by male–male articles. With the change in demographics of the field, it remains to be seen if this pattern will persist.

Our study has several strengths, including the use of a large number of publications with several sensitivity analyses to account for potential biases. However, we were limited by several factors. In particular, we experienced a poor response rate from the surveys sent to epidemiology departments and societies. Though we accounted for some of this potential bias using a detailed sensitivity analysis, we were not able to account for the lack of information from societies and some departments regarding age, race, rank, and membership type. This lack of information did not allow us to make inference on possible racial disparities as well. To this end, we encourage societies and departments to collect this data over time so that patterns can be identified. Additionally, the list of those invited to participate in the surveys was not exhaustive. Results could have been different if all field-specific epidemiology societies and journals were included. Furthermore, not all epidemiologists are directly affiliated with epidemiology departments, and would not be included in the demographic profile of the field. Therefore, our results may not be generalizable to sub-disciplines within epidemiology.

In conclusion, we found a greater proportion of female epidemiologists in early career positions, and further evidence of a potential gender disparity in publication metrics in epidemiology with articles written primarily by female first authors, but less likely to be written by female last authors, and large differences in citations by gender. If epidemiology will be practiced by a female majority, it remains to be seen if there will be a corresponding increase of women in senior authorship and leadership positions, or if the differences observed in these positions and in publication metrics will persist. Identifying and addressing gender and minority disparities in epidemiology presents an opportunity to advance the field as a whole.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Sources of financial support: This work was supported by the Intramural Research Program of the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health.

We would like to acknowledge Ellen Chaljub for her data collection support, Aijun Ye for statistical support, Pauline Mendola for commenting on previous versions of this manuscript, and Lisa Federer of the NIH Library for her assistance with editing this manuscript.

References

1. Larivière V, Ni C, Gingras Y, Cronin B, Sugimoto C. Gender Disparities in Science. *Nature*. 2013
2. van den Besselaar P, Sandstrom U. Gender differences in research performance and its impact on careers: a longitudinal case study. *Scientometrics*. 2016; 106:143–162. [PubMed: 26798162]
3. West JD, Jacquet J, King MM, Correll SJ, Bergstrom CT. The role of gender in scholarly authorship. *PLoS One*. 2013; 8:e66212. [PubMed: 23894278]
4. Symonds MR, Gemmell NJ, Braisher TL, Gorringer KL, Elgar MA. Gender differences in publication output: towards an unbiased metric of research performance. *PLoS One*. 2006; 1:e127. [PubMed: 17205131]
5. Bentley PJ. Cross-country differences in publishing productivity of academics in research universities. *Scientometrics*. 2015; 102:865–883.
6. Dickersin K, Fredman L, Flegal KM, Scott JD, Crawley B. Is There a Sex Bias in Choosing Editors? *Epidemiology Journals as an Example*. *JAMA*. 1998
7. Carter-Pokras OD, Spirtas R, Bethune L, Mays V, Freeman VL, Cozier YC. The training of epidemiologists and diversity in epidemiology: findings from the 2006 Congress of Epidemiology survey. *Ann Epidemiol*. 2009; 19:268–75. [PubMed: 19344867]
8. Abramo G, Cicero T, D'Angelo CA. Assessing the varying level of impact measurement accuracy as a function of the citation window length. *J Informetr*. 2011; 5:659–667.
9. Wang J. Citation time window choice for research impact evaluation. *Scientometrics*. 2012; 94:851–872.
10. Thomson Reuters. [Accessed July 15, 2016] Searching the Document Type Field. http://images.webofknowledge.com/WOKRS59B4/help/WOS/hs_document_type.html
11. Association of Schools & Programs of Public Health. Academic Program Finder. Association of Schools & Programs of Public Health; http://www.aspph.org/program-finder/?area_of_study=11°ree=PhD;DrPH;ScD+or+SD [Accessed April 19, 2016]
12. Council on Education for Public Health. Accredited Schools & Programs. Council on Education for Public Health; <http://ceph.org/accredited/> [Accessed April 19, 2016]
13. Medicine USNLo. [Accessed April 19, 2016] Medical subject headings (MeSH) in MEDLINE/ PubMed: A tutorial. <https://www.nlm.nih.gov/bsd/disted/meshtutorial/principlesofmedlinesubjectindexing/majortopics/>
14. Jenuwine ES, Floyd JA. Comparison of Medical Subject Headings and text-word searches in MEDLINE to retrieve studies on sleep in healthy individuals. *J Med Libr Assoc*. 2004; 92:349–53. [PubMed: 15243641]
15. Chang AA, Heskett KM, Davidson TM. Searching the literature using medical subject headings versus text word with PubMed. *Laryngoscope*. 2006; 116:336–40. [PubMed: 16467730]
16. Nelson, SJ. Medical Terminologies That Work: The Example of MeSH. *Proceedings of the 10th International Symposium on Pervasive Symptoms, Algorithms, and Networks (ISPAN 2009)*; Kaohsiung, Taiwan. 2009. p. 380-384.
17. World Health Organization. [Accessed April 19, 2016] Health statistics and information systems: Definition of region groupings. http://www.who.int/healthinfo/global_burden_disease/definition_regions/en/
18. Rorstad K, Aksnes DW. Publication rate expressed by age, gender and academic position - A large-scale analysis of Norwegian academic staff. *J Informetr*. 2015; 9:317–333.
19. Antonio-García MT, López-Navarro I, Rey-Rocha J. Determinants of success for biomedical researchers: a perception-based study in a health science research environment. *Scientometrics*. 2014; 101:1747–1779.
20. Mauleón E, Bordons M, Oppenheim C. The effect of gender on research staff success in life sciences in the Spanish National Research Council. *Res Eval*. 2008; 17:213–225.

21. Hirsch JE. An index to quantify an individual's scientific research output. *Proc Natl Acad Sci U S A*. 2005; 102:16569–72. [PubMed: 16275915]
22. Bornmann L, Daniel H-D. What do we know about the index? *J Am Soc Inf Sci Technol*. 2007; 58:1381–1385.
23. Dickersin K, Fredman L, Flegal KM, Scott JD, Crawley B. Is there a sex bias in choosing editors?: Epidemiology journals as an example. *JAMA*. 1998; 280:260–264. [PubMed: 9676675]
24. Cooper EA, Radonjic A. Gender representation in the vision sciences: A longitudinal study. *J Vis*. 2016; 16:17.
25. National Science Foundation, National Center for Science and Engineering Statistics. Women, Minorities, and Persons with Disabilities in Science and Engineering: 2015. Arlington, VA: 2015. Special Report NSF 15–311 <http://www.nsf.gov/statistics/wmpd/> [Accessed July 15, 2016]
26. Ding WW, Murray F, Stuart TE. Gender differences in patenting in the academic life sciences. *Science*. 2006; 313:665–667. [PubMed: 16888138]
27. Amering M, Schrank B, Sibitz I. The Gender Gap in High-Impact Psychiatry Journals. *Acad Med*. 2011; 86:946–952. [PubMed: 21694564]
28. Long MT, Leszczynski A, Thompson KD, Wasan SK, Calderwood AH. Female authorship in major academic gastroenterology journals: a look over 20 years. *Gastrointest Endosc*. 2015; 81:1440–1447. [PubMed: 25887727]
29. Kaufman RR, Chevan J. The Gender Gap in Peer-Reviewed Publications by Physical Therapy Faculty Members: A Productivity Puzzle. *Phys Ther*. 2011; 91:122–131. [PubMed: 21127164]
30. Liang T, Zhang C, Khara RM, Harris AC. Assessing the Gap in Female Authorship in Radiology: Trends Over the Past Two Decades. *J Am Coll Radiol*. 2015; 12:735–741.
31. van Arensbergen P, van der Weijden I, van den Besselaar P. Gender differences in scientific productivity: a persisting phenomenon? *Scientometrics*. 2012; 93:857–868. [PubMed: 23162173]
32. Cikara M, Rudman L, Fiske S. Dearth by a Thousand Cuts?: Accounting for Gender Differences in Top-Ranked Publication Rates in Social Psychology. *J Soc Issues*. 2012; 68:263–285.
33. Geraci L, Balsis S, Busch AJB. Gender and the h index in psychology. *Scientometrics*. 2015
34. King DA. The scientific impact of nations. *Nature*. 2004; 430:311–316. [PubMed: 15254529]
35. Braisher TL, Symonds MR, Gemmell NJ. Publication success in Nature and Science is not gender dependent. *Bioessays*. 2005; 27:858–9. [PubMed: 16015605]
36. Gilbert JR, Williams ES, Lundberg GD. Is There Gender Bias in JAMA's Peer Review Process? *JAMA*. 1994:272.
37. Holliday EB, Jaggi R, Wilson LD, Choi M, Thomas CR Jr, Fuller CD. Gender differences in publication productivity, academic position, career duration, and funding among U.S. academic radiation oncology faculty. *Acad Med*. 2014; 89:767–73. [PubMed: 24667510]
38. Abramo G, D'Angelo CA, Murgia G. Gender differences in research collaboration. *J Informetr*. 2013; 7:811–822.
39. Casadevall A, Handelsman J. The presence of female conveners correlates with a higher proportion of female speakers at scientific symposia. *mBio*. 2014; 5(1):e00846–13. [PubMed: 24399856]

Figure 1.

Percentage of female faculty among responding and non-responding departments overall, by rank, and reweighted to account for potential non-response bias

Figure 2.
Distribution in total citations by gender of the first and last authors among all articles published in six epidemiology journals from 2008–2012

Figure 3.

Collaboration network by gender of the 100 most prolific authors (regardless of authorship position) in six epidemiology journals from 2008–2012

Table 1

Demographic characteristics of epidemiology departments, societies, and editorial boards by gender

Category	No. (%)			
	Total	Female	Male	Unknown
Professors/Members (n =29 departments)				
Gender, ^a n (%)	680	374 (55)	306 (45)	
Tenure Status				
Tenured	310	147 (47)	163 (53)	
Tenure-track	170	106 (62)	64 (38)	
Non-tenure track	157	88 (56)	69 (44)	
Rank				
Full	278	115 (41)	163 (59)	
Associate	171	94 (55)	77 (45)	
Assistant	177	124 (70)	53 (30)	
Age				
35 years and younger	96	71 (74)	25 (26)	
36 – 50 years old	285	167 (59)	118 (41)	
51 years and older	299	136 (45)	163 (55)	
Race				
Black/African	44 (6)	33 (75)	11 (25)	
Hispanic or Latino	23 (3)	9 (39)	14 (61)	
White	588 (79)	294 (50)	294 (50)	
Other	91 (12)	47 (52)	44 (48)	
Students (n = 24 departments)				
Gender	338	247 (73)	91 (27)	
Race				
Black/African	28 (8)	21 (75)	7 (25)	
Hispanic or Latino	10 (3)	5 (50)	5 (50)	
White	209 (62)	161 (77)	48 (23)	
Other	91 (27)	60 (66)	31 (34)	
Editorial Boards (n=6 journals) ^b				
Gender	299	75 (25)	192 (64)	32 (11)
Position				
Current Editor in Chief	8 (5)	0 (0)	8 (100)	0 (0)
Emeritus Editor in Chief	8 (5)	0 (0)	6 (75)	2 (25)

Category	No. (%)			
	Total	Female	Male	Unknown
Editorial Board	135 (90)	29 (22)	86 (64)	20 (15)
Members				
Associate Editors	115 (71)	36 (31)	76 (66)	3 (3)
Policy Advisory Boards	13 (8)	3 (23)	8 (62)	2 (15)
Editorial Consultants Board	30 (18)	6 (20)	19 (63)	5 (17)
Other Editors	5 (3)	2 (40)	1 (20)	2 (40)
Societies (n = 4 societies)				
Gender (%)		^c (54)	^c (46)	

^aCalculated from reported numbers of professors by rank, the most consistent and complete response.

^bAmerican Journal of Epidemiology's Associate Editors include two Associate Editors in Residence. The European Journal of Epidemiology had two categories for editors, editors and editorial board; editors were described as "Associate Editors". The International Journal of Epidemiology category "editors" was added to the editorial board. The "other" category includes managing, assistant, and academic editors.

^cTotal membership of societies is unknown. Percentages are based on figures provided by the societies.

Table 2
Distribution and characteristics of 4149 articles from 2008–2012 from six epidemiology journals by gender of the first and last author

Category	Total	Female First Author	Male First Author	Unknown First Author	Total	Female Last Author	Male Last Author	Unknown Last Author
Gender, n(%)		1968 (47)	1753 (42)	428 (10)		1233 (30)	2603 (63)	313 (8)
Publication year (%)								
2008	849	387 (46)	353 (42)	109 (13)		220 (26)	522 (61)	107 (13)
2009	861	407 (47)	366 (43)	88 (10)		263 (31)	542 (63)	56 (7)
2010	851	410 (48)	350 (41)	91 (11)		256 (30)	543 (64)	52 (6)
2011	819	393 (48)	366 (45)	60 (7)		267 (33)	521 (64)	31 (4)
2012	769	371 (48)	318 (41)	80 (10)		227 (30)	475 (62)	67 (9)
Field category								
Methods	1253	595 (47)	530 (42)	128 (10)		358 (29)	809 (65)	86 (7)
Other	2896	1373 (47)	1223 (42)	300 (10)		875 (30)	1794 (62)	227 (8)
WHO Region								
Americas	2044	1050 (51)	856 (42)	138 (7)	2065	689 (33)	1296 (63)	80 (4)
Europe	1594	742 (47)	695 (44)	157 (10)	1525	405 (27)	1004 (66)	116 (8)
West Pacific	414	154 (37)	149 (36)	111 (27)	393	103 (26)	204 (52)	86 (22)
Other	97	22 (23)	53 (55)	22 (23)	166	36 (22)	99 (60)	31 (19)
International								
One Country	2687	1323 (49)	1101 (41)	263 (10)		878 (33)	1615 (60)	194 (7)
Multinational	1462	645 (44)	652 (45)	165 (11)		355 (24)	988 (68)	119 (8)
Journal								
1	1393	733 (53)	532 (38)	128 (9)		489 (35)	804 (58)	100 (7)
2	580	289 (50)	229 (39)	62 (11)		193 (33)	343 (59)	44 (8)
3	475	231 (49)	218 (46)	26 (6)		123 (26)	344 (72)	8 (2)
4	419	188 (45)	176 (42)	55 (13)		108 (26)	267 (64)	44 (11)
5	643	252 (39)	288 (45)	103 (16)		165 (26)	405 (63)	73 (11)

Category	Total	Female First Author	Male First Author	Unknown First Author	Total	Female Last Author	Male Last Author	Unknown Last Author
6	639	275 (43)	310 (49)	54 (8)		155 (24)	440 (69)	44 (7)
Number of co-authors, n(%)								
< 4	879	334 (38)	469 (53)	76 (9)		247 (28)	579 (66)	53 (6)
4	3270	1634 (50)	1284 (39)	352 (11)		986 (30)	2024 (62)	260 (8)
Average number of citations (n)		23.6±25.6	27.8±47.9	21.2±27.5		23.6±28.1	26.1±41.4	22.9±27.5

Table 3

Predictors of gender of the first and last author among 4,149 articles from 2008–2012 from six epidemiology journals, adjusted for all of the other factors

Categories	Predictors of First Author – OR (95% CI)		Predictors of Last Author – OR (95% CI)	
	Last=F	Last=M	First=F	First=M
Publication Year				
2008	Reference	Reference	Reference	Reference
2009	1.2 (0.79,1.7)	0.97 (0.75,1.2)	1.2 (0.92,1.7)	1.0 (0.72,1.5)
2010	0.99 (0.67,1.5)	1.1 (0.84,1.4)	1.1 (0.85,1.5)	1.2 (0.85,1.8)
2011	0.87 (0.59,1.3)	0.98 (0.76,1.3)	1.2 (0.85,1.6)	1.3 (0.91,1.9)
2012	0.93 (0.62,1.4)	1.1 (0.86,1.4)	1.1 (0.81,1.5)	1.3 (0.91,1.9)
Field category				
Methods	Reference	Reference	Reference	Reference
Others	0.98 (0.75,1.3)	0.95 (0.80,1.1)	1.1 (0.93,1.4)	1.1 (0.83,1.4)
Journal				
1	Reference	Reference	Reference	Reference
2	1.2 (0.80,1.7)	0.94 (0.72,1.2)	0.97 (0.73,1.3)	0.80 (0.56,1.2)
3	0.91 (0.60,1.4)	0.82 (0.63,1.1)	0.65 (0.47,0.90)	0.61 (0.41,0.89)
4	1.3 (0.78,2.1)	0.68 (0.50,0.9)	1.2 (0.80,1.7)	0.63 (0.41,0.97)
5	0.89 (0.60,1.3)	0.68 (0.52,0.88)	0.88 (0.64,1.2)	0.68 (0.48,0.97)
6	1.1 (0.73,1.6)	0.63 (0.49,0.81)	0.83 (0.61,1.1)	0.47 (0.33,0.67)
Number of Coauthors				
<4	Reference	Reference	Reference	Reference
4	1.2 (0.88,1.6)	2.1 (1.7,2.6)	0.77 (0.60,0.99)	1.4 (1.1,1.8)
Region				
Americas	Reference	Reference	Reference	Reference
Europe	0.69 (0.52,0.92)	1.1 (0.90,1.3)	0.60 (0.48,0.75)	0.94 (0.72,1.2)
Others	0.34 (0.12,0.92)	0.43 (0.23,0.81)	0.67 (0.26,1.7)	0.78 (0.37,1.6)
West Pacific	0.74 (0.47,1.2)	0.92 (0.66,1.3)	0.93 (0.65,1.4)	1.1 (0.74,1.7)

Table 4

Predictors of total citation count among 4,149 articles in six epidemiology journals from 2008–2012, adjusting for all other factors as well as a proxy measure of years in the field

Category	Adjusted β (95% CI)	Adjusted with sensitivity analysis to account for h-index β (95% CI)
First and Last Author		
Gender		
MM	Reference	Reference
MF	−6.5 (−11, −2.5)	−6.4 (−11, −2.3)
FM	−7.0 (−9.8, −4.1)	−6.9 (−10, −3.7)
FF	−5.7 (−9.0, −2.5)	−5.6 (−9.2, −2.1)
Any Unknown	−8.0 (−12, −4.3)	−8.0 (−12, −3.9)
Publication Year		
2008	Reference	Reference
2009	−5.0 (−8.4, −1.6)	−5.0 (−8.7, −1.3)
2010	−9.6 (−13, −6.2)	−9.6 (−13, −6.1)
2011	−14 (−17, −10)	−14 (−17, −10)
2012	−20 (−23, −16)	−20 (−23, −16)
Field category		
Methods	Reference	Reference
Other	−0.1 (−2.5, 2.3)	−0.2 (−2.6, 2.3)
Journal		
1	Reference	Reference
2	−13 (−16, −9.3)	−13 (−17, −9.3)
3	−1.3 (−5.0, 2.5)	−1.3 (−5.1, 2.5)
4	−10 (−14, −5.8)	−9.9 (−14, −5.5)
5	3.6 (0.1, 7.1)	3.5 (−0.1, 7.2)
6	−4.9 (−8.4, −1.5)	−4.9 (−8.4, −1.4)
Number of Coauthors		
<4	Reference	Reference
4	3.1 (0.3, 5.8)	3.0 (−0.01, 5.9)
Region		
Americas	Reference	Reference
Europe	−1.4 (−4.0, 1.1)	−1.4 (−4.1, 1.2)
Others	−8.7 (−16, −1.2)	−7.9 (−20, 4.2)
West Pacific	−5.8 (−9.8, −1.9)	−5.8 (−10, −1.6)
Log of First Author h-index	NA	0.1 (−2.0, 2.2)
Log of Last Author h-index	NA	0.2 (−8.2, 8.5)

CI, confidence interval; FF, female first and female last author; FM, female first and male last author; MF, male first and female last author; MM, male first and male last author

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript