nature human behaviour

LETTERS

https://doi.org/10.1038/s41562-017-0235-x

One and a half million medical papers reveal a link between author gender and attention to gender and sex analysis

Mathias Wullum Nielsen[1]*, Jens Peter Andersen[2], Londa Schiebinger[1] and Jesper W. Schneider[2]

Gender and sex analysis is increasingly recognized as a key factor in creating better medical research and health care[1–7]. Using a sample of more than 1.5 million medical research papers, our study examined the potential link between women's participation in medical science and attention to gender-related and sex-related factors in disease-specific research. Adjusting for variations across countries, disease topics and medical research areas, we compared the participation of women authors in studies that do and do not involve gender and sex analysis. Overall, our results show a robust positive correlation between women's authorship and the likelihood of a study including gender and sex analysis. These findings corroborate discussions of how women's participation in medical science links to research outcomes, and show the mutual benefits of promoting both the scientific advancement of women and the integration of gender and sex analysis into medical research.

Despite a burgeoning scholarship, gender and sex differences remain unaddressed in large areas of medical research. Consider, for instance, the paradigmatic example of biological sex differences in cardiovascular disease — the leading cause of death among adult populations in the Western world. Since the 1980s, the annual number of cardiovascular disease-related deaths has been higher for women than for men; however, women continue to be underrepresented as participants in clinical trials[8–10]. A similar pattern is found in cancer research, where male research subjects dominate the trials[11], or in studies of cell tissue and laboratory animals, in which the sex of the subject is often unrecorded[12,13]. Even in research that involves both sexes, data are often not analysed by sex, and results may therefore not be accurate for either male or female subjects[14]. Both women and men face disadvantages under these circumstances. Around one-third of osteoporosis-related hip fractures occur in elderly men, but osteoporosis research tends to focus on women, resulting in osteoporosis in men being underdiagnosed, undertreated and underreported[15].

Not only biological sex but also gender (that is, the social attitudes and behaviours associated with being a woman or a man) is known to be a crucial determinant of human health[8,9]. Women and men are exposed to different occupational hazards (for example, ergonomic demands and psychosocial stressors) and differ on a wide array of health-related lifestyle behaviours (such as exercise, alcohol and tobacco use)[16,17]. Furthermore, gendered psychosocial factors have been found to predict recurrent outcomes in patients with acute coronary syndrome and to moderate patient perceptions of pain[18,19]. Despite robust evidence, associations between gender, biological sex and health outcomes remain largely neglected in the literature[20], with potentially life-threatening and costly consequences. Of the ten drugs withdrawn from the US market between 1997 and 2000, eight involved health risks for women that may have been avoided if more attention had been devoted to gender-related and sex-related factors[21].

These examples demonstrate how gender and sex analysis (GSA) has the potential to improve medical diagnosis and treatment. Here, we define GSA as scientific approaches that are aimed at understanding how social and behavioural differences between women, men and gender-diverse people (gender analysis) and biological differences between female and male research subjects (sex analysis) relate to health outcomes.

A growing number of medical scholars and journal editors already acknowledge the importance of GSA in fostering excellence in medical science and health care[1–7]. Science agencies also subscribe to this idea (for an overview of national policy efforts, see ref. [22]). Through the Horizon 2020 programme, the European Commission has committed itself to "integrating the gender dimension into the research and innovation content (…) in order to improve the quality of research and stimulate innovation"[23]. Similarly, the US National Institutes of Health (NIH) recognize that both gender and sex "play a role in how health and disease processes differ among individuals," and have implemented guidelines recommending that "sex as a biological variable will be factored into research designs, analyses, and reporting in vertebrate animal and human studies"[24].

Both the European Commission and the NIH prioritize policies: (1) to advance the careers of women scientists, and (2) to encourage GSA in research design. However, we know little about how these policy objectives may be linked. Are women and men equally likely to integrate GSA into their research designs?

While recognizing the crucial role played by funding agencies and scholarly journals in putting gender and sex on the scholarly agenda, our study is unique in analysing potential gender differences in medical scholars' involvement in GSA. If gender variations are detected in our data, we are not proposing that they stem from innate differences in the scientific styles or preferences of women and men. Instead, we follow tenets in the literature on cultural diversity, and see gender as a cultural category that shapes cognition, experience and perspective in the workplace, with implications for research interests and focus[25–27].

Sociological research illustrates how societal norms and expectations operate to cultivate gender-differentiated career aspirations in higher education and in the labour market[28]. Existing research, for instance, documents clear gender differences in the selection of

[1]History of Science, Stanford University, Stanford, CA, USA. [2]Danish Centre for Studies in Research and Research Policy, Department of Political Science, Aarhus University, Aarhus, Denmark. *e-mail: mwn@ps.au.dk

fields and medical specialties by students[28,29]; however, gender norms and expectations may also spur variations in researchers' choice of different forms of research within a given field or specialty[27]. Indeed, this idea already finds some support in the literature. Using meta-data from more than 8 million articles in JSTOR, notable differences were observed in the primary areas of specialization by women and men in a range of scholarly disciplines spanning the natural sciences, social sciences and humanities[30]. Using data from Sociological Abstracts, another study found women sociologists to be overrepresented in research areas such as gender, race, family and medicine, whereas men dominated in political, comparative and economic sociology[31]. In a third study, notable gender variations were observed in the primary subfields in economic research: health, education, welfare and labour economics were popular topics among women, whereas agricultural economics, fluctuations and business cycles, general equilibrium, comparative systems and corporate finance were popular among men[32]. Finally, recent studies document a clear overrepresentation of women authors in social science journals on gender, feminism and sexuality[33,34].

Given the traditional notion of gender-related and sex-related research as a woman-dominated domain in academia, we expect the involvement of women researchers in GSA to be more congruent with prevailing gender norms and expectations than the involvement of men. Broader cultural influences about appropriate gender-typed work may, in other words, draw a disproportionate number of women towards (and men away from) this form of research. A pioneering study found that women medical investigators are more likely to address gender and sex in successful research proposals for the Canadian National Institutes of Health[35]. However, the results reported in this study are descriptive and do not adjust for potential spurious associations that result from differences in participation by women and men investigators across diseases and medical subfields.

Our study has been designed to accommodate these potential biases. Using a global sample of more than 1.5 million medical research papers and adjusting for variations in the gender composition of author teams across disease topics, countries and medical research areas, we compared women's general participation and share of first and last authorships in studies that do and do not involve GSA. Our prior conjectures were that:

(1) the likelihood of a study involving GSA increases with the proportion of women among its authors;
(2) the likelihood of a study involving GSA increases if the first author is a woman; and
(3) the likelihood of a study involving GSA increases if the last author is a woman.

We focused on disease-specific medical research using peer-reviewed scholarly articles published in the period 2008–2015 as our basic unit of analysis. To identify studies using GSA, we used the GenderMed Database[36,37], which is an extensive bibliographic archive of the existing medical literature analysing gender and sex differences. To determine the gender of authors, we used the name-to-gender assignment algorithm Gender API (https://gender-api.com/). The analysis took place in two steps. First, we used descriptive statistics to document global gender disparities in authorship participation in disease-specific medical research. Second, we used three logistic regression analyses to estimate how the overall representation of women in the author group and as first and last authors influenced the likelihood of a study systematically using GSA. All models were estimated in a Bayesian framework with weakly informed default prior distributions[38] (see Methods for more details).

Figure 1 provides descriptive information on women's representation per author group and share of first and last authorships in studies published from 2008 to 2015. Women comprise 40% of first authors, 27% of last authors and, on average, 35% of authors per author group in the full sample. Importantly, women's participation as first and last authors varies across geographical groupings (see Supplementary Table 1). Gender disparities are largest in East Asia (women first authors: 24%, women last authors: 16%) and smallest in Latin America (women first authors: 52%, women last authors: 40%). North American and Western European figures lie close to the global average, which is not surprising, given that 67% of first authors and 69% of last authors in our sample are affiliated with institutions in North America or Western Europe.

As a second step in the analysis, we found differences in women's (and men's) full-group participation and share of first and last authorships in studies that do and do not involve GSA. Table 1 shows the odds ratios, standard deviations and 95% credible intervals (CI) for the logistic regressions that predict the likelihood of a study using GSA (outcome variable: non-GSA = 0, GSA = 1) (see Methods for variable specifications). Model 1 and model 2 specify the effects that are attributable to female first and last authors, and model 3 determines the affect effect of women's overall representation in the author group. Conditioned on our model specifications and prior assumptions, women's participation positively predicts the likelihood of a study using GSA for all three groups. This is in accordance with our initial conjectures. Odds ratios are 1.66 (95% CI: 1.53–1.79) for first authors (f_first), 1.56 (95% CI: 1.44–1.68) for last authors (f_last) and 3.14 (95% CI: 2.74–3.59) for full-group representation (fw).

Figure 2 shows the estimated marginal means and 95% CI for the main predictors. The estimated marginal means enable us to specify mean-based differences in the estimated participation of women (and men) authors across GSA and non-GSA, while adjusting for covariation attributable to all other variables in the models. As illustrated in the figure, women comprise 40% of first authors in non-GSA studies and 49% in GSA studies. Differences in women's participation in non-GSA studies and GSA studies for last authors and full author groups are 27% versus 35% and 35% versus 42%, respectively. The relative difference is largest for last authors (30%) and smallest for full author groups (20%). This indicates that the

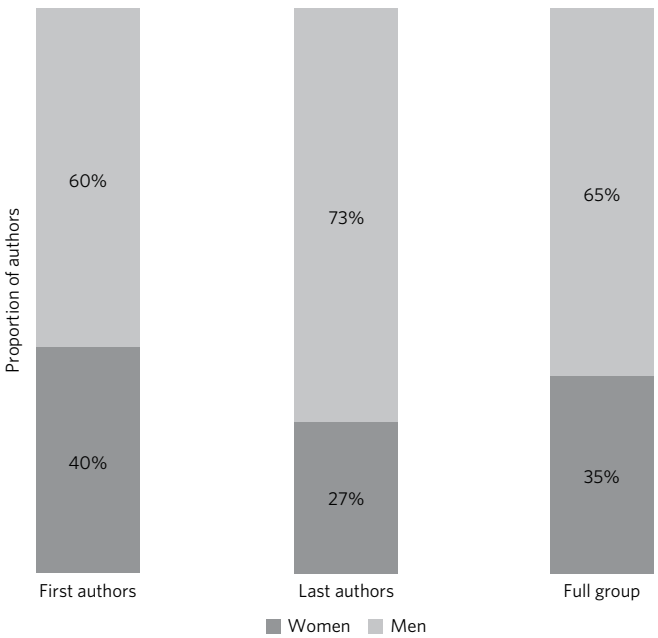

Fig. 1 | The global share of women as first authors, last authors and full-group participation in disease-specific medical research. Women comprise 40% of first authors (men 60%), 27% of last authors (men 73%) and 35% of authors overall per paper (men 65%). $n = 1{,}542{,}690$.

Table 1 | Binary logistic regression model predicting GSA

	Model 1 (first author)			Model 2 (last author)			Model 3 (full group)		
Parameter	**Odds ratio**	**s.d.**	**95% CI**	**Odds ratio**	**s.d.**	**95% CI**	**Odds ratio**	**s.d.**	**95% CI**
f_first	1.66	0.07	1.53–1.79						
f_first country	2.75	0.80	1.50–4.62						
f_first MeSH	6.13	1.67	3.55–10.00						
f_first SC	2.11	0.47	1.36–3.19						
f_last				1.56	0.06	1.44–1.68			
f_last country				4.03	1.31	2.02–7.13			
f_last MeSH				4.76	1.55	2.48–8.49			
f_last SC				4.32	1.09	2.60–6.84			
fw							3.14	0.22	2.74–3.59
fw country							2.19	0.85	0.99–4.26
fw MeSH							4.86	1.58	2.55–8.58
fw SC							1.98	0.52	1.16–3.18
Arab States	2.31	0.46	1.55–3.34	2.41	0.48	1.62–3.48	2.26	0.45	1.50–3.26
East Asia	1.79	0.22	1.40–2.26	1.96	0.25	1.54–2.49	1.90	0.25	1.47–2.44
Latin America	1.19	0.19	0.85–1.60	1.31	0.20	0.95–1.76	1.20	0.19	0.86–1.62
Oceania	1.25	0.18	0.93–1.66	1.49	0.22	1.11–1.97	1.33	0.20	1.00–1.75
South and West Asia	1.27	0.21	0.91–1.73	1.33	0.23	0.96–1.83	1.29	0.22	0.99–1.77
South-Central and Eastern Europe	1.36	0.20	1.00–1.80	1.50	0.21	1.10–1.97	1.35	0.20	0.92–1.77
Sub-Saharan Africa	3.03	0.58	2.01–4.28	3.26	0.62	2.19–4.61	3.17	0.60	2.13–4.48
North America	2.08	0.22	1.69–2.56	2.37	0.25	1.93–2.90	2.12	0.22	1.72–2.60
Western Europe	2.00	0.22	1.62–2.46	2.53	0.28	2.05–3.12	2.16	0.23	1.75–2.65
$n = 1{,}513{,}638$									

Posterior summaries of odds ratios with 95% CI for Bayesian logistic regression models with Cauchy informative priors predicting GSA. Commonwealth of Independent States is the reference group for the geographical variables. For more model specifications, see Supplementary Tables 2–4.

effect attributable to women's participation is strongest when women serve as leaders of the author group.

In comparison, logistic regression models that excluded all variables but the main predictors (f_first, f_last and fw) have odds ratios of 1.91 (95% CI: 1.77–2.06) for first authors, 1.76 (95% CI: 1.63–1.89) for last authors and 3.87 (95% CI: 3.43–4.36) for full author groups (for model specifications, see Supplementary Tables 6–8). Furthermore, using arithmetic means, differences in the average participation of women authors in non-GSA and GSA studies are 40% versus 53% for first authors, 27% versus 37% for last authors and 35% versus 45% for full author groups.

Two control variables in models 1, 2 and 3 also deserve attention. As shown in Table 1, GSA studies are more likely to be carried out in disease-specific research areas with a high general representation of women. Odds ratios for this factor are 6.13 (95% CI: 3.55–10.00) for first authors (f_first MeSH), 4.76 (95% CI: 2.48–8.49) for last authors (f_last MeSH) and 4.86 (95% CI: 2.55–8.58) for full author groups (fw MeSH). Women's representation in the broader research areas (or specialties) circumscribing a given disease topic also has a positive effect. Odds ratios for this factor are 2.11 (95% CI: 1.36–3.19) for first authors (f_first SC), 4.32 (95% CI: 2.60–6.84) for last authors (f_last SC) and 1.98 (95% CI: 1.16–3.18) for full author groups (fw SC). This indicates that not only the gender composition of author groups but also the general gender composition of the context in which they operate play an important role in predicting a study's focus on GSA. However, we do not know whether this finding results from the fact that women tend to more often specialize in topics, in which disease-specific gender-related and sex-related variations are likely to exist.

In summary, our results provide global evidence linking the likelihood of a study involving GSA to the presence of women in the author group, especially in leading positions as first and last authors.

Canadian researchers have already documented gender differences in funding applicants' attention to GSA; however, their analysis did not adjust for potential spurious associations resulting from differences in women's and men's participation across diseases and medical subfields[35]. We show that adjusting for such factors (and geographical variations) leads to more robust comparisons. Specifically, the odds ratios for the main predictors (f_first, f_last and fw) are reduced from 1.91 to 1.66 for first authors, 1.76 to 1.56 for last authors and 3.87 to 3.14 for full author groups, when covariates adjusting for women's participation across geographical groupings, disease topics and medical subfields are factored into the logistic regression models.

Furthermore, our analysis documents notable global gender disparities in authorship participation. Women comprise 40% of first authors, 27% of last authors and 35% of the authors per author group in disease-specific research. Given the findings of this paper, the modest share of women last authors is particularly troublesome, as last authors typically take the lead in identifying, planning and developing research questions in the health-related fields.

Our study adds to the existing literature in several important ways. First, it empirically links the efforts of science policy-makers to increase the numbers of women in academic medicine and to promote GSA, therefore providing a strong argument for both policy objectives. In this regard, it is important to note that GSA constitutes a small subset of the total number of articles published in the

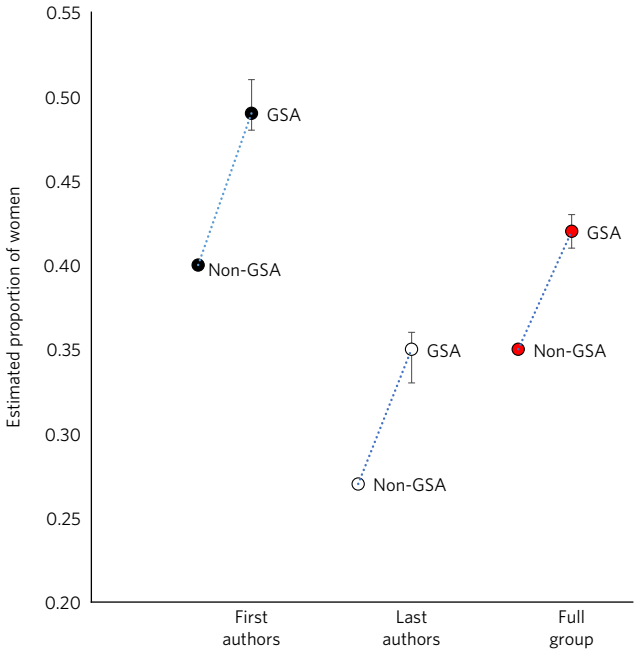

Fig. 2 | Plot of estimated marginal means. The estimated marginal means for f_first, f_last and fw in models 1, 2 and 3 are shown. Error bars represent 95% CI (for estimate specifications, see Supplementary Table 5). The plots visualize the participation of women (relative to men) as first authors, last authors and overall representation in the byline for studies that do and do not involve GSA. The figure shows that women's estimated share of authorships is higher in GSA studies than in non-GSA studies for all three author variables.

disease-specific medical literature. The majority of women and men are not engaged in this form of research.

Second, our findings highlight the importance of devoting more systematic attention to the link between gender diversity and research outcomes in academia. In a recent opinion piece in *Proceedings of the National Academy of Sciences*, the NIH Director of Scientific Workforce Diversity Hannah Valantine and the NIH Director Francis Collins encouraged the scientific community to develop approaches that are specifically designed to document the effect of diversity on the quality and outputs of academic medicine[39]. Our study pushes the research agenda forward by empirically demonstrating how gender diversity can expand health solutions by diversifying research methods to include GSA. Thus, expanding gender equality may have broader implications for knowledge and health outcomes than previously suspected.

Future research might investigate whether similar relationships between the participation of women in research and the likelihood of including GSA can be detected in other fields. Evidence from the social sciences, where women are relatively well-represented, suggests that this may be the case[31,33,34]; but in the STEM (science, technology, engineering and mathematics)-related areas, such as engineering and computer science, where GSA is less prevalent and men dominate research teams, this question remains unexplored. However, there is no reason to believe that the situation in STEM would be any different from the medical sciences.

Like any other study, ours is characterized by certain caveats and limitations. First, the search algorithm used to establish the corpus of GSA studies may, despite a systematic and thorough screening strategy, not capture the full range of potentially relevant publications[36,37]. Consequently, an unspecified number of 'false negatives' (that is, eligible GSA studies that are not specified as such) may exist in our baseline group of non-GSA studies. However, as GSA-related research constitutes a small subset of the total number of articles published in the disease-specific medical literature, we do not expect this to influence our results in any noteworthy way. Although if this were the case, it would probably imply even stronger support for our conjectures, given the higher level of participation by women in GSA studies.

Second, status-related factors, which are not addressed in this paper, may play a part in explaining our findings. We know from social science research that when women enter male-dominated fields, subtle forms of gender segregation tend to persist. Women may, for instance, self-select or be 'ghettoized' into less competitive and prestigious work areas[40,41]. To address such issues, future studies could investigate whether the relative status of GSA and funding opportunities (or lack hereof) are pushing a disproportionate number of women towards (and men away from) this form of research. In a supplementary analysis (see Supplementary Methods), we take a first step towards revealing potential status differentials. Specifically, we compare the distribution of GSA and non-GSA publications according to journal impact. Despite a vast array of problems with such measures, metrics of journal impact are widely used as proxies of achievement and status in the medical sciences[42,43], and most medical scholars acknowledge the existence of a journal hierarchy in their disciplines. As demonstrated in Supplementary Table 9 and Supplementary Fig. 1, our data provide no indications of any systematic status variations between the journals in which GSA and non-GSA studies are published, but more research is needed to reveal the influence of other potential status-related and resource-related differences.

Finally, engaging GSA requires expertise. Anyone, irrespective of gender, can be trained to do this effectively. Indeed, granting agencies and universities have begun to initiate workshops and incentive programmes to integrate GSA into research design, and future studies could investigate the effect of such programmes in enhancing the external validity and applicability of scientific research with human outcomes.

Our study establishes an empirical link between gender diversity in the scientific workforce and research outcomes. Our findings show a symbiotic relationship between increasing the numbers of women in academic medicine and enhancing excellence in research by incorporating GSA. Hence, our study provides empirical evidence for science policy-makers to promote both the scientific careers of women and GSA in research design. Taken together, these objectives support the twin goals of diversity and excellence in science.

Methods

Data. Data for this study were harvested from PubMed's MEDLINE. MEDLINE is one of the most exhaustive databases of medical journal literature and offers a systematic hierarchy of Medical Subject Headings (MeSH) for indexing scholarly articles, for example, on the basis of disease topics. We gathered bibliographic metadata for all Medline articles with disease-specific MeSH terms for the period 2008–2015, which resulted in 2,512,371 records (see Supplementary Methods). To obtain information on author first names and country of institutional affiliation, 2,124,998 of these records were matched to records in the citation database Web of Science (WoS) (see Supplementary Methods). This information was used to determine the gender of authors using the name-to-gender assignment algorithm Gender API (Gender API. Gender API — determines the gender of a first name. *Gender API* https://gender-api.com/ (2016)) (see Supplementary Methods). For each first name and nation pair, Gender API provided an estimate specifying the certainty of the given name-to-gender assignment. To validate its accuracy, we conducted a manual quality-control check of 500 randomly selected authors from the data set (see Supplementary Methods and Supplementary Table 10). Papers that lacked full first name information for one or more authors were excluded from the analysis. Supplementary Fig. 2 details the data inclusion and exclusion steps, including the assignment of gender to author first names, leading to the final sample of 1,542,690 documents (61.4% of the total population). Supplementary Fig. 3 shows the distribution of fw scores as a function of the number of authors per paper in the data set.

To identify studies using GSA, we matched the WoS sample with the GenderMed Database[36,37], which is an extensive bibliographic archive of the existing medical literature analysing gender and sex differences. Based on a

screening of more than 13 million MEDLINE abstracts and 40,000 full-text manuscripts, approximately 13,000 studies, extending back to 1975, have been made available (see Supplementary Methods). GenderMed includes 4,830 studies with MeSH terms subordinate to a disease category for the period 2008–2015, of which 3,394 (70.3% of the population) matched with our WoS sample (see Supplementary Methods). A manual quality check of 500 GenderMed articles was carried out to verify the accuracy of the database in identifying research involving GSA (see Supplementary Methods).

The GenderMed Database limits its scope to selected diseases that field experts have deemed epidemiologically relevant for GSA[37]. Thus, we excluded all studies in our final sample ($n=1{,}542{,}690$) that did not overlap with studies in the GenderMed subsample ($n=3{,}394$) with respect to disease-specific MeSH terms. This exclusion resulted in a reduced sample of 1,513,638 unique disease-specific papers, which was used in the logistic regression analyses (see Supplementary Methods).

Outcome variable. GSA is the binary outcome variable in models 1, 2 and 3. The variable is used to predict the likelihood of a study systematically using GSA. Specifically, it dissociates the GenderMed sample from the remaining articles in the WoS database (non-GSA = 0, GSA = 1).

Main predictors and covariates. For each article in the full data set ($n=1{,}542{,}690$), we computed the weighted indicator fw to specify the general participation of women authors in a given study. fw values range from 0 to 1, with values closer to 1 indicating a higher proportion of women in the author group (see Supplementary Methods). To account for within-group variations in women's (and men's) representation across countries, disease topics and medical subject areas, we also computed average fw scores for the following variables: MeSH disease terms (fw MeSH), WoS subject categories classified based on journal information (fw SC) and last author country (fw country). fw MeSH was included to adjust for differences in the average gender composition of authors across disease-specific research topics, while fw SC was constructed to account for covariation attributable to the varying participation of women across medical research areas and specialties (See Supplementary Methods). The control variable fw SC is important, as a given disease term may be addressed by researchers in different research areas and specialties. A complementary method using MeSH disease terms to capture medical research specialties was examined but did not yield any improvements compared with the approach based on the WoS subject categories[44].

Based on Gender API estimates, we also calculated gender indices for first and last authors (f_first and f_last). Furthermore, we computed average f_first and f_last scores for MeSH disease terms (f_first MeSH and f_last MeSH), WoS subject categories (f_first SC and f_last SC) and author country (f_first country and f_last country) (see Supplementary Methods). An additional analysis was carried out to estimate the prevalence of non-traditional author listings based on alphabetical order, as these may bias the odds-ratio estimations for the key predictors f_first and f_last in the analysis. We found the intentional use of alphabetized author listings to be very rare and unlikely to influence our results in any notable way (see Supplementary Methods).

Additional control variables were constructed to help prevent potential misspecifications in the logistic regression models. To capture the influence of cultural and socioeconomic factors across area-specific contexts, we included ten categorical variables based on geographical groupings (for specifications on country groupings see Supplementary Table 13). Furthermore, author country information was used to compute ranked variables capturing covariation that is attributable to national differences in the overall level of gender equality and annual research and development expenditure related to health, but these variables were excluded from the final models due to trivial variance (see Supplementary Methods) (variable specifications are available in Supplementary Table 14).

Statistical models. All models were estimated in a Bayesian framework with weakly informed default prior distributions[38]. A Bayesian framework is transparent and enables flexible estimation of parameters based on prior information and data. Proper uncertainties in estimates were calculated, which provided intuitive interpretations and direct inferences of hypotheses conditioned on the data. A simple Bayesian logistic modelling approach to estimate posterior distributions of the coefficients was applied. A basic generic binominal model with a logit link function was used for all specifications. Let Y_i be the outcome variable for observation i which takes the value of either 0 or 1. The stochastic component of the model is given by

$$Y_i \sim \text{Bernoulli}(\pi_i) = \pi_i^{Y_i}(1-\pi_1)^{1-Y_i}, \quad (1)$$

where $\pi_1 = \Pr(Y_1 = 1)$

The systematic component is given by

$$\pi_i = \frac{1}{1+\exp(-x_i\beta)},$$

where x_i is the vector of k covariates for observation i and β is the vector of coefficients.

We assigned default independent Cauchy prior distributions with location parameter 0 (μ) and scale parameter 2.5 (σ) to all coefficients $\beta_0, \beta_i\, i = 1, \ldots, k$. This prior is proposed as a default choice because it is weakly informative, but allows inferences to be made even in the presence of complete separation. Given the size of the data sets used in models 1, 2 and 3, the expected influence of the weakly informed priors on the estimated posterior parameters will be less pronounced. The posterior distributions were estimated empirically by Markov Chain Monte Carlo (MCMC) techniques using the R package 'MCMCpack'[45,46]. All models used 5,000 burn-in and 50,000 Metropolis iterations for the sampler, with a tuning parameter of 0.25. To summarize and visualize findings from the three models, we estimated marginal means for the main predictors using the 'lsmeans' package in R[47]. The descriptive analysis of gender disparities in authorship was based on the full sample of 1,542,690 papers, whereas the reduced sample of 1,513,638 was used in the regression models.

Code availability. The code used to generate the logistic regression analysis and estimated marginal means is available at https://osf.io/r6hd9/

Data availability. All data that support our main findings are publicly available at https://osf.io/r6hd9/

Received: 5 May 2017; Accepted: 27 September 2017;
Published online: 6 November 2017

References

1. Arnold, A. P. Promoting the understanding of sex differences to enhance equity and excellence in biomedical science. *Biol. Sex Differ.* **1**, 1 (2010).
2. Schiebinger, L., Leopold, S. S. & Miller, V. M. Editorial policies for sex and gender analysis. *Lancet* **388**, 2841–2842 (2016).
3. *Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals* (International Committee of Medical Journal Editor, 2016); http://www.icmje.org/news-and-editorials/icmje-recommendations_annotated_dec16.pdf.
4. Heidari, S., Babor, T. F., De Castro, P., Tort, S. & Curno, M. Sex and gender equity in research: rationale for the SAGER guidelines and recommended use. *Res. Integr. Peer Rev.* **1**, 2 (2016).
5. Miller, V. M. In pursuit of scientific excellence: sex matters. *Physiol. Genomics* **44**, 485–486 (2012).
6. Nieuwenhoven, L. & Klinge, I. Scientific excellence in applying sex- and gender-sensitive methods in biomedical and health research. *J. Womens Health* **19**, 313–321 (2010).
7. Johnson, J. L., Greaves, L. & Repta, R. Better science with sex and gender: facilitating the use of a sex and gender-based analysis in health research. *Int. J. Equity Health* **8**, 14 (2009).
8. Oertelt-Prigione, S. & Regitz-Zagrosek, V. (eds) *Sex and Gender Aspects in Clinical Medicine* (Springer, London, 2012).
9. Kim, E. S. H. & Menon, V. Status of women in cardiovascular clinical trials. *Arterioscler. Thromb. Vasc. Biol.* **29**, 279–283 (2009).
10. Mosca, L., Hammond, G., Mochari-Greenberger, H., Towfighi, A. & Albert, M. A. Fifteen-year trends in awareness of heart disease in women results of a 2012 American Heart Association national survey. *Circulation* **127**, 1254–1263 (2013).
11. Kwiatkowski, K., Coe, K., Bailar, J. C. & Swanson, G. M. Inclusion of minorities and women in cancer clinical trials, a decade later: have we improved? *Cancer* **119**, 2956–2963 (2013).
12. Beery, A. K. & Zucker, I. Sex bias in neuroscience and biomedical research. *Neurosci. Biobehav. Rev.* **35**, 565–572 (2011).
13. Shah, K., McCormack, C. E. & Bradbury, N. A. Do you know the sex of your cells? *Am. J. Physiol. Cell Physiol.* **306**, C3–C18 (2014).
14. Klein, S. L. et al. Sex differences in the incidence and case fatality rates from hemorrhagic fever with renal syndrome in China, 2004–2008. *Clin. Infect. Dis.* **52**, 1414–1421 (2011).
15. Adler, R. A. Osteoporosis in men: a review. *Bone Res.* **2**, 14001 (2014).
16. Smith, P. M. & Koehoorn, M. Measuring gender when you don't have a gender measure: constructing a gender index using survey data. *Int. J. Equity Health* **15**, 82 (2016).
17. Courtenay, W. H. Behavioral factors associated with disease, injury, and death among men: evidence and implications for prevention. *J. Mens Stud.* **9**, 81–142 (2000).
18. Alabas, O. A., Tashani, O. A., Tabasam, G. & Johnson, M. I. Gender role affects experimental pain responses: a systematic review with meta-analysis. *Eur. J. Pain* **16**, 1211–1223 (2012).
19. Pelletier, R. et al. Sex versus gender-related characteristics: which predicts outcome after acute coronary syndrome in the young? *J. Am. Coll. Cardiol.* **67**, 127–135 (2016).
20. Schiebinger, L. & Stefanick, M. L. Gender matters in biological research and medical practice. *J. Am. Coll. Cardiol.* **67**, 136–138 (2016).

21. US General Accounting Office *Drug Safety: Most Drugs Withdrawn in Recent Years had Greater Health Risks for Women* (Government Publishing Office, Washington DC, 2001).
22. Schiebinger, L. et al. *Sex and Gender Analysis Policies of Major Granting Agencies* (Gendered Innovations in Science, Health & Medicine, Engineering, and Environment, 2017); http://genderedinnovations.stanford.edu/sex-and-gender-analysis-policies-major-granting-agencies.html
23. *Gender Equality in Horizon 2020* Version 2 (European Commission, 2016); http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/gender/h2020-hi-guide-gender_en.pdf
24. *Consideration of Sex as a Biological Variable in NIH-Funded Research* (National Institutes of Health, 2015); https://grants.nih.gov/grants/guide/notice-files/NOT-OD-15-102.html.
25. Ely, R. J. & Thomas, D. A. Cultural diversity at work: the effects of diversity perspectives on work group processes and outcomes. *Adm. Sci. Q.* **46**, 229–273 (2001).
26. Page, S. E. *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies* (Princeton Univ. Press, Princeton, NJ, 2008).
27. Nielsen, M. W. et al. Opinion: gender diversity leads to better science. *Proc. Natl Acad. Sci. USA* **114**, 1740–1742 (2017).
28. Charles, M. & Bradley, K. Indulging our gendered selves? Sex segregation by field of study in 44 countries. *Am. J. Sociol.* **114**, 924–976 (2009).
29. Alers, M., van Leerdam, L., Dielissen, P. & Lagro-Janssen, A. Gendered specialities during medical education: a literature review. *Perspect. Med. Educ.* **3**, 163–178 (2014).
30. West, J. D., Jacquet, J., King, M. M., Correll, S. J. & Bergstrom, C. T. The role of gender in scholarly authorship. *PLoS ONE* **8**, e66212 (2013).
31. Light, R. in *Networks, Work, and Inequality* (ed. Mcdonald, S.) 239–268 (Research in the Sociology of Work Vol. 24, Emerald Group Publishing, Bingley, 2013).
32. Dolado, J. J., Felgueroso, F. & Almunia, M. Are men and women-economists evenly distributed across research fields? Some new empirical evidence. *SERIEs* **3**, 367–393 (2012).
33. Kretschmer, H., Kundra, R., Beaver, D. D. & Kretschmer, T. Gender bias in journals of gender studies. *Scientometrics* **93**, 135–150 (2012).
34. Söderlund, T. & Madison, G. Characteristics of gender studies publications: a bibliometric analysis based on a Swedish population database. *Scientometrics* **105**, 1347–1387 (2015).
35. Johnson, J., Sharman, Z., Vissandjee, B. & Stewart, D. E. Does a change in health research funding policy related to the integration of sex and gender have an impact? *PLoS ONE* **9**, e99900 (2014).
36. Oertelt-Prigione, S., Parol, R., Krohn, S., Preissner, R. & Regitz-Zagrosek, V. Analysis of sex and gender-specific research reveals a common increase in publications and marked differences between disciplines. *BMC Med.* **8**, 70 (2010).
37. Oertelt-Prigione, S., Gohlke, B. O., Dunkel, M., Preissner, R. & Regitz-Zagrosek, V. GenderMedDB: an interactive database of sex and gender-specific medical literature. *Biol. Sex Differ.* **5**, 7 (2014).
38. Gelman, A., Jakulin, A., Piitau, M. G. & Su, Y.-S. A weakly informative default prior distribution for logistic and other regression models. *Ann. Appl. Stat.* **2**, 1360–1383 (2008).
39. Valantine, H. A. & Collins, F. S. National Institutes of Health addresses the science of diversity. *Proc. Natl Acad. Sci. USA*. **112**, 12240–12242 (2015).
40. Gneezy, U., Niederle, M. & Rustichini, A. Performance in competitive environments: gender differences. *Q. J. Econ.* **118**, 1049–1074 (2003).
41. Reskin, B. F. & Roos, P. A. (eds) *Job Queues, Gender Queues: Explaining Women's Inroads into Male Occupations* (Temple Univ. Press, Philadelphia, PA, 1990).
42. Patel, V. M. et al. How has healthcare research performance been assessed? A systematic review. *J. R. Soc. Med.* **104**, 251–261 (2011).
43. Young, N. S., Ioannidis, J. P. A. & Al-Ubaydli, O. Why current publication practices may distort science. *PLoS Med.* **5**, e201 (2008).
44. Darmoni, S. J. et al. A MEDLINE categorization algorithm. *BMC Med. Inform. Decis. Mak.* **6**, 7 (2006).
45. Martin, A. D., Quinn, K. M. & Park, J. H. MCMCpack: Markov Chain Monte Carlo in R. *J. Stat. Softw.* **42**, 1–21 (2011).
46. Martin, A. D., Quinn, K. M. & Park, J. H. Package 'MCMC-pack' v.1.3-9 (R Foundation for Statistical Computing, 2017); https://cran.r-project.org/web/packages/MCMCpack/MCMCpack.pdf
47. Lenth, R. Package 'lsmeans' Version 2.2 (The Comprehensive R Archive Network, 2016); https://cran.r-project.org/web/packages/lsmeans/lsmeans.pdf

Acknowledgements

We thank S. Oertelt-Prigione and the Institute of Gender in Medicine, Charité-Universitätsmedizin Berlin, Germany, for data acquisition from the GenderMed database. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Competing interests

The authors declare no competing interests.

Author contributions

The research was designed by M.W.N. The database was constructed by J.P.A. and M.W.N. The data were analysed by M.W.N., J.W.S. and J.P.A. All authors contributed to writing the paper.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41562-017-0235-x.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to M.W.N.

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