BMJ Open

# Longitudinal analyses of gender differences in first authorship publications related to COVID-19

Carolin Lerchenmüller ,[1,2] Leo Schmallenbach ,[3] Anupam B Jena,[4,5,6] Marc J Lerchenmueller [3]

**To cite:** Lerchenmüller C, Schmallenbach L, Jena AB, *et al*. Longitudinal analyses of gender differences in first authorship publications related to COVID-19. *BMJ Open* 2021;**11**:e045176. doi:10.1136/bmjopen-2020-045176

► Prepublication history and additional material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2020-045176).

ABJ and MJL contributed equally.

Received 27 September 2020
Revised 01 March 2021
Accepted 05 March 2021



For numbered affiliations see end of article.

**Correspondence to**
Dr Carolin Lerchenmüller; carolin.lerchenmueller@med.uni-heidelberg.de

## ABSTRACT

**Objective** Concerns have been raised that the COVID-19 pandemic has shifted research productivity to the disadvantage of women in academia, particularly in early career stages. In this study, we aimed to assess the pandemic's effect on women's COVID-19-related publishing over the first year of the pandemic.

**Methods and results** We compared the gender distribution of first authorships for 42 898 publications on COVID-19 from 1 February 2020 to 31 January 2021 to 483 232 publications appearing in the same journals during the same period the year prior. We found that the gender gap—the percentage of articles on which men versus women were first authors—widened by 14 percentage points during the COVID-19 pandemic, despite many pertinent research fields showing near equal proportions of men and women first authors publishing in the same fields before the pandemic. Longitudinal analyses revealed that the significant initial expansions of the gender gap began to trend backwards to expected values over time in many fields. As women may have been differentially affected depending on their geography, we also assessed the gender distribution of first authorships grouped by countries and geographical areas. While we observed a significant reduction of the shares of women first authors in almost all countries, longitudinal analyses confirmed a resolving trend over time.

**Conclusion** The reduction in women's COVID-19-related research output appears particularly concerning as many disciplines informing the response to the pandemic had near equal gender shares of first authorship in the year prior to the pandemic. The acute productivity drain with the onset of the pandemic magnifies deep-rooted obstacles on the way to gender equity in scientific contribution.

## Strengths and limitations of this study

- The COVID-19 pandemic is an exogenous source of variation that allows the examination of differential effects of the pandemic on women's and men's publishing activity.
- We used a retrospective cohort design, comparing author gender for COVID-19 articles with articles in similar fields published during the year prior to the pandemic.
- Data on affiliations, publishing journals and dates enable analyses of gender differences in publication rates by geography, scientific discipline and over time.
- This large-scale archival study did not allow disentangling the mechanisms that underpin gender differences in publishing rates associated with the pandemic.
- The methodology relied on a probabilistic algorithm to assign gender to thousands of authors, bearing a residual risk of gender misclassification.

## INTRODUCTION

Women are integral to productive and innovative science communities.[1 2] Nonetheless, women remain under-represented in prestigious author positions on publications in the life sciences and medicine,[3] are less likely to be promoted to higher academic ranks and are paid less, despite the continuously growing number of women academics.[4] Projections indicate that this gap will persist if targeted interventions are not implemented.[5]

It is in this setting that concerns have been raised that research and expert reporting on the COVID-19 pandemic has disproportionately involved men as scientific authors. For example, women submitted fewer manuscripts overall, were less available for peer review[6–8] and attended fewer funding panel meetings.[7] Also, women first authorship was significantly reduced on preprints and publications about COVID-19 in the USA[9 10] and globally.[11]

It has been suggested that this might, at least in part, be due to an exacerbation of pre-existing work–family conflicts, especially for early-career mothers in academia.[12] With lockdown measures to prevent uncontrolled spread of the coronavirus came not only remote working, but also closures of childcare services like daycares and schools transferring teaching responsibilities often to mothers, without the possibility to involve family members in childcare that belong to the ageing population who are particularly

vulnerable to severe illness from the coronavirus.[12–16] Evidence from Germany showed, for example, that women not only took over the physical load of increased childcare and household responsibilities, but also the mental load associated with taking care of the family during a pandemic.[17] Even prior to the pandemic, research has shown that childrearing and household work were tasks largely taken care of by women, thereby impacting women's academic careers more than men's. For example, parental leaves taken by men often result in increased productivity, while no such phenomenon can be observed for women.[18]

The proliferation of COVID-19-related publications provides a unique window into these gendered dynamics for two reasons. First, COVID-19 publications have been produced rapidly under unusual conditions that likely disfavour women scientists relative to usual conditions that can serve as a control. Second, COVID-19 publications are mainly, though not exclusively, produced in the life sciences and medicine where long-standing authorship norms reserve the first author position to early career investigators leading the project, which allows estimating the repercussions for this group in particular. Therefore, we assessed the pandemic's effect on women's COVID-19-related scientific publishing over the first year of the pandemic by analyses of first authorships in a longitudinal approach. We further performed analyses to quantify the effect per scientific specialty and country affiliation, as women may have been differentially affected across specialties and geographical areas.

## METHODS

### Study design and data

We use a retrospective cohort design, comparing the gender composition on the author byline of 42 898 PubMed indexed life science articles that included the term 'COVID' in the title and/or abstract, and that were published between February 2020 and January 2021 to a set of 483 232 control articles published in the same journals a year earlier (see online supplemental material 1 for details on data and methods). The onset of the COVID-19 pandemic in early 2020 (and ensuing countermeasures like lockdowns, remote work and so on) serves as a source of exogenous variation that affects authors of COVID-19 articles but not authors who published research in the same journals prior to the COVID-19 outbreak, that is, our control group. To assess possible effect stratification across research areas, we allocated 3426 represented journals to scientific disciplines based on the disciplines provided in the Clarivate Journal Citation Report using unique International Standard Serial Numbers as a crosswalk. We restricted this analysis to disciplines with at least 50 publications per reported time period to provide more precise estimates. We obtained detailed affiliation records to determine the geographical locale of first authors for country-specific analyses. We restricted this analysis to countries with at least 50 publications, and continents with at least 10 publications per reported time point, to increase precision of estimates (see online supplemental figure 1 for details on the sample construction).

We further made use of a long-standing authorship norm in the life sciences, according to which the first author is usually the junior author who executed the research, while the last author is generally the senior author who funded and may have conceived of the research. To designate the probable gender of thousands of these authors in our dataset, we use the genderize.io database that draws on a number of official sources, like Social Security Administration records and social media profiles, to assign a probability that a given forename is more likely held by men or women. For our analysis, we only included cases where the algorithm assigned a 90% or greater probability to the individual being of a specific gender (see also online supplemental figures 2–4). Overall, our applied inclusion criteria did not introduce tangible selection bias in terms of fields or countries represented (see online supplemental tables 1 and 2).

### Outcome measures

We calculated the gender gap in academic authorships as the absolute percentage point difference between men and women authors. For example, if men and women accounted for 55% and 45% of first author positions, respectively, the absolute gender difference would be 10 percentage points. We also offered parametric analyses in the supplement, analysing the effect of authoring during the pandemic versus not (a binary independent variable) on the likelihood that the first author was a woman versus a man (our outcome) (online supplemental tables 3–5).

### Sensitivity analyses

We conducted additional sensitivity analyses (online supplemental tables 6 and 7), including varying the confidence with which gender could be inferred, excluding articles with group authorships, rerunning our analyses for the full set of articles (ie, without applying sampling restrictions) and comparing sampled with non-sampled observations.

### 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

### One-year gender differences in first and last authorships related to COVID-19 publications

On average, men accounted for 54.9% and women for 45.1% of first authorships in 483 232 articles, published

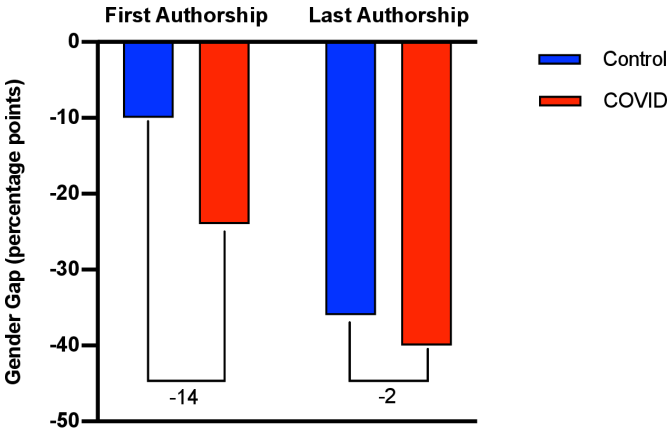



**Figure 1** Overall gender differences in first and last authorships related to COVID-19 publications. Gender gap in first and last authorships for COVID-19 publications (red) compared with the gender gap for control publications appearing in the same journals a year earlier (blue).

before the pandemic (between February 2019 and January 2020), for an absolute gender gap of 9.8 percentage points. In contrast, men and women accounted for 62.3% and 37.7% of first authorships on COVID-related publications, for an absolute gender gap of 24.6 percentage points. The gender gap therefore widened by approximately 14.8 percentage points in disciplines related to the COVID-19 pandemic (figure 1). Smaller changes were observed in last authorships. For example, while the gender gap in last authorships before the pandemic was expectedly much larger (approximately 36 percentage points), the effect of the pandemic on last author publishing by women was less pronounced (although statistically significant given the large sample size). On average, women accounted for 31.9% of last authorships in articles published before the pandemic, compared with 30.5% of last authorships for COVID-19-related publications, representing a widening of the last author gender gap by approximately 2.8 percentage points (figure 1). In line with previous data, our results therefore indicate that junior women investigators were disproportionally affected by the pandemic.[9–11 13 15] As women may have been differently affected depending on their field of study or depending on geography, we further sought to investigate gender differences particularly in first authorships according to field of study and country.

### Longitudinal analysis of gender differences in first authorships on COVID-19 publications, by scientific discipline

As the first author gender gap in publications generally varies across fields,[19] we calculated the first author gender gap by discipline. We allocated the journals that published on COVID-19 to scientific fields based on the disciplines represented in the Clarivate Journal Citation Report.[20] As depicted in the first column of the heatmap in figure 2, disciplines producing most COVID-19-relevant publications had shares of first authorship pre-COVID-19 by women of 45% (ie, an average gender gap of approximately 10 percentage points across all disciplines). During the pandemic, however, publications related to COVID-19 had an average 1-year gender gap of approximately 24 percentage points (ie, the share of first authorships from women for publications related to COVID-19 was 38%), a deviation of 14 percentage points (figure 2). This effect was most prominent in the first months of COVID-19 publishing, from February to May of 2020, when the gender gap rose to 36 percentage points (corresponding to a share of women first authorships of 32%). In the following months from June to September 2020, the share of women first authors slowly increased again to an average of 37%, and to 41% from October 2020 to January 2021, reducing the gender gap to 26 and 18 percentage points, respectively. However, this still represented a significant deviation from the pre-pandemic gender gap of 8 percentage points.

Interestingly, many of the disciplines that produced most COVID-19 publications had equal or near-equal gender shares in the year before the pandemic. For example, in the fields of virology, immunology, infectious diseases and general/internal medicine, the shares of women first authors prior to COVID-19 were 50%, 52%, 48% and 44%, respectively. The share of women first authors in other relevant fields like public, environmental and occupational health was even higher (59%). In many relevant fields, the gender rift in first authorships for COVID-19-related publications widened significantly above the pre-COVID-19 average. For example, while women were more likely to be first authors on publications within public, environmental and occupational health (difference of 18 percentage points) before the pandemic, the gap changed by 30 percentage points so that women were now less likely to publish research within this field as first authors of COVID-19-related publications (gender gap of 12 percentage points). In biochemistry and molecular biology, the gender gap in first authorships increased by 34 percentage points, from 2 percentage points pre-COVID-19 to 36 percentage points for COVID-19 publications. In virology, the gender gap increased by 26 percentage points, from equal shares (no gap) to 26 percentage points for COVID-19 publications. Other COVID-19-relevant fields were much less prone to changes in the first author gender gap, for example, surgery, and cardiac and cardiovascular systems, in which the gender gap prior to COVID-19 was 44 percentage points compared with 46 percentage points for COVID-19-related studies, respectively (meaning women's first authorship shares were as low as 28% and 27%, respectively) (figure 2).

### Longitudinal analyses of gender differences in first authorships on COVID-19 publications, by affiliated geographical area

Since the pandemic has affected countries differently, we further performed analyses based on the country affiliation of the first author. Women's research productivity

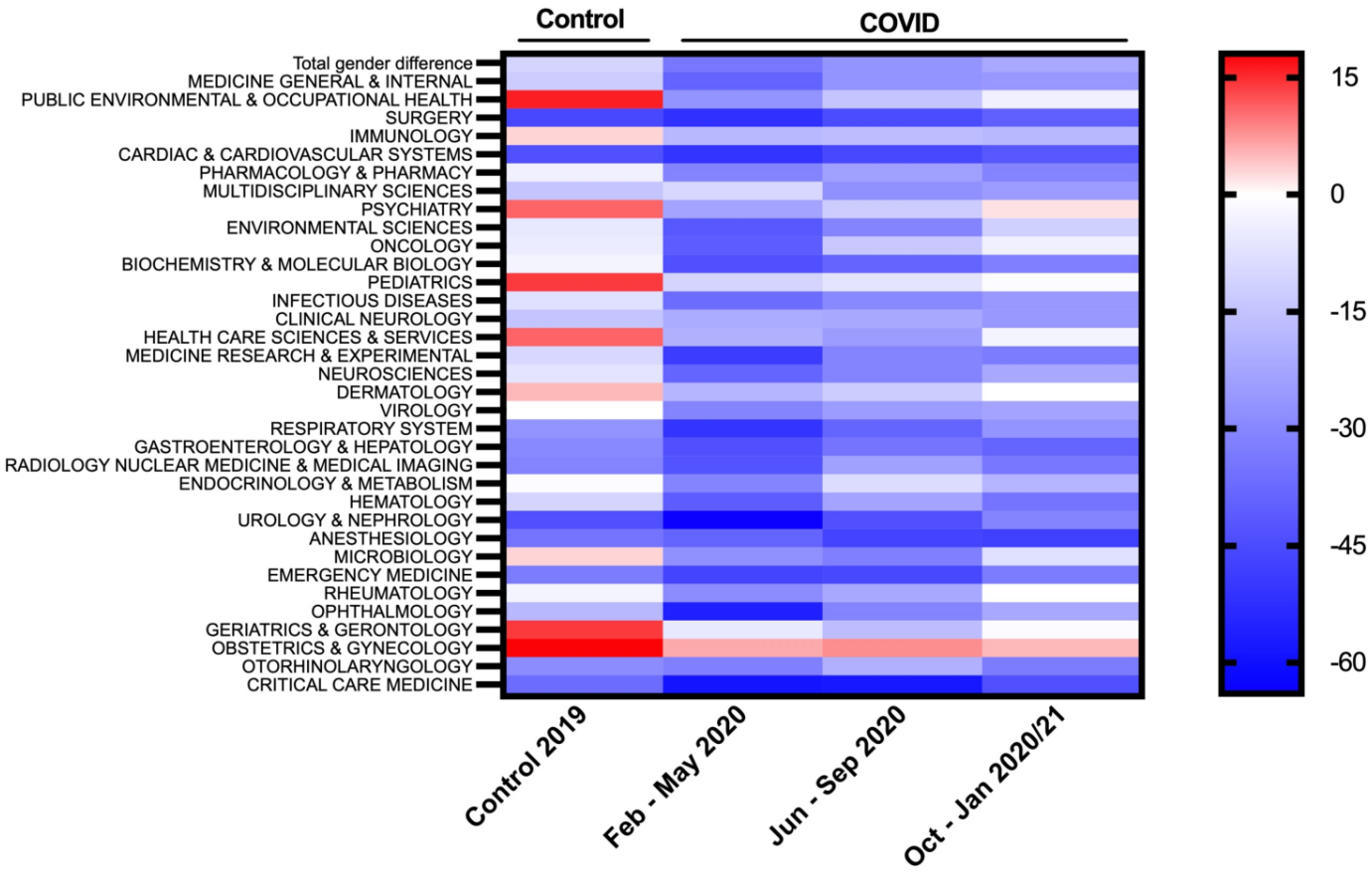



**Figure 2** Time-resolved gender differences in first authorship shares on COVID-19 publications, by scientific discipline. Heatmap depicting the gender gap in first authorships for COVID-19 publications and control publications from the same disciplines appearing in the same journals during the same period in the year prior to the pandemic. Fields sorted in descending order by number of publications. Red indicates an over-representation of women first authors, white indicates gender parity, blue indicates an over-representation of men first authors (in percentage points).

went down in almost all countries (figure 3A). For example, in the USA, which accounted for ~26% of all COVID-19 publications between February 2020 and January 2021, women's first authorship share decreased from a share of 44% to a share of 41% (corresponding to a widening of the gender gap by 6 percentage points). In European countries that were hit earlier by the pandemic than the USA, women's representation in authorships was also more affected. In Italy, for example, women's share of first authorships decreased from 49% before the pandemic to 35% for COVID-19-related publications, an increase in the gender gap of 28 percentage points because of the pandemic, with the overall number of publications from Italy accounting for 10% of total publications on COVID-19. The increase in the first authorship gender gap was also substantive in Brazil (30 percentage points), and Mexico (35 percentage points), Australia (14 percentage points) and India (22 percentage points). Only very few countries showed no change in the first authorship gender gap, including China (no change), South Korea (decreased by 3 percentage points) or Taiwan (decreased by 2 percentage points).

We further performed a granular time-resolved (per 2 months) analysis of women's first authorship shares grouped by continents. Our data showed that the largest reduction in women's first authorship shares happened early in last spring (April and May 2020). In Europe, for example, the gender gap increased by 18 percentage points, in North America by 8 percentage points, in Latin America by 28 percentage points, in Australia and Oceania by 15 percentage points, in Africa by 18 percentage points, and in Asia by 7 percentage points (figure 3B). Similar to our analysis of the difference in first authorship shares over time by field, we found that the gap began to close again over time and seemed to have reached baseline levels in North America and Oceania. In all the other geographical areas, the gender gap has yet to reach the levels expected from the year prior to the pandemic, with Africa and Latin America being farthest from the baseline (12 percentage points, respectively) (figure 3B), while COVID-19-related research output has been relatively stable from April 2020 to January 2021 within each geographic region.

## DISCUSSION

Our results provide evidence for the COVID-19 pandemic's effect on women's publishing productivity across

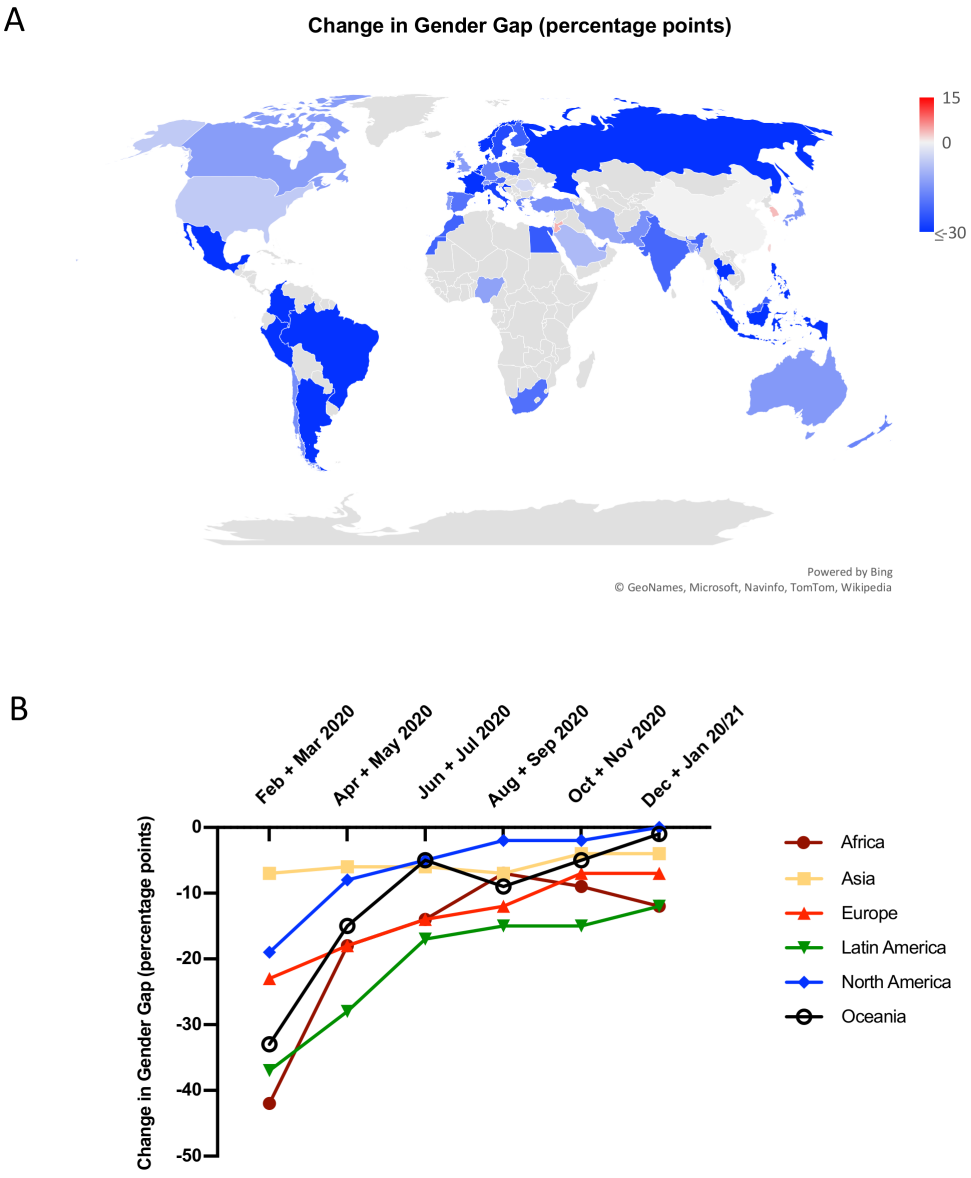



**Figure 3** Difference in first authorship gender gap, by country/geographical area. (A) World map depicting the deviation in the gender gap in first authorships for COVID-19 publications when compared with the expected gender gap derived from control publications from the same countries appearing in the same journals a year earlier. Red indicates an over-representation of women first authors, white indicates gender parity, blue indicates an over-representation of men first authors (in percentage points). (B) Time-resolved deviation in the gender gap in first authorships for COVID-19 publications when compared with the expected gender gap derived from control publications from the same geographical area appearing in the same journals a year earlier (in percentage points).

disciplines, worldwide and over time. In line with our hypothesis, we found that the relative increase in the gender gap was more pronounced for women in the first author position. In light of previous research and observations,[13 15 18 21] we suspected that the overcontribution of women to household and childrearing responsibilities—that leaves less opportunity to participate in writing, submitting and publishing research related to COVID-19—led to reduced productivity of early-career women investigators. However, aside from time constraints that disproportionally changed for women in an earlier stage of their career, other reasons are possible for the significant difference of women's productivity with regard to COVID-19. For example, since COVID-19 emerged as a high-profile, and very publishable subject, it is possible that it was easier for men, who are still more likely to be in leadership and well-funded positions in academia to pick up the topic quickly. Also, women were less likely to function as expert reviewers on articles related to COVID-19,[6–8] known to potentially exacerbate a pre-existing gender bias in the peer-review and publishing process.[22–24] These might contribute to a vicious cycle that hindered access to COVID-19 publishing especially for women in the early stages of their career.[11] The exact determination of underlying mechanisms, however, warrants future research that might also benefit from longer time-series data.

Interestingly, we found that decreased publishing activity was specifically significant in fields that had a relatively equal share between women and men as first authors prior to the pandemic. Those were also fields that had a high overall productivity among COVID-19 articles. This is a reason for concern insofar as the current scientific/medical response to one of the most incisive global crises could be overly dominated by men and missing expert voices by women that would usually be a vital part of this research.

Applying longitudinal analyses, we found that the much-increased gender gap in relevant fields was particularly noticeable early on and continuously trended back towards the baseline since then. However, recovery remains slow, with potential reasons being the extent of the impact, the fact that many fields with an above average share of women first authors were affected, but also because some of the most affected fields require in-person work, for example, in wet labs (eg, biochemistry, molecular biology, microbiology).

Analysing the change in the gender gap globally, we similarly found that the shares of women first authors declined across almost all geographic areas early in the course of the pandemic in spring of 2020. Since then, the gender gap slowly began to close again towards the expected baseline and even reached baseline in North America and Oceania. Asia is the only continent where no significant reduction in women's first authorships was noted. In China, there was no change in gender composition among COVID-19 publication when compared with control publications. In Taiwan and South Korea, a small increase in women first author shares was noted. Given that these countries together accounted for 4.2% of COVID-19 publications in our dataset, and that gender designation algorithms tend to offer lower probability gender designations for Asian forenames, we are reluctant to conclusively interpret these findings. Of note, we applied a uniform probability threshold of 90% for designating an author's gender as a conservative measure.

Even though our longitudinal analysis leads us to be cautiously optimistic that the impact of the pandemic on women's COVID-19 research activity might have been temporary, we speculate that the absence of many expert women voices during the initial response to the pandemic impacted not only the individual researchers, but also society as a whole. While our study focused on COVID-19 publications, the dynamics reported here may be amplified in research beyond COVID-19. Effects of the pandemic on early-stage or ongoing projects are likely to show with a time-delay and potentially have long-lasting consequences jeopardising efforts toward equity in academia. For example, women at earlier career stages have not been able to allocate enough time to their research, manuscript and grant writing, were bound to remote working instead of in-person work, were potentially less likely to be allocated to leading roles on projects given the circumstances and so on.

To avoid long-term impacts on the academic advancement and scientific contributions, the disproportionate impact of the pandemic on early-career women investigators needs serious consideration and immediate actions. In a first step we would suggest for an open discourse about how the pandemic has highlighted systemic and structural barriers preventing gender equity in academia.[25] Naturally, pre-existing inequities must be evaluated, and a long-term strategy has to be established to support equity in science.[26] But more acutely, COVID-19-related gender inequities need to be addressed with direct measures, monetary and non-monetary, on both the political (eg, federal funding agencies) and institutional level. For example, financial support for postdoctoral/graduate students could help to facilitate a research setback in a recently established laboratory. Modifications for grant deadlines, timelines, extensions for granted expenses, as well as additional (bridge) funding programmes, are likely warranted for early-career mentored/independent investigators.[27] Extension of tenure evaluation and promotion should be considered on the individual level accounting for constraints posed by COVID-19 for junior faculty (for both men and women). Resources for childcare should be provided for parents, additional funds for expanded childcare arrangements could help to reallocate time to regular professional duties. However, it is as necessary to normalise the increased stress of living through a pandemic that not only affects professional obligations and goals, but also other family members, and no penalty should be awarded for caregivers but measures mentioned above should rather allow for extra quality time.[28]

Our study had several limitations. One limitation is that part of our large-scale study design was based on field association by Clarivate Journal categories, which bears a potential risk of misclassification or inclusion of articles from journals that might not follow the norm of author ordering with regard to contribution that we assume for the life sciences and medicine. For example, one of the most affected fields in our data analysis is public, environmental and occupational health, where such norms might not be generalisable. However, previous research about the topic in public health, for example, also applied said authorship order norms and given the high relevance of the field and results, we decided to present the data.[29] Next, we relied on authors' first names to designate their likely gender, which bears the risk of gender misclassification, particularly across different geographies. We attempted to minimise this risk by applying a 90% probability requirement, however, a certain level of uncertainty remains.[30] Also, by design of the gender designation algorithms grouping into two categories, namely 'men' and 'women', we cannot separate out an effect for scholars who are non-binary, transgender men and women. Along those lines, by virtue of the large-scale nature of our study, we acknowledge that we cannot draw conclusions for researchers on the individual level as well as confirm the assumed career stage. While our study focused on

gender disparities for COVID-19-related research, it is important to note that, beyond gender diversity,[31] ethnic and cultural diversity benefit science on multiple levels.[32] With our analyses, however, we cannot comment on how the pandemic might have affected ethnic and cultural diversity with regard to COVID-19-related research or if populations under-represented in academic life sciences were similarly affected. Lastly, in this observational study, we cannot causally decipher the underlying mechanisms leading to women being under-represented on COVID-19-related research, also precluding a definitive explanation for the trend back to baseline over time and what the potentially successful measures were or could be.

In conclusion, we found that women first authors have been under-represented in COVID-19-related research, particularly at the beginning of the pandemic, despite having nearly equal first authorship shares as men in pertinent fields prior to the pandemic.

**Author affiliations**
[1]Department of Cardiology, Angiology, Pulmonology, University Hospital Heidelberg, Heidelberg, Germany
[2]German Center for Heart and Cardiovascular Research (DZHK), Heidelberg/Mannheim, Germany
[3]Area Management, University of Mannheim, Mannheim, Germany
[4]Department of Health Care Policy, Harvard Medical School, Boston, Massachusetts, USA
[5]Department of Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA
[6]National Bureau of Economic Research, Cambridge, Massachusetts, USA

**Twitter** Carolin Lerchenmüller @CLerchenmueller, Leo Schmallenbach @LSchmallenbach, Anupam B Jena @AnupamBJena and Marc J Lerchenmueller @MLerchenmueller

**Acknowledgements** The authors acknowledge support by the German state of Baden-Württemberg through bwHPC—high performance cluster computing resources. The authors also thank Tobias Engel and Felix Fieberg for supporting data collection and literature review.

**Contributors** LS and ML collected data. LS, ML and CL performed data analysis. CL, ML and AJ wrote the manuscript, LS edited the manuscript. All authors read and approved the final manuscript.

**Funding** Support was provided by the Office of the Director, National Institutes of Health (1DP5OD017897, AJ). Joachim Herz Foundation (LS) and University of Mannheim (FAiR@UMA) (ML).

**Disclaimer** The funding sources had no role in the design and conduct of the study; collection, management, analysis and interpretation of the data; and preparation, review or approval of the manuscript.

**Competing interests** AJ reports receiving 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. Also unrelated to this work, ML reports serving on the board of AaviGen.

**Patient consent for publication** Not required.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** All data and source code generated and analysed during the current study can be accessed athttps://doi.org/10.7910/DVN/OOTXIW (Harvard Dataverse).



**ORCID iDs**
Carolin Lerchenmüller http://orcid.org/0000-0002-5620-0285
Leo Schmallenbach http://orcid.org/0000-0001-5302-2578
Marc J Lerchenmueller http://orcid.org/0000-0001-8875-3602

## REFERENCES

1 World Economic Forum. The global gender gap report, 2017. Available: http://www3weforumorg/docs/WEF_GGGR_2017pdf [Accessed 08 June 2018].
2 O'Brien KR, Scheffer M, van Nes EH, *et al*. How to break the cycle of low workforce diversity: a model for change. *PLoS One* 2015;10:ARTN e0133208.
3 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.
4 Blumenthal DM, Olenski AR, Yeh RW, *et al*. Sex differences in faculty RANK among academic cardiologists in the United States. *Circulation* 2017;135:506–17.
5 Holman L, Stuart-Fox D, Hauser CE. The gender gap in science: how long until women are equally represented? *PLoS Biol* 2018;16:e2004956.
6 Squazzoni F, Bravo G, Grimaldo F, *et al*. Only second-class tickets for women in the COVID-19 race. A study on manuscript submissions and reviews in 2329 Elsevier journals. *SSRN Journal* 2020 (published Online First: October 16, 2020).
7 Krukowski RA, Jagsi R, Cardel MI. Academic productivity differences by gender and child age in science, technology, engineering, mathematics, and medicine faculty during the COVID-19 pandemic. *J Womens Health* 2021;30:341–7.
8 Mogensen MA, Lee CI, Carlos RC. The impact of the COVID-19 pandemic on Journal scholarly activity among female contributors. *J Am Coll Radiol* 2021. doi:10.1016/j.jacr.2021.01.011. [Epub ahead of print: 26 Jan 2021].
9 Andersen JP, Nielsen MW, Simone NL, *et al*. COVID-19 medical papers have fewer women first authors than expected. *Elife* 2020;9. doi:10.7554/eLife.58807. [Epub ahead of print: 15 06 2020].
10 Kibbe MR. Consequences of the COVID-19 pandemic on manuscript submissions by women. *JAMA Surg* 2020;155:803–4.
11 Pinho-Gomes A-C, Peters S, Thompson K, *et al*. Where are the women? gender inequalities in COVID-19 research authorship. *BMJ Glob Health* 2020;5:e002922.
12 Yildirim TM, Eslen-Ziya H. The differential impact of COVID-19 on the work conditions of women and men academics during the Lockdown. *Gend Work Organ* 202010.1111/gwao.12529. [Epub ahead of print: 19 08 2020] (published Online First: 2020/09/10).
13 Staniscuaski F, Reichert F, Werneck FP, *et al*. Impact of COVID-19 on academic mothers. *Science* 2020;368:724.
14 Oleschuk M. Gender equity considerations for Tenure and promotion during COVID-19. *Can Rev Sociol* 2020;57:502–15.
15 Kitchener C. Women academics seem to be submitting fewer papers during coronavirus. 'Never seen anything like it,' says one editor. *The Lily* 2020;24.
16 Myers KR, Tham WY, Yin Y, *et al*. Unequal effects of the COVID-19 pandemic on scientists. *Nat Hum Behav* 2020;4:880–3.
17 Czymara CS, Langenkamp A, Cano T. Cause for concerns: gender inequality in experiencing the COVID-19 lockdown in Germany. *European Societies* 2020:1–14.
18 Antecol H, Bedard K, Stearns J. Equal but inequitable: who benefits from gender-neutral Tenure clock stopping policies? *American Economic Review* 2018;108:2420–41.
19 Jagsi R, Guancial EA, Worobey CC, *et al*. The "gender gap" in authorship of academic medical literature--a 35-year perspective. *N Engl J Med* 2006;355:281–7.
20 Clarivate. Journal citation reports 2018, 2018. Available: https://clarivate.com/webofsciencegroup/wp-content/uploads/sites/2/2019/10/Crv_JCR_Full-Marketing-List_A4_2018_v4.pdf

21 BCG. Easing the covid-19 burden on working parents, 2020. Available: https://www.bcg.com/en-us/publications/2020/helping-working-parents-ease-the-burden-of-covid-19.aspx
22 Clark J, Horton R. What is the Lancet doing about gender and diversity? *Lancet* 2019;393:508–10.
23 Holmes A, Hardy S. *Gender bias in peer review–opening up the black box*. London, UK: London School of Economics, 2019.
24 Pinho-Gomes A-C. The time is ripe for addressing gender inequalities in the authorship of scientific papers. *Am J Public Health* 2021;111:15–16.
25 Coe IR, Wiley R, Bekker L-G. Organisational best practices towards gender equality in science and medicine. *Lancet* 2019;393:587–93.
26 Malisch JL, Harris BN, Sherrer SM, *et al*. Opinion: in the wake of COVID-19, academia needs new solutions to ensure gender equity. *Proc Natl Acad Sci U S A* 2020;117:15378–81.
27 Gibson EM, Bennett FC, Gillespie SM, *et al*. How support of early career researchers can reset science in the Post-COVID19 world. *Cell* 2020;181:1445–9.
28 Spinelli M, Lionetti F, Pastore M, *et al*. Parents' stress and children's psychological problems in families facing the COVID-19 outbreak in Italy. *Front Psychol* 2020;11:1713. doi:10.3389/fpsyg.2020.01713
29 Bittante C, Beatrice G, Carletti L, *et al*. Gender disparity in authorships of manuscripts on the COVID-19 outbreak. *Z Gesundh Wiss* 2020:1–2 (published Online First: 2020/08/25).
30 Lerchenmueller MJ, Sorenson O, Jena AB. Gender differences in how scientists present the importance of their research: observational study. *BMJ* 2019;367:l6573.
31 Nielsen MW, Alegria S, Börjeson L, *et al*. Opinion: gender diversity leads to better science. *Proc Natl Acad Sci U S A* 2017;114:1740–2.
32 Science benefits from diversity. *Nature* 2018;558:5.

**Supplementary materials for**

# Longitudinal Analyses of Gender Differences in First Authorship Publications Related to COVID-19

Carolin Lerchenmüller, MD, Leo Schmallenbach, MSc,
Anupam B. Jena, MD, PhD, Marc J. Lerchenmüller, MPH, PhD

**Table of Contents**

**Additional Information on Data**

We merged several databases to analyze potential gender differences in first authorships of COVID publications relative to a set of control publications in the same journals and within the same time period one year earlier. First, we extracted all articles from the PubMed database for which the term "COVID" appeared in the title or abstract and obtained all available article characteristics including, among others, the names of all authors, country affiliation per author, the journal ISSN (International Standard Serial Number), and time of publication (months and year). The U.S. National Library of Medicine maintains the PubMed XML database and a detailed data inventory can be found online (https://www.nlm.nih.gov/databases/download/pubmed_medline.html). We obtained the journals' major scientific discipline from the Clarivate Journal Citation Report of 2018 via the unique journal ISSNs. We used journal names as a crosswalk to identify publications that appeared a year earlier in the exact same journals as the COVID articles.

An overview of the sample creation is provided in **Figure S1**. In service of estimation accuracy, we included only journals that are listed in Clarivate. By construction that excludes all COVID publications in journals that had no publication on record in PubMed for 2019. These journals likely only came into being in 2020. We restricted our search query to articles published between February 1st of 2020 and January 31st of 2021, since these months were the most productive in terms of COVID publishing and we sought to mitigate seasonal influences, like gender differences in teaching load at certain times of year.

We used the forenames recorded in PubMed to designate the gender of authors (PubMed started to systematically record forenames in 2002). We determined the probable gender of the authors through the Genderize database, an established approach that allows gender assignment for a large number of authors. At the time of initial submission, Genderize included 86,710 distinct forenames drawn from 74 countries and 81 languages. Recent 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 Genderize provided the most accurate estimates of gender (*1*). Our underlying code for calling the Genderize database with a large set of forenames has been posted to Figshare (*2*). Genderize uses a variety of information, such as social media records, to assign a probability that an individual with a particular forename is a man or a woman. For example, Genderize designates the forename "Chris" as male with 93% probability based on 8,631 verified records in the database. We considered gender determined if Genderize assigned a probability of greater than 90%. Applying this threshold, we designated the gender for more than 72% of the authors in our dataset. However, there is variation across author origins (**Figure S2**). For example, we designated the gender for 84% of authors with an affiliation from North America and for 52% of authors with an affiliation from Asia. The lower accuracy for authors from Asia is a common challenge in name-based gender designation and a limitation to our analysis of authors from these countries. Yet, there is no difference in the accuracy of gender designation across men and women authors (**Figure S3**) or COVID and non-COVID

articles (**Figure S4**). Hence, there is no reason to be concerned that the gender designation would systematically bias our results. Additionally, our main findings do not change when setting different gender designation thresholds.

Next, we compared the distribution of disciplines producing COVID research relative to the articles in the control sample (**Table S1**). Ranking the disciplines in terms of publication output, and testing a Spearman Rank correlation, we obtain a coefficient of greater 0.80. While this correlation would generally be considered strong (*3*) lending credence to our basic design, it does not consider the possibility that men and women may sort differently into these fields. However, our **Figure 2** in the main text documents that it is primarily fields where women tend to be well represented that produce COVID research.

To execute country-level analyses, we use regular expressions to extract the full country name or country codes from affiliation data for the first author. We also ranked countries by productivity for COVID-articles and control articles, obtaining a Spearman rank correlation of 0.94, again supporting our approach of using non-COVID articles in the prior year as a control group (**Table S2**). This also mitigates concerns that countries with larger gender gaps in general produce more COVID research.

**Additional information on methods**

Measurement

To assess the effect of the COVID pandemic on the gender gap in publishing, we reported unadjusted differences in the percent of women first authorships versus male first authorships for COVID and non-COVID publications. This straightforward metric provides a direct and easy to understand measure of how the COVID pandemic impacts women's versus men's publication productivity.

$$\begin{aligned}\Delta GenderGap \; = & \{FirstAuthor_{Female} - \; FirstAuthor_{Male} \mid COVID\} \\ & - \; \{FirstAuthor_{Female} - \; FirstAuthor_{Male} \mid Non - COVID\}\end{aligned}$$

To conduct subgroup analysis for discipline and country, we calculated the change in the gender gap based on the percent of first authorships by men and women for the specific discipline and country.

Estimation

In addition to the unadjusted differences, we also provided adjusted differences in first authorships from women and men obtained from linear probability models (**Table S3**), adjusting for the number of authors on a publication, the month of publication, the field of research and country. We run the same analysis for last authorships from women and

men (**Table S4**). Both regression analyses support the descriptive evidence presented in Figure 1 of the main text. Logistic regression as an alternative estimation model has two disadvantages in our analysis. First, the large number of fixed effects when including countries and discipline dummies, for example, raises the possibility of incidental parameters bias and could prevent the convergence of some of our models. Second, logistic regressions can overestimate effect sizes as a result of the high leverage of marginal cases (i.e., identifying larger gender differences than reported in the main text), whereas linear probability models average across observations and produce more conservative results (see also **Table S5**).

We provided adjusted estimates in the supplement as one might be concerned, for example, that men are more numerous in fields that produce COVID research. This would also lower women's observed COVID productivity but not due to pandemic related constraints as hypothesized, but rather due to underlying structural differences in subspecialties. Of note, the descriptive data paint a different picture, such that women tend to be at least equal if not overrepresented in the most productive COVID disciplines.

We conducted four robustness checks to establish the reliability of our findings (**Table S6**). In the first two robustness checks, we vary the threshold applied to the accuracy of the gender designation. In Model 1, we consider all authors, for which gender was assigned with a probability higher than chance (>50%). In Model 2, we only consider authors, for which the gender designation accuracy was reported with 100%. Both models show very similar estimates for the decrease in women authorship on COVID publications (8.2%-points vs. 9.0%-points). Next, we excluded articles from the analysis, for which collective authorship was indicated in PubMed. This concerns roughly 8% of articles but excluding them does not alter the effect estimate. Last, we reran the analysis on the full sample, that is including COVID articles published in journals, which are not listed in Clarivate's journal citation report and for which the first author's gender could be designated. As we do not know the disciplines these journals fall into, we include journal instead of discipline fixed effects in this last model specification. Again, the results are consistent with our previous analysis. Accordingly, a descriptive comparison of the articles in- and excluded from the analysis shows that they are near identical with respect to the representation of women first and last authors (**Table S7**).

## References

1. C. N. G. detection, (http://codingnews.info/post/genderdetection (accessed 11/15/17) (2015)).
2. Lerchenmueller. Marc, *Genderize_unlimited_API_request*. (2017).
3. M. G. Pagano, Kimberlee, *Principle of Biostatistics*. (Brooks/Cole, ed. Secon Edition, 2000).

**Figure S1:** Sample construction for COVID articles and non-COVID (control) articles

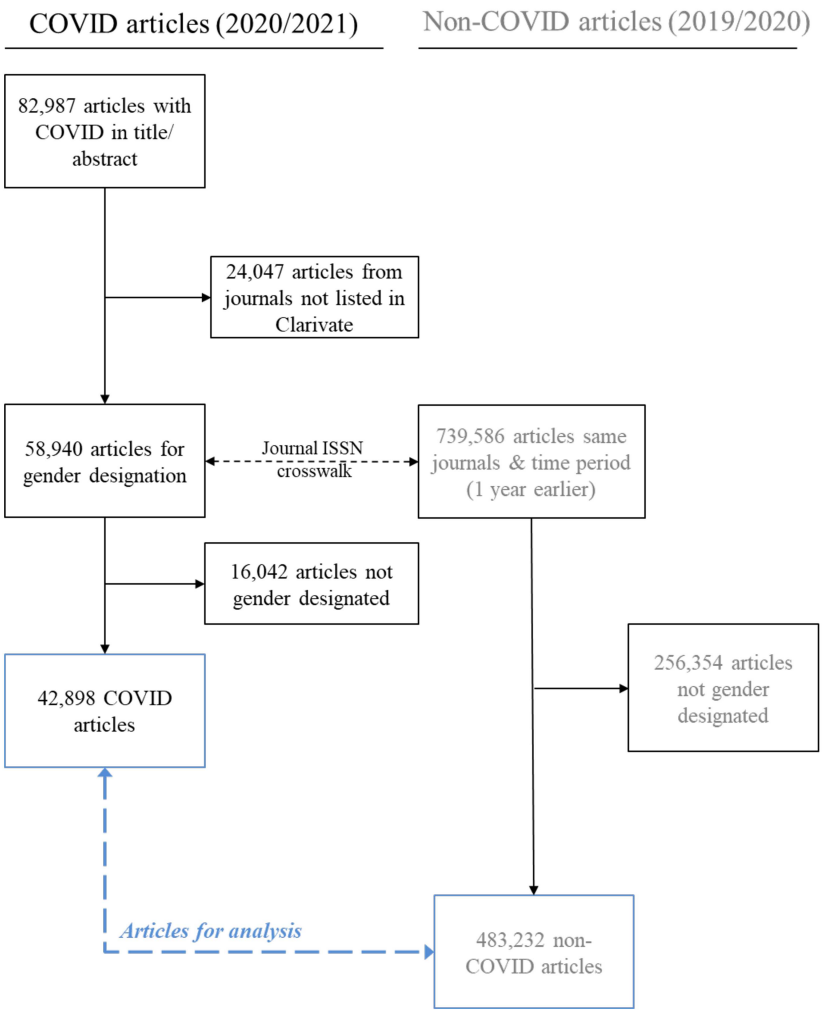
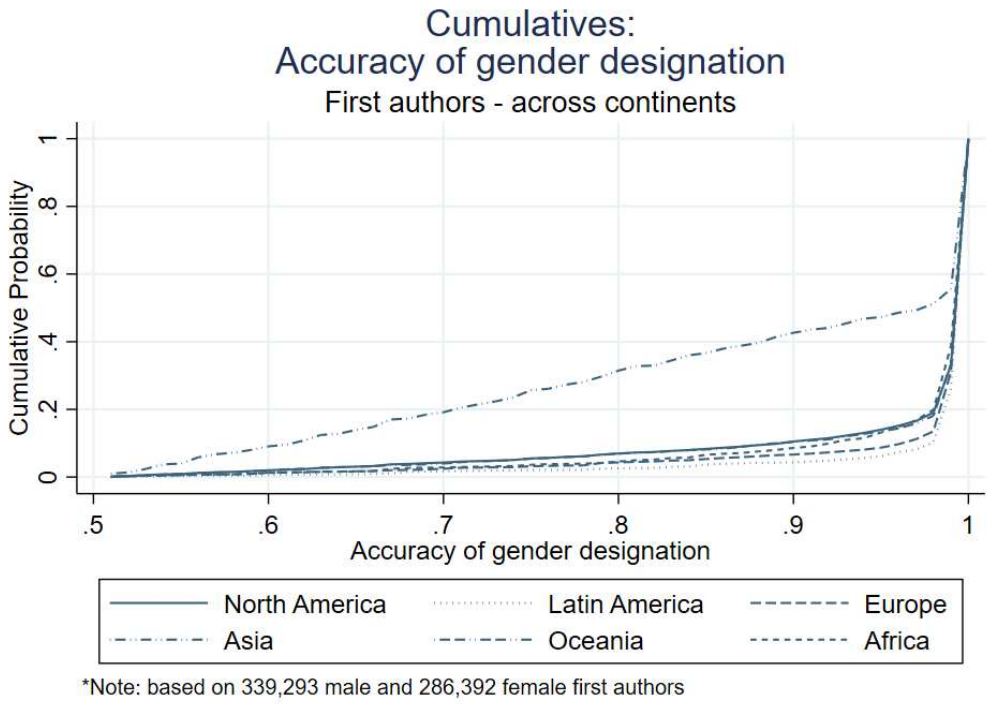
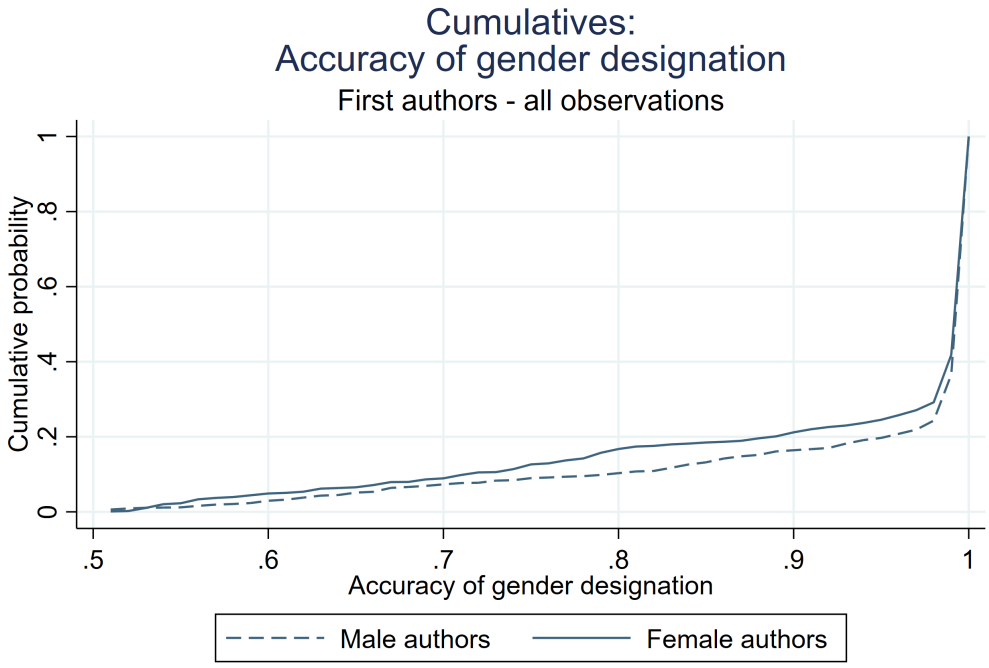


**Figure S2:** Gender designation accuracy for first authors from North America, Latin America, Europe, Asia, Oceania, and Africa separately




*Note: based on 339,293 male and 286,392 female first authors

**Figure S3:** Gender designation accuracy for all articles in the sample (COVID articles and non-COVID articles) for women and men first authors




*Note: based on 349,489 male 296,937 female first authors

**Figure S4:** Gender designation accuracy for women first authors for COVID articles and non-COVID (control) articles separately

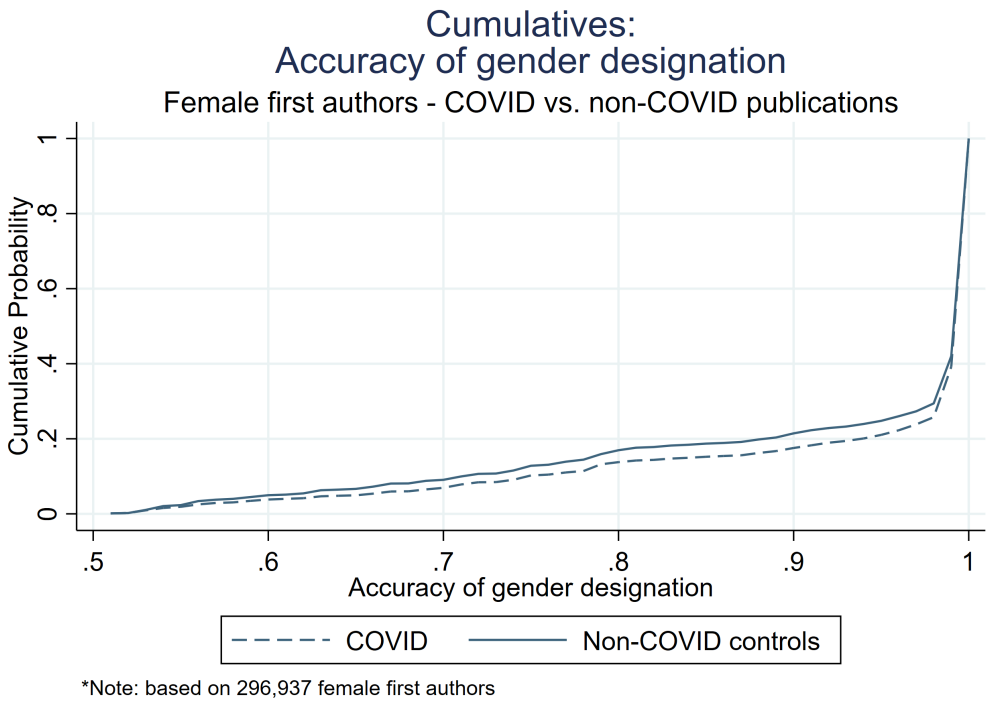


**Table S1:** Major disciplines for COVID articles and non-COVID (control) articles and concordance statistic

| Discpline | Non-COVID | | | COVID | | | Total | | |
|---|---|---|---|---|---|---|---|---|---|
| | Total | % | Rank | Total | % | Rank | Total | % | Rank |
| MEDICINE GENERAL & INTERNAL | 23,163 | 4.79% | 5 | 4,229 | 9.88% | 1 | 27,392 | 5.21% | 4 |
| PUBLIC ENVIRONMENTAL & OCCUPATIONAL HEALTH | 17,860 | 3.70% | 8 | 3,530 | 8.24% | 2 | 21,390 | 4.07% | 6 |
| SURGERY | 26,457 | 5.48% | 3 | 2,254 | 5.26% | 3 | 28,711 | 5.46% | 3 |
| IMMUNOLOGY | 11,151 | 2.31% | 12 | 1,925 | 4.50% | 4 | 13,076 | 2.49% | 11 |
| CARDIAC & CARDIOVASCULAR SYSTEMS | 16,196 | 3.35% | 9 | 1,479 | 3.45% | 5 | 17,675 | 3.36% | 9 |
| PHARMACOLOGY & PHARMACY | 18,408 | 3.81% | 7 | 1,439 | 3.36% | 6 | 19,847 | 3.77% | 8 |
| MULTIDISCIPLINARY SCIENCES | 38,524 | 7.97% | 1 | 1,435 | 3.35% | 7 | 39,959 | 7.60% | 1 |
| PSYCHIATRY | 9,890 | 2.05% | 14 | 1,333 | 3.11% | 8 | 11,223 | 2.13% | 14 |
| ENVIRONMENTAL SCIENCES | 19,607 | 4.06% | 6 | 1,278 | 2.98% | 9 | 20,885 | 3.97% | 7 |
| ONCOLOGY | 23,207 | 4.80% | 4 | 1,159 | 2.71% | 10 | 24,366 | 4.63% | 5 |
| BIOCHEMISTRY & MOLECULAR BIOLOGY | 30,224 | 6.25% | 2 | 1,118 | 2.61% | 11 | 31,342 | 5.96% | 2 |
| PEDIATRICS | 10,280 | 2.13% | 13 | 1,105 | 2.58% | 12 | 11,385 | 2.16% | 13 |
| INFECTIOUS DISEASES | 3,143 | 0.65% | 41 | 1,001 | 2.34% | 13 | 4,144 | 0.79% | 37 |
| CLINICAL NEUROLOGY | 9,614 | 1.99% | 15 | 922 | 2.15% | 14 | 10,536 | 2.00% | 15 |
| HEALTH CARE SCIENCES & SERVICES | 5,209 | 1.08% | 29 | 921 | 2.15% | 15 | 6,130 | 1.17% | 27 |
| MEDICINE RESEARCH & EXPERIMENTAL | 5,645 | 1.17% | 25 | 887 | 2.07% | 16 | 6,532 | 1.24% | 25 |
| NEUROSCIENCES | 14,611 | 3.02% | 10 | 835 | 1.95% | 17 | 15,446 | 2.94% | 10 |
| DERMATOLOGY | 5,452 | 1.13% | 26 | 777 | 1.81% | 18 | 6,229 | 1.18% | 26 |
| VIROLOGY | 2,198 | 0.45% | 47 | 771 | 1.80% | 19 | 2,969 | 0.56% | 43 |
| RESPIRATORY SYSTEM | 4,074 | 0.84% | 35 | 732 | 1.71% | 20 | 4,806 | 0.91% | 32 |
| GASTROENTEROLOGY & HEPATOLOGY | 6,367 | 1.32% | 23 | 719 | 1.68% | 21 | 7,086 | 1.35% | 22 |
| RADIOLOGY NUCLEAR MEDICINE & MEDICAL IMAGING | 7,092 | 1.47% | 20 | 715 | 1.67% | 22 | 7,807 | 1.48% | 19 |
| ENDOCRINOLOGY & METABOLISM | 6,468 | 1.34% | 22 | 628 | 1.47% | 23 | 7,096 | 1.35% | 21 |
| HEMATOLOGY | 4,449 | 0.92% | 31 | 618 | 1.44% | 24 | 5,067 | 0.96% | 31 |
| UROLOGY & NEPHROLOGY | 7,369 | 1.52% | 18 | 581 | 1.36% | 25 | 7,950 | 1.51% | 18 |
| ANESTHESIOLOGY | 2,100 | 0.43% | 48 | 519 | 1.21% | 26 | 2,619 | 0.50% | 44 |
| MICROBIOLOGY | 6,086 | 1.26% | 24 | 499 | 1.17% | 27 | 6,585 | 1.25% | 24 |
| NURSING | 4,602 | 0.95% | 30 | 484 | 1.13% | 28 | 5,086 | 0.97% | 30 |
| EMERGENCY MEDICINE | 1,958 | 0.41% | 52 | 459 | 1.07% | 29 | 2,417 | 0.46% | 48 |
| RHEUMATOLOGY | 3,146 | 0.65% | 40 | 455 | 1.06% | 30 | 3,601 | 0.68% | 40 |
| PSYCHOLOGY MULTIDISCIPLINARY | 3,271 | 0.68% | 38 | 413 | 0.96% | 31 | 3,684 | 0.70% | 39 |
| OPHTHALMOLOGY | 5,303 | 1.10% | 27 | 412 | 0.96% | 32 | 5,715 | 1.09% | 28 |
| GERIATRICS & GERONTOLOGY | 1,818 | 0.38% | 55 | 392 | 0.92% | 33 | 2,210 | 0.42% | 52 |
| OBSTETRICS & GYNECOLOGY | 4,388 | 0.91% | 32 | 353 | 0.82% | 34 | 4,741 | 0.90% | 33 |
| CELL BIOLOGY | 8,490 | 1.76% | 16 | 352 | 0.82% | 35 | 8,842 | 1.68% | 16 |
| OTORHINOLARYNGOLOGY | 2,098 | 0.43% | 49 | 303 | 0.71% | 36 | 2,401 | 0.46% | 49 |
| CRITICAL CARE MEDICINE | 2,018 | 0.42% | 50 | 282 | 0.66% | 37 | 2,300 | 0.44% | 50 |
| DENTISTRY ORAL SURGERY & MEDICINE | 3,212 | 0.66% | 39 | 275 | 0.64% | 38 | 3,487 | 0.66% | 41 |
| ECONOMICS | 495 | 0.10% | 81 | 274 | 0.64% | 39 | 769 | 0.15% | 74 |
| BIOTECHNOLOGY & APPLIED MICROBIOLOGY | 4,021 | 0.83% | 36 | 246 | 0.57% | 40 | 4,267 | 0.81% | 34 |
| NUTRITION & DIETETICS | 5,241 | 1.08% | 28 | 229 | 0.53% | 41 | 5,470 | 1.04% | 29 |
| PSYCHOLOGY CLINICAL | 2,392 | 0.50% | 43 | 210 | 0.49% | 42 | 2,602 | 0.49% | 45 |
| ETHICS | 674 | 0.14% | 75 | 195 | 0.46% | 43 | 869 | 0.17% | 71 |
| MEDICAL LABORATORY TECHNOLOGY | 1,217 | 0.25% | 63 | 185 | 0.43% | 44 | 1,402 | 0.27% | 61 |
| SPORT SCIENCES | 3,065 | 0.63% | 42 | 182 | 0.43% | 45 | 3,247 | 0.62% | 42 |
| PATHOLOGY | 2,011 | 0.42% | 51 | 181 | 0.42% | 46 | 2,192 | 0.42% | 53 |
| GENETICS & HEREDITY | 7,184 | 1.49% | 19 | 178 | 0.42% | 47 | 7,362 | 1.40% | 20 |
| ORTHOPEDICS | 2,345 | 0.49% | 44 | 168 | 0.39% | 48 | 2,513 | 0.48% | 46 |
| PERIPHERAL VASCULAR DISEASE | 1,700 | 0.35% | 57 | 133 | 0.31% | 49 | 1,833 | 0.35% | 56 |
| MATHEMATICS INTERDISCIPLINARY APPLICATIONS | 70 | 0.01% | 120 | 126 | 0.29% | 50 | 196 | 0.04% | 104 |

| Spearman Rank Correlation - all disciplines | |
|---|---|
| coefficient (rs) | 0.807 |
| N | 148 |
| T statistic | 16.537 |
| DF | 146 |
| p-value | 0.000 |

| Spearman Rank Correlation - top 50 disciplines | |
|---|---|
| coefficient (rs) | 0.738 |
| N | 50 |
| T statistic | 7.588 |
| DF | 48 |
| p-value | 0.000 |

## **Table S2:** Major countries for COVID articles and non-COVID (control) articles and concordance statistics

| Country (First Author) | Non-COVID | | | COVID | | | Total | | |
|---|---|---|---|---|---|---|---|---|---|
| | Total | % | Rank | Total | % | Rank | Total | % | Rank |
| United States | 120,478 | 26.65% | 1 | 11,066 | 26.12% | 1 | 131,544 | 26.75% | 1 |
| Italy | 22,670 | 5.01% | 5 | 4,309 | 10.15% | 2 | 26,979 | 5.49% | 4 |
| United Kingdom | 27,994 | 6.19% | 3 | 3,157 | 8.13% | 3 | 31,151 | 6.33% | 2 |
| India | 11,388 | 2.52% | 12 | 1,778 | 7.43% | 4 | 13,166 | 2.68% | 12 |
| Spain | 15,273 | 3.38% | 9 | 1,672 | 4.19% | 5 | 16,945 | 3.45% | 9 |
| China | 29,078 | 6.43% | 2 | 1,377 | 3.77% | 6 | 30,455 | 6.19% | 3 |
| Canada | 16,654 | 3.68% | 7 | 1,324 | 3.11% | 7 | 17,978 | 3.66% | 7 |
| France | 13,837 | 3.06% | 10 | 1,313 | 3.02% | 8 | 15,150 | 3.08% | 10 |
| Germany | 24,413 | 5.40% | 4 | 1,161 | 2.64% | 9 | 25,574 | 5.20% | 5 |
| Brasil | 13,219 | 2.92% | 11 | 1,160 | 2.64% | 10 | 14,379 | 2.92% | 11 |
| Australia | 16,132 | 3.57% | 8 | 1,052 | 2.50% | 11 | 17,184 | 3.49% | 8 |
| Iran | 6,757 | 1.49% | 15 | 878 | 2.07% | 12 | 7,635 | 1.55% | 14 |
| Turkey | 6,134 | 1.36% | 18 | 851 | 1.97% | 13 | 6,985 | 1.42% | 16 |
| Japan | 18,952 | 4.19% | 6 | 499 | 1.30% | 14 | 19,451 | 3.96% | 6 |
| Netherlands | 9,702 | 2.15% | 13 | 425 | 1.24% | 15 | 10,127 | 2.06% | 13 |
| Switzerland | 6,455 | 1.43% | 16 | 422 | 1.00% | 16 | 6,877 | 1.40% | 17 |
| Singapore | 1,823 | 0.40% | 36 | 405 | 0.99% | 17 | 2,228 | 0.45% | 32 |
| Israel | 3,813 | 0.84% | 22 | 378 | 0.96% | 18 | 4,191 | 0.85% | 22 |
| Saudi Arabia | 2,020 | 0.45% | 33 | 341 | 0.96% | 19 | 2,361 | 0.48% | 30 |
| Greece | 2,605 | 0.58% | 27 | 323 | 0.86% | 20 | 2,928 | 0.60% | 26 |
| Belgium | 4,132 | 0.91% | 21 | 314 | 0.82% | 21 | 4,446 | 0.90% | 21 |
| Pakistan | 1,769 | 0.39% | 38 | 286 | 0.80% | 22 | 2,055 | 0.42% | 36 |
| Mexico | 3,186 | 0.70% | 24 | 281 | 0.74% | 23 | 3,467 | 0.71% | 24 |
| Egypt | 2,541 | 0.56% | 28 | 254 | 0.71% | 24 | 2,795 | 0.57% | 28 |
| Poland | 6,318 | 1.40% | 17 | 253 | 0.66% | 25 | 6,571 | 1.34% | 18 |
| Hong Kong | 1,285 | 0.28% | 41 | 248 | 0.63% | 26 | 1,533 | 0.31% | 41 |
| Ireland | 2,100 | 0.46% | 31 | 230 | 0.59% | 27 | 2,330 | 0.47% | 31 |
| Austria | 3,313 | 0.73% | 23 | 203 | 0.58% | 28 | 3,516 | 0.71% | 23 |
| South Korea | 7,400 | 1.64% | 14 | 199 | 0.52% | 29 | 7,599 | 1.55% | 15 |
| Sweden | 5,841 | 1.29% | 19 | 192 | 0.45% | 30 | 6,033 | 1.23% | 19 |
| Bangladesh | 318 | 0.07% | 60 | 154 | 0.45% | 31 | 472 | 0.10% | 55 |
| Portugal | 3,007 | 0.67% | 25 | 151 | 0.36% | 32 | 3,158 | 0.64% | 25 |
| Denmark | 4,599 | 1.02% | 20 | 143 | 0.35% | 33 | 4,742 | 0.96% | 20 |
| South Africa | 1,771 | 0.39% | 37 | 142 | 0.35% | 34 | 1,913 | 0.39% | 38 |
| United Arab Emirates | 532 | 0.12% | 51 | 129 | 0.34% | 35 | 661 | 0.13% | 50 |
| Colombia | 869 | 0.19% | 45 | 120 | 0.33% | 36 | 989 | 0.20% | 45 |
| Chile | 1,520 | 0.34% | 40 | 114 | 0.29% | 37 | 1,634 | 0.33% | 40 |
| Taiwan | 2,101 | 0.46% | 30 | 112 | 0.27% | 38 | 2,213 | 0.45% | 33 |
| Norway | 2,812 | 0.62% | 26 | 106 | 0.26% | 39 | 2,918 | 0.59% | 27 |
| Malaysia | 1,224 | 0.27% | 42 | 94 | 0.25% | 40 | 1,318 | 0.27% | 42 |
| Peru | 346 | 0.08% | 58 | 93 | 0.22% | 41 | 439 | 0.09% | 56 |
| Argentina | 1,739 | 0.38% | 39 | 89 | 0.20% | 42 | 1,828 | 0.37% | 39 |
| Romania | 1,142 | 0.25% | 43 | 84 | 0.20% | 43 | 1,226 | 0.25% | 43 |
| Russia | 2,083 | 0.46% | 32 | 83 | 0.20% | 44 | 2,166 | 0.44% | 34 |
| Lebanon | 679 | 0.15% | 49 | 82 | 0.20% | 45 | 761 | 0.15% | 49 |
| New Zealand | 2,011 | 0.44% | 34 | 80 | 0.19% | 46 | 2,091 | 0.43% | 35 |
| Nigeria | 501 | 0.11% | 52 | 75 | 0.19% | 47 | 576 | 0.12% | 52 |
| Indonesia | 289 | 0.06% | 64 | 73 | 0.17% | 48 | 362 | 0.07% | 60 |
| Jordan | 345 | 0.08% | 59 | 68 | 0.17% | 49 | 413 | 0.08% | 58 |
| Morocco | 304 | 0.07% | 63 | 66 | 0.17% | 50 | 370 | 0.08% | 59 |

| Spearman Rank Correlation - all countries | |
|---|---|
| coefficient (rs) | 0.93 |
| N | 167 |
| T statistic | 32.32 |
| DF | 165 |
| p-value | 0.000 |

| Spearman Rank Correlation - top 50 countries | |
|---|---|
| coefficient (rs) | 0.85 |
| N | 50 |
| T statistic | 10.97 |
| DF | 48 |
| p-value | 0.000 |

**Table S3:** Hierarchical linear probability model for the likelihood of women first authorship for COVID articles versus non-COVID (control) articles

| *Dependent variable: First Author Female* | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| **COVID** | **-0.074*** ** | **-0.074*** ** | **-0.075*** ** | **-0.086*** ** | **-0.089*** ** |
| | **(0.00)** | **(0.00)** | **(0.00)** | **(0.00)** | **(0.00)** |
| number of authors | | 0.002*** | 0.002*** | 0.002*** | 0.002*** |
| | | (0.00) | (0.00) | (0.00) | (0.00) |
| publication month fixed effects (12) | | | Included | Included | Included |
| discipline fixed effects (148) | | | | Included | Included |
| country fixed effects (167) | | | | | Included |
| constant | 0.451*** | 0.442*** | 0.454*** | 0.269*** | 0.036 |
| | (0.00) | (0.00) | (0.00) | (0.02) | (0.10) |
| R-squared | 0.002 | 0.002 | 0.002 | 0.043 | 0.060 |
| Adjusted R-squared | 0.002 | 0.002 | 0.002 | 0.043 | 0.059 |
| Observations | 526,130 | 526,130 | 526,130 | 526,130 | 491,912 |

Note: standard errors in brackets, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

**Table S4:** Hierarchical linear probability model for the likelihood of women last authorship for COVID articles versus non-COVID (control) articles

| *Dependent variable: Last Author Female* | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| **COVID** | **-0.014*** ** | **-0.014*** ** | **-0.015*** ** | **-0.033*** ** | **-0.037*** ** |
| | **(0.00)** | **(0.00)** | **(0.00)** | **(0.00)** | **(0.00)** |
| number of authors | | -0.002*** | -0.002*** | -0.001*** | -0.001*** |
| | | (0.00) | (0.00) | (0.00) | (0.00) |
| publication month fixed effects (12) | | | Included | Included | Included |
| discipline fixed effects (148) | | | | Included | Included |
| country fixed effects (167) | | | | | Included |
| constant | 0.319*** | 0.332*** | 0.342*** | 0.232*** | 0.081 |
| | (0.00) | (0.00) | (0.00) | (0.02) | (0.09) |
| R-squared | 0.000 | 0.001 | 0.001 | 0.043 | 0.060 |
| Adjusted R-squared | 0.000 | 0.001 | 0.001 | 0.043 | 0.059 |
| Observations | 539,103 | 539,103 | 539,103 | 539,103 | 504,148 |

Note: standard errors in brackets, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

**Table S5:** Hierarchical logit regression for the likelihood of women first authorship for COVID articles versus non-COVID (control) articles

| *Dependent variable: First Author Female* | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| **COVID** | **0.738*** ** | **0.737*** ** | **0.733*** ** | **0.690*** ** | **0.678*** ** |
| | **(0.01)** | **(0.01)** | **(0.01)** | **(0.01)** | **(0.01)** |
| number of authors | | 1.006*** | 1.006*** | 1.009*** | 1.009*** |
| | | (0.00) | (0.00) | (0.00) | (0.00) |
| publication month fixed effects (12) | | | Included | Included | Included |
| discipline fixed effects (148) | | | | Included | Included |
| country fixed effects (167) | | | | | Included |
| constant | 0.821*** | 0.790*** | 0.829*** | 0.365*** | 0.117*** |
| | (0.00) | (0.00) | (0.01) | (0.03) | (0.06) |
| observations | 526,130 | 526,130 | 526,130 | 526,112 | 491,837 |

Note: Coefficients reported as odds ratios, standard errors in brackets, * p < 0.05, ** p < 0.01, *** p < 0.001

**Table S6:** Robustness checks

| *Dependent variable: First Author Female* | Accuracy of gender designation > 50% (1) | Accuracy of gender designation = 100% (2) | Exluding collective authorships (3) | Full sample (4) |
|---|---|---|---|---|
| **COVID** | **-0.082*** ** | **-0.090*** ** | **-0.089*** ** | **-0.079*** ** |
| | **(0.00)** | **(0.00)** | **(0.00)** | **(0.00)** |
| number of authors | 0.002*** | 0.002*** | 0.003*** | 0.002*** |
| | (0.00) | (0.00) | (0.00) | (0.00) |
| publication month fixed effects (12) | Included | Included | Included | Included |
| discipline fixed effects (148) | Included | Included | Included | Included |
| country fixed effects (167) | Included | Included | Included | |
| journal fixed effects (5,101) | | | | Included |
| constant | 0.095 | 0.040 | 0.031 | 0.215* |
| | (0.10) | (0.10) | (0.10) | (0.10) |
| R-squared | 0.049 | 0.067 | 0.060 | 0.091 |
| Adjusted R-squared | 0.048 | 0.066 | 0.059 | 0.082 |
| Observations | 607,598 | 443,711 | 483,308 | 507,653 |

Note: standard errors in brackets, * $p < 0.05$, ** $p< 0.01$, *** $p < 0.001$

**Table S7:** Descriptive statistics of articles included in the analysis versus articles not included in the analysis

| | Included in analysis | | Excluded from analysis | | t-test | |
|---|---|---|---|---|---|---|
| **Variable** | **Mean** | **Std. Dev.** | **Mean** | **Std. Dev.** | **Difference** | **t-statistic** |
| First Author Female | 0.38 | 0.48 | 0.38 | 0.49 | 0.00 | 0.45 |
| Last Author Female | 0.32 | 0.47 | 0.33 | 0.47 | 0.01 | 2.27 |
| Publication Month | 7.35 | 3.15 | 7.51 | 3.25 | 0.15 | 5.35 |
| Number of Authors | 6.42 | 8.56 | 5.75 | 6.25 | -0.68 | -10.76 |
| North America | 0.29 | 0.45 | 0.24 | 0.43 | -0.05 | -12.70 |
| Europe | 0.35 | 0.48 | 0.28 | 0.45 | 0.07 | -16.15 |
| Asia | 0.19 | 0.39 | 0.23 | 0.42 | 0.04 | 9.69 |
| Latin America | 0.05 | 0.21 | 0.04 | 0.20 | 0.00 | -1.44 |
| Oceania | 0.03 | 0.16 | 0.02 | 0.14 | -0.01 | -4.22 |
| Africa | 0.02 | 0.14 | 0.03 | 0.17 | 0.01 | 6.20 |
| **Observations** | **42,898** | | **17,445** | | **60,343** | |