J Data Sci. Info. Citation Studies., 2025; 4(1):103-106.
https://www.jcitation.org

Short Communication

Female Representation in Retracted Scientific Publications: A Study of Authorship and Collaboration

Kiran Sharma[1,2,*], Satyam Mukherjee[3], Harsha V. Garine[4]

[1]Schools of Engineering and Technology, BML Munjal University, Gurugram, Haryana, INDIA.
[2]Center for Advanced Data and Computational Science, BML Munjal University, Gurugram, Haryana, INDIA.
[3]School of Management and Entrepreneurship, Shiv Nadar University, Greater Noida, Uttar Pradesh, INDIA.
[4]Charles W. Davidson College of Engineering, San Jose State University, California, UNITED STATES.

ABSTRACT

Numerous studies in the literature have emphasized the underrepresentation of women in the scientific field, leading to their underrepresentation in prestigious publications, authorship roles and collaborations. At the same time, the involvement of women in scientific misconduct is also increasingly attracting the attention of the scientific community. This motivated the present study to explore female collaboration and authorship positions in retracted publications. To conduct the analysis, 3750 retracted scientific papers were extracted from the Web of Science and the respective gender of each author was identified. The evaluation included the year-wise representation of female authors, females at various authorship positions, collaboration and female-to-male odds ratio. In all, 26.43% of authorship is held by women and the share of male-female collaborative retracted publications is 55.11%. To conclude, women are less likely to hold the last authorship and more likely to hold the middle authorship position in retracted publications.

Keywords: Female Authorship, Gender Disparity, Male-Female Collaboration, Retraction.

Correspondence:
Dr. Kiran Sharma[1,2]
[1]Schools of Engineering and Technology, BML Munjal University, Gurugram-122413, Haryana, INDIA.
[2]Center for Advanced Data and Computational Science, BML Munjal University, Gurugram-122413, Haryana, INDIA.
Email: kiran.sharma@bmu.edu.in

Received: 12-11-2024;
Revised: 02-01-2025;
Accepted: 31-03-2025.

INTRODUCTION

Extant literature in STEM suggests gender disparity acts as a deterrent for women in the scientific workforce (Vásárhelyi *et al.*, 2021; Fang *et al.*, 2013; Aiston and Jung, 2015). The existing "glass ceiling" prohibits women from receiving scientific recognition, delays hiring and promotion and receives fewer citations compared to their male counterparts. On the other hand, with the rise of retracted publications, the representation of females in such retracted publications also caught the attention of the scientific community (Van Noorden, 2011; Pinho-Gomes, 2023).

Recent studies have highlighted a growing trend in scientific misconduct (Steen, 2011; Stehen, 2013; Fang *et al.*, 2012; Campos-Varela, 2019). Publications may be retracted for various reasons, ranging from unintentional editorial errors to intentional fraud, such as data or image manipulation, fabricated conclusions, or duplicate publications (Sharma, 2024; Khurana *et al.*, 2024). Regardless of the underlying reasons, retracted publications can lead to a loss of reputation in the scientific community (Sharma and Mukherjee, 2024; Azoulay *et al.*, 2017).

With the rise of collaboration in scientific research, the involvement of women scholars has increased in recent years and consequently, women are also represented in retracted papers (Sharma, 2021; Peng *et al.*, 2022). However, in collaborative work, an author's contribution is often measured by their authorship position, with first and last or corresponding authors typically bearing the primary responsibility for research integrity. The contributions of authors in other positions are generally considered to be less significant in comparison. Therefore, the present study aims to investigate the position of female authors in collaborative retracted papers (Zheng, 2023). This raises the question: Are women as likely as men to be associated with erroneous publications and what roles do they play in such retracted works? (Agogino, 2007).

Our research has two primary objectives. First, we examine the involvement of women in retracted publications. Second, we assess the proportion of female authorship positions and their collaboration in these retracted works. The structure of the study is as follows: Section 2 discusses data collection and filtering methods. The results are detailed in Section 3 and the conclusion is provided in Section 4.

OPEN ACCESS

Manuscript TECHNOMEDIA

DOI: 10.5530/jcitation.20250143

Publishing Partner : Manuscript Technomedia. [www.mstechnomedia.com]

METHODOLOGY

Data acquisition and filtration

We performed a Web of Science (WoS) search for the period from 1989 to 2021 to retrieve all articles listed as "Retracted". A total of 5845 research papers were retracted from the portal. These papers were published in more than 1897 journals. The data has all the necessary meta-data for the analysis, which includes author information, publisher details, year of publication, citation count and discipline of research. The papers retracted before 2006 did not provide the complete author's details in the database. Hence, we removed those papers from our study and left with 3750 papers. The authors' names and respective positions from 3750 papers were extracted to get the information of 30412 authors. Further, the authors with initials in their names were removed and 19164 authors were filtered. For each entry, the data set was prepared on the following variables: manuscript title, year of retraction, first, middle and last names of the authors, journal impact factor, country of provenance of all authors and discipline of publication. Figure 1 provides a schematic representation of the methodology of data extraction, filtration and gender.

Gender estimation

An online database named "Gender API" was used to estimate the author's gender (*https://gender-api.com/).* This third-party Gender API estimates gender by first name and country. This API has been used by many researchers prior in their work to examine the gender disparity in the authorship of academic articles. The gender estimation was performed based on the author's first name and the country for the accuracy of search results. The gender API provides an accuracy score for the prediction of gender across each name from its sampled data. The accuracy score varies from 0 (not confirmed) to 100 (highly confirmed). The threshold for author selection was chosen with accuracy scores above 60. Hence, 17634 authors' genders with a given accuracy were filtered. Further, to cross-verify the data, a random sample of 10% of the data was selected and a manual recheck was performed on the authors' gender from their scholar or institute profiles.

RESULTS

First, we provide a visual summary of women's representation in retracted publications. Figure 2 provides descriptive information on the women's representation year-wise retracted articles from 2006 to 2020 (See Figure 2a) and at different authorship positions: first, last, middle and total (See Figure 2b). We observe that between 2006 and 2020, the percentage of women in retracted publications increased slightly from 21% to 29%, while a slight drop was observed in the percentage of male authors. Further, women comprise 25.54% of first authors, 28.71% of middle authors and 19.19% of last authors (See Figure 2b). In the full author's sample, on average, 26.43% of women appeared in retracted publications.

In sciences, the first authorship position is usually considered as a junior author who executed the research, while the last authorship position points towards the author who leads and funds the research- the senior author. The male-female collaboration at different levels also points toward the active involvement of women in the research. To analyse gender-specific collaboration in retracted publications, papers were categorised as M-M (Male-Male), F-F (Female-Female) and M-F (Male-Female). 41.23% of retractions were reported from M-M collaboration, 3.66% from F-F and 55.11% from M-F collaboration (Figure

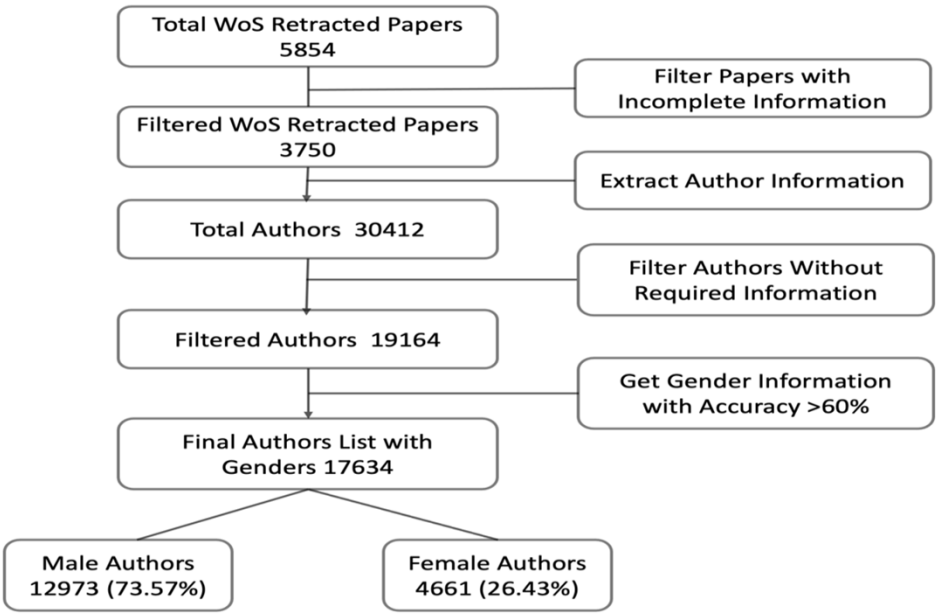

Figure 1: The Flowchart Represents the Data Selection, Filtration and Gender Estimation Process.

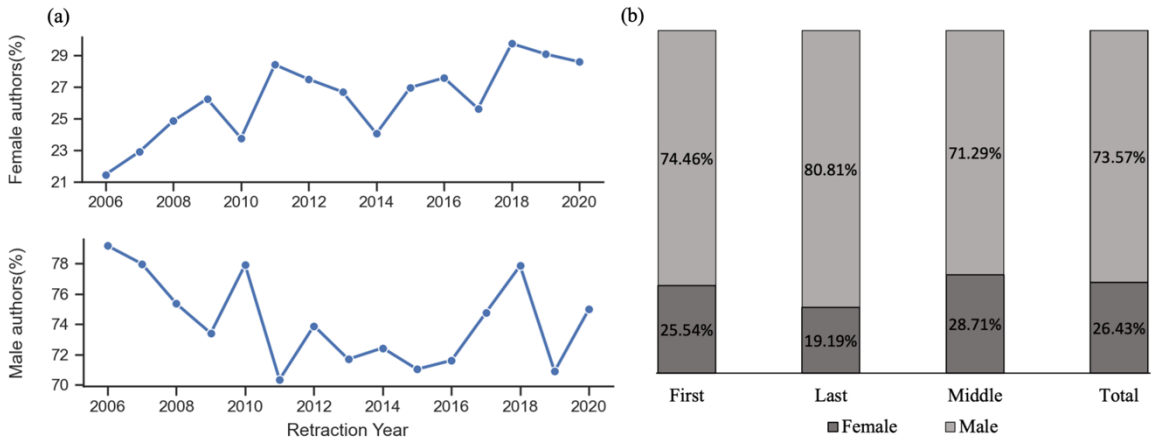

Figure 2: (a) Year-wise male and female contribution trend in retracted publications. (b) The global share of females as first, last and middle authorship positions. Females comprise 26.43% of total authorship than males (73.57%).

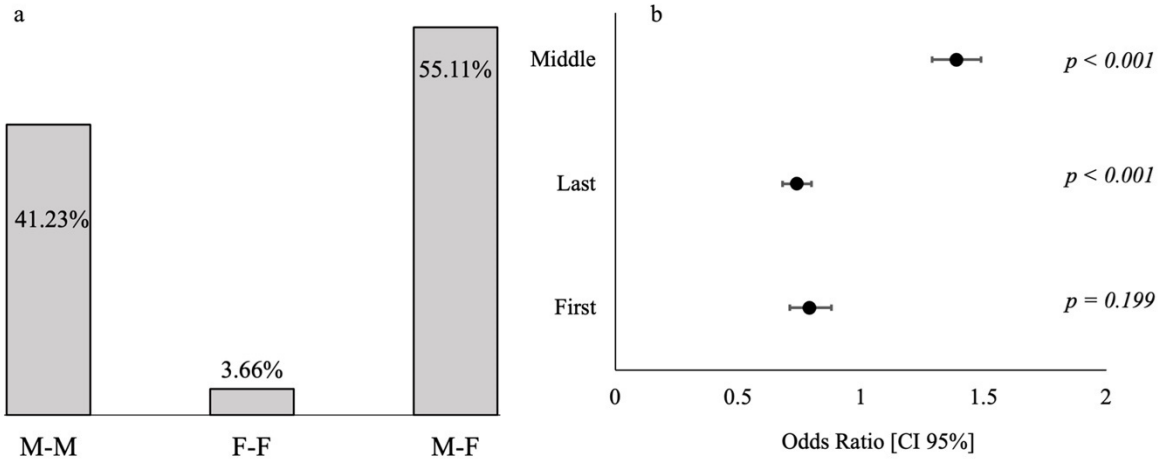

Figure 3: (a) Gender-specific collaboration in retracted publications where M-M represents all male authors' collaboration, F-F represents all female authors' collaboration and M-F represents male-female collaboration. (b) Female to Male odds ratio. Females are less likely to be last and more likely to be the middle author.

3(a)). Although the retraction reported from the male-female collaboration is higher, women are less likely to hold the last authorship position (OR=0.62, p<0.001) and more likely to hold the middle authorship position (OR=1.38, p<0.001) (Figure 3(b)).

Finally, we implement a multinomial logit model to estimate the strength and significance of the relationship between the author's gender and authorship position. The dependent variable is a binary variable, which takes the value of 1 if the author's gender is female and 0 otherwise. The independent variable is a categorical variable: First author, middle author and last author. Additionally, we account for the unobserved heterogeneity in the data by including fixed effects of publication years, scientific fields and countries. Table 1 presents the multinomial logit estimate comparing the number of female authors to male authors for the last author and middle author positions relative to the first author position. The multinomial logit for female authors relative to male authors is 0.138 times higher for being in the middle position relative to the first position (p<0.01). Similarly, the number of female authors is 0.385 units lower than the number of male authors for being in the last position relative to the first position.

DISCUSSION AND CONCLUSION

Gender disparity in science is widely studied by researchers worldwide. This study investigated women participation in various authorship positions and their likelihood of being in that position. Also, the proportion of male-female collaboration in such a retracted paper is being analyzed. Out of 3750 retracted papers, the gender of 19164 authors was identified. Overall, female representation is less in retracted papers and the share of male-female collaborative retracted publications is more than half of the total share. In order to investigate the female contribution at various authorship positions: first, last and middle author, the odds ratio was calculated. It was observed that women are less likely to hold the last authorship and more likely to hold the middle authorship position.

Table 1: Estimates from the multinomial regressions ** <0.01; * <0.001.**

Independent variable (Female dummy)	Base category (First author female)
Last author	-0.385***
	(0.061)
Middle author	0.138**
	(0.046)
Fixed effects	
Publication Year	Y
Fields	Y
Country	Y
#Observations	17634
Log-likelihood	-15567.73
Prob>Chi^2	<0.001

Although we studied the largest set of retracted papers to date, our work is not without limitations. First, for retracted articles published before 2006, data extraction is not feasible since the author names were predominantly abbreviated with initial letters, making first-name-based gender determination impracticable. Our findings are solely based on the subset of the whole retracted database. Second, we did not compare the study with the controlled group; however, we showed the impact of authorship position with respect to the controlled variable, like the publication year, filed and country in Table 1. Third, we assume that the first author is the main author and the last is the lead author, which may not always hold true. Fourth, the use of Gender API is to identify gender and selection of cutoffs. A higher threshold will reduce the count of the number of authors.

Our results highlight women's participation and roles in scientific misconduct. First, it shows the important timeline of female involvement in the retracted publication, which shows a positive trend. Second, more than 50% of papers retracted were written in collaboration among males and females. Third, females are overall more likely to participate as middle authors than the first and last. This shows that women participate less in scientific misconduct.

ACKNOWLEDGEMENT

We acknowledge BML Munjal University, India, for financial support to buy a subscription to Gender-API and Jawaharlal Nehru University, India, for data access from the Web of Science.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

REFERENCES

Agogino, A. (2007). Beyond bias and barriers: Fulfilling the potential of women in academic science and engineering. National Academies Press.

Aiston, S. J., & Jung, J. (2015). Women academics and research productivity: An international comparison. Gender and Education, 27(3), 205-220. https://doi.org/10.1080/09540253.2015.1024617

Azoulay, P., Bonatti, A., & Krieger, J. L. (2017). The career effects of scandal: Evidence from scientific retractions. Research Policy, 46(9), 1552-1569. https://doi.org/10.1016/j.respol.2017.07.003

Campos-Varela, I., & Ruano-Raviña, A. (2019). Misconduct as the main cause for retraction. A descriptive study of retracted publications and their authors. Gaceta Sanitaria, 33(4), 356-360. https://doi.org/10.1016/j.gaceta.2018.01.009

Fang, F. C., Bennett, J. W., & Casadevall, A. (2013). Males are overrepresented among life science researchers committing scientific misconduct. mBio, 4(1), e00640-e00612. https://doi.org/10.1128/mBio.00640-12

Fang, F. C., Steen, R. G., & Casadevall, A. (2012). Misconduct accounts for the majority of retracted scientific publications. Proceedings of the National Academy of Sciences of the United States of America, 109(42), 17028-17033. https://doi.org/10.1073/pnas.1212247109

Khurana, P., Sharma, K., & Uddin, Z. (2024). Unraveling retraction dynamics in COVID-19 research: Patterns, reasons and implications. Accountability in Research, 1-24. https://doi.org/10.1080/08989621.2024.2379906

Peng, H., Romero, D. M., & Horvát, E.-Á. (2022). Dynamics of cross-platform attention to retracted papers. Proceedings of the National Academy of Sciences of the United States of America, 119(25), e2119086119. https://doi.org/10.1073/pnas.2119086119

Pinho-Gomes, A.-C., Hockham, C., & Woodward, M. (2023). Women's representation as authors of retracted papers in the biomedical sciences. PLOS ONE, 18(5), e0284403. https://doi.org/10.1371/journal.pone.0284403

Sharma, K. (2021). Team size and retracted citations reveal the patterns of retractions from 1981 to 2020. Scientometrics, 126(10), 8363-8374. https://doi.org/10.1007/s11192-021-04125-4

Sharma, K. (2024). Over two decades of scientific misconduct in India: Retraction reasons and journal quality among inter-country and intra-country institutional collaboration. Scientometrics, 1–23. https://doi.org/10.1007/s11192-024-05192-z

Sharma, K., & Mukherjee, S. (2024). The ripple effect of retraction on an author's collaboration network. Journal of Computational Social Science, 7(2), 1519-1531. https://doi.org/10.1007/s42001-024-00278-3

Steen, R. G. (2011). Retractions in the scientific literature: Is the incidence of research fraud increasing? Journal of Medical Ethics, 37(4), 249–253. https://doi.org/10.1136/jme.2010.040923

Steen, R. G., Casadevall, A., & Fang, F. C. (2013). Why has the number of scientific retractions increased? PLOS ONE, 8(7), e68397. https://doi.org/10.1371/journal.pone.0068397

Van Noorden, R. (2011). Science publishing: The trouble with retractions. Nature, 478(7367), 26-28. https://doi.org/10.1038/478026a

Vásárhelyi, O., Zakhlebin, I., Milojević, S., & Horvát, E.-Á. (2021). Gender inequities in the online dissemination of scholars' work. Proceedings of the National Academy of Sciences of the United States of America, 118(39), e2102945118. https://doi.org/10.1073/pnas.2102945118

Zheng, E.-T., Fu, H.-Z., & Fang, Z. (2023). Do men commit more scientific misconduct than women? Evidence from retracted articles. In 27th International Conference on Science, Technology and Innovation Indicators (STI 2023). https://doi.org/10.55835/643feae1d7c4812e2f078f27

Cite this article: Sharma K, Mukherjee S, Garine HV. Female Representation in Retracted Scientific Publications: A Study of Authorship and Collaboration. Journal of Data Science, Informetrics, and Citation Studies. 2025;4(1):103-6.