Supplementary material

Gender Equity of Authorship in Pulmonary Medicine Over the Past Decade

|  |  |
| --- | --- |
| 1. **Methods** |  |
| 1.1 Python program principle | p 2 |
| 1.2 Search strategy | p 3 |
| 1. **Results** |  |
| 2.1 eFigure 1. Overall distribution of the percentage of female authors in terms of regional differences | p 4 |
| 2.2 eFigure 2. Line graphs of temporal trends and linear regression equations of female first and last authors over time in different regions | p 5 |
| 2.3 eFigure 3. Overall distribution of the percentage of female authors in terms of country differences | p 6 |
| 2.4 eFigure 4. Line graphs of temporal trends and linear regression equations of female first and last authors over time in different countries | p 7 |
| 2.5 eFigure 5. Percentage of female first and last authors in different journals | p 8 |
| 2.6 eFigure 6. Line graphs of temporal trends and linear regression equations of female first and last authors over time in different journals | p 9 |
| 2.7 eFigure 7. Line graphs of temporal trends and linear regression equations of different gender pairs over time | p 10 |
| 2.8 eFigure 8. Top 100 most prolific authors of the respiratory system between 2012 and 2021 | p 11 |
| 2.9 eTable 1. Gender Disparity in Citations | p 12 |
| 2.10 eTable 2. Total number of citations by first and last authorship gender pairs | p 13 |
| 2.11 eFigure 9. Evaluation on the proportions of female first authors over time in different disciplines | p 14 |
| 2.12 eFigure 10. Evaluation on the proportions of female last authors over time in different disciplines | p 15 |

**1. Methods**

**1.1** **Python program principle**

1. The first author and last author extraction: according to the character split of the author list information, we got a new author list and took the first item for the first author and the last item for the last author; 2. The authors’ address information extraction: according to the character split of the authors’ address information, from the split address information list to take out the address information corresponding to the first author and last authors’ name. When the same author has more than one address, the first address would be extracted by default; 3. The nationality information of author extraction: the extracted authors’ address information (the address information obtained in item 2) would be split by characters again, and the last item of the address field obtained after the split would be the nationality information.

**1.2 Evaluation on Gender Diversity of Authorship in Clinical Medicine**

The literature search strategy was ((((((gender[Title/Abstract]) OR ("Gender Identity"[Mesh])) OR (sex[Title/Abstract])) OR ((women[Title/Abstract]) OR (woman[Title/Abstract]))) OR (female[Title/Abstract])) AND ((author[Title/Abstract]) OR ((authorship\*[Title/Abstract]) OR ("Authorship"[Mesh])))) AND ((journal[Title/Abstract]) OR ((publication\*[Title/Abstract]) OR ("Publications"[Mesh]))).

We excluded studies that 1) did not report the percentage of female authorship; 2) only investigated specific populations, such as faculty in one university, members of one medical society, or medical graduates; 3) only investigated a specific disease or a type of surgery; 4) presented results only in a figure, and did not provide numerical data.

Two investigators independently screened study titles and abstracts, and reviewed full texts to select studies in the systematic review portion of this study. Any disagreement regarding study selection, data extraction, or quality assessments was resolved by discussion with a third investigator.

**2. Results**

**2.1 eFigure 1.**

****

**eFigure 1.** Overall distribution of the percentage of female authors in terms of regional differences. **(A)** Percentage of first and last authors from different regions. **(B)** Percentage of female first and last authors in different regions.

**2.2 eFigure 2.**

****

**eFigure 2.** Line graphs of temporal trends and linear regression equations of female first and last authors over time in different regions.

**2.3 eFigure 3.**

****

**eFigure 3.** Overall distribution of the percentage of female authors in terms of country differences. **(A)** Percentage of first and last authors from different countries. **(B)** Percentage of female first and last authors in different countries.

**2.4 eFigure 4.**

****

**eFigure 4.** Line graphs of temporal trends and linear regression equations of female first and last authors over time in different countries.

**2.5 eFigure 5.**

****

**eFigure 5.** Percentage of female first and last authors in different journals.

**2.6 eFigure 6.**

****

**eFigure 6.** Line graphs of temporal trends and linear regression equations of female first and last authors over time in different journals.

**2.7 eFigure 7.**

****

**eFigure 7.** Line graphs of temporal trends and linear regression equations of different gender pairs over time.

**2.8 eFigure 8.**

****

**eFigure 8.** Top 100 most prolific authors of the respiratory system between 2012 and 2021.

**2.9** **eTable 1. Gender Disparity in Citations.**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Citations by first author gender, median [IQR], No.** | | | | **Citations by last author gender, median [IQR], No.** | | | |
| **Year** | **Female** | | **Male** | ***P*** | **Female** | | **Male** | ***P*** |
| Overall | | 25[9-54] | 29[11-64] | *<0.001* | | 24[9-56] | 28[10-61] | *<0.001* |
| 2012 | | 41[22-78] | 42[21-87] | *0.455* | | 41[18-82] | 42[22-85] | *0.546* |
| 2013 | | 41[21-73] | 43[22-83] | *0.292* | | 43[21-76] | 42[22-79] | *0.910* |
| 2014 | | 40[19-76] | 41[21-77] | *0.577* | | 37[20-68] | 42[20-80] | *0.124* |
| 2015 | | 36[18-72] | 37[19-80] | *0.457* | | 36[18-70] | 37[19-79] | *0.521* |
| 2016 | | 32[17-59] | 37[17-72] | *0.068* | | 36[16-68] | 35[17-68] | *0.683* |
| 2017 | | 32[17-52] | 34[16-71] | *0.074* | | 32[17-52] | 33[16-66] | *0.250* |
| 2018 | | 22[12-44] | 25[13-49] | *0.047* | | 22[12-48] | 25[13-48] | *0.248* |
| 2019 | | 17[8-33] | 20[9-41] | *0.010* | | 16[8-34] | 19[9-38] | *0.105* |
| 2020 | | 9[4-24] | 12[5-40] | *<0.001* | | 9[4-27] | 11[4-35] | *0.009* |
| 2021 | | 2[0-9] | 3[1-18] | *<0.001* | | 3[1-15] | 3[1-13] | *0.562* |

Abbreviations: IQR, interquartile range.

**2.10** **eTable 2. Total number of citations by first and last authorship gender pairs.**

|  |  |  |
| --- | --- | --- |
| **Author pair** | **Citations, median [IQR], No.** | ***P*** |
| Female-female | 22 [7-49] | *<0.001* |
| Female-male | 27 [10-57] |
| Male-female | 27 [10-63] |
| Male-male | 29 [11-65] |

Abbreviations: IQR, interquartile range.

**2.11 eFigure 9.**

****

**eFigure 9.** Evaluation on the proportions of female first authors over time in different disciplines.

**2.12 eFigure 10.**

****

**eFigure 10.** Evaluation on the proportions of female last authors over time in different disciplines.