

The Lancet. Volume 391, No. 10140, p2601, 30 June 2018.

DOI: [https://doi.org/10.1016/S0140-6736\(18\)31139-5](https://doi.org/10.1016/S0140-6736(18)31139-5)

Author Gender in The Lancet journals

Julio González-Alvarez¹

¹ Department of Basic, Clinical, and Biological Psychology.
University Jaume I. Castellón (Spain).

short title: Author Gender in The Lancet journals (2014-17).

Correspondence should be addressed at:

Julio González-Alvarez Ph D.
Department of Basic, Clinical, and Biological Psychology.
University Jaume I.
F. Health Sciences.
12071-Castellón de la Plana, Spain..

Email: gonzalez@psb.uji.es

Fax: +34 964 729267 Phone: +34 964 729670

orcid.org/0000-0002-0389-5263.

Note: The author will be away from the office from 2-11 (March).

Abstract

Despite important advances in recent decades, gender inequality persists in science. In this Comment, the current gender composition of the authors published in *The Lancet* journals is analyzed briefly. In general terms, women represent about one-third of article authorships, with the noteworthy exception of *The Lancet Psychiatry* (45.2%). Female representation among first authors is 51.1% in *The Lancet Psychiatry* and 42.9% in *The Lancet Global Health*, higher than the overall percentages.

A common feature (except for *The Lancet HIV* and, to a lesser extent, *The Lancet Global Health*) is a more pronounced gender gap in the last (senior) position, which indicates that age might be a factor (although not the only one) modulating gender asymmetry in *The Lancet* journals.

Despite significant advances made in recent decades, gender inequality persists in many scientific fields¹, including medicine^{2,3} and global health areas. In this special issue dedicated to *Women in science, medicine and global health*, *The Lancet* focuses on looking forward to help to understand and remove women's existing disadvantage in science. Nevertheless, it is worthwhile to dedicate a few words about the current starting point in terms of women's representation in *The Lancet* journals.

Here, I present a snapshot of the gender of authors who publish in *The Lancet* journals. The study was based on Thomson Reuters' WoS database. All the articles published in the past four years (2014-17) in the ten *The-Lancet* journals included in the database were extracted. A total of 3,321 articles, signed by 68,846 authorships (making an average of about twenty authors per paper), was obtained. Gender could be identified in 60,642 authorships (88.1% of the total; see methodological details in the Supplementary Material): 40,464 (66.7%) corresponding to men, and 20,178 (33.3%) corresponding to women (the gender percentages always refer to the known-gender total).

The author data were further segregated according to the journals and the order in the paper by-lines. Figure 1 shows the overall percentage of women authors in each journal, and the corresponding female percentages as first or last author (numerical values in the Supplementary Material). Overall, women represent about one third of the authors in *The Lancet* (31.8%) and six other journals. *The Lancet Psychiatry* (45.2%), *The Lancet Global Health* (39.8%), and *The Lancet HIV* (38.8%) stand out because they present a lower gender imbalance (particularly the former).

Female representation as first author is 51.1% in *The Lancet Psychiatry* and 42.9% in *The Lancet Global Health*, in both cases higher than the overall percentage. The first author is often the researcher who has made the most significant contribution to the work, especially in terms of performance and time dedicated. These figures could indicate a future trend towards a

greater incorporation of new female researchers in these fields. A common feature (except for *The Lancet HIV* and, to a lesser extent, *The Lancet Global Health*), also found in other studies^{1,4}, is that women appear selectively underrepresented in the last author position. In many fields, including health and behavioral sciences, the last position is usually reserved for the senior or leading member of a research group⁵, generally a scientist with a consolidated (and presumably long) career. A more pronounced gender gap in the senior position could indicate that age might be an important factor (although probably not the only one) modulating gender asymmetry in *The Lancet* journals.

References

¹ Larivière V, Ni C, Gingras Y, Cronin B, Sugimoto CR. Bibliometrics: global gender disparities in science. *Nature*. 2013;504(7479):211-213.

² Filardo G, da Graca B, Sass DM, Pollock BD, Smith EB, Martinez MA. Trends and comparison of female first authorship in high impact medical journals: observational study (1994-2014). *BMJ*. 2016;352:i847. doi: 10.1136/bmj.i847.

³ 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;355:281-287.

⁴ González-Alvarez J, Cervera-Crespo T. Research production in high-impact journals of contemporary neuroscience: a gender analysis. *J Informetrics*. 2017;11(1):232-243.

⁵ Waltman L. An empirical analysis of the use of alphabetical authorship in scientific publishing. *J Informetrics*. 2012;6(4):700-711.

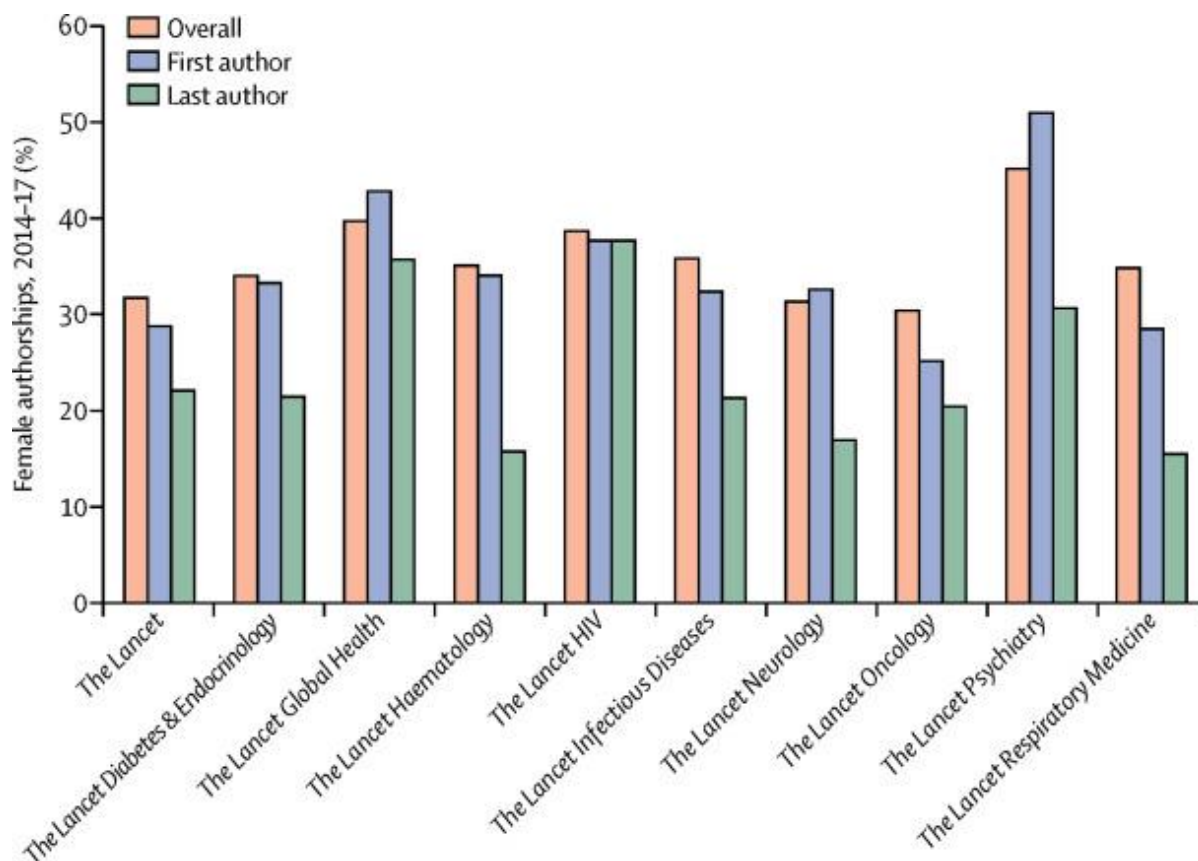


Figure 1. Representation of women as authors of articles published in *The Lancet* journals during the years 2014-17. The figure displays overall percentages and percentages of women as first or last author. Numerical values are available in the Supplementary Material.

Julio González-Alvarez Ph.D.

The author states no conflict of interest

Author Gender in The Lancet journals

Appendix: Supplementary Material

Method

Database.

This study was based on Thomson Reuters' WoS (Web of Science) database, specifically the Science Citation Index Expanded (SCI-EXPANDED). All the articles (Document Type= Article) published during the years 2014-17 in the ten The-Lancet journals included in SCI-EXPANDED were extracted. The journals were *The Lancet*, *The Lancet Diabetes & Endocrinology*, *The Lancet Global Health*, *The Lancet Haematology*, *The Lancet HIV*, *The Lancet Infectious Diseases*, *The Lancet Neurology*, *The Lancet Oncology*, *The Lancet Psychiatry*, and *The Lancet Respiratory Medicine*.

Gender identification of Authorships.

It was based on the first names of the authorships. The Web of Science (WoS) database, like most scientific databases, does not provide the authors' gender. However, in 2008 the WoS began to include the authors' full names (field tag AF: Author Full Name), although a small proportion of records still display only the authors' initials. After a preprocess of normalization that eliminated initials accompanying given names and replaced hyphens with spaces, all the

authors' first names were matched through GenderChecker, a database that includes 102,142 worldwide names, classified as male, female, or unisex (acquired from <http://genderchecker.com/>). This database is being used in research¹⁻⁴ and, according to the website, by the UN Refugee Agency.

In order to increase the number of observations, we followed Larivière et al.'s procedure⁵ (see also⁶⁻⁷) and the names classified as unisex by GenderChecker were subsequently matched to the 1990 US Census. This census presents lists of first names and their frequencies associated with males and females from the US population (no other subsequent census has offered frequencies associated with gender). When a name classified as unisex by GenderChecker presented a rate above 90% (vs. 10%) associated with a specific gender in the US Census, the name was finally classified as belonging to that gender. For example, 'Aaron' appeared in the US Census 7,209 times linked to men and 64 times linked to women (99.1% vs. 0.9%); therefore, it was considered a male name. Conversely, the name 'Carmen' accounted for 6,210 women and 330 men (95% vs. 5%); consequently, it was classified as a female name. Moreover, to maximize the number of observations, we manually identified the gender of a large quantity of authors by locating biographical information or a photo on the Internet.

Procedure

Each variable of interest (author full name, title of the article, year and journal of publication, etc.) was extracted using the BibExcel program⁸. This software is a toolbox for bibliometricians that creates a file in which the values of an extracted variable are associated with each individual article (identified with a number). Finally, the values of all the variables studied were merged in a master Excel database to perform the bibliometric and statistical analyses.

Validation Study

To test the accuracy of our data, we followed a similar procedure to Larivière et al.⁵ These authors selected a random sample of authorships to manually check their gender identity. We randomly selected 383 authorships from the total of 68,846 authorships and manually identified the gender of each author by locating some biographical information or a photo on the Internet (see also⁶⁻⁷). According to the formula:

$$\text{Sample size} = \frac{\frac{z^2 \times p(1-p)}{e^2}}{1 + \left(\frac{z^2 \times p(1-p)}{e^2 N}\right)}$$

where N = population size, e = margin of error, and z = z-score, 383 is the sample size corresponding to N= 68,846, e = 0.05, and a 95% confidence interval (online calculator at <https://www.surveymonkey.com/mp/sample-size-calculator/>).

After several attempts, the gender of 15 authorships could not be identified and they were replaced by other 15 items randomly selected. The 383 authorships corresponded to 250 male (65.7%) and 132 female (34.6%) authorships, yielding a proportion close to the overall rate of 66.7% males and 33.3% females. The relative difference between the observed frequencies in the validation study (250 males, 132 females) and the expected frequencies according to the overall proportion (255 males, 127 females) was not significant, $X^2(1) = 0.29$; $p = 0.587$.

Data

A total of 3,321 articles published in The Lancet journals (2014-17) were extracted from the SCI-EXPANDED database. Table S1 shows the distribution of the number of articles published by each journal.

The articles were signed by 68,846 authorships, making an average of 20.7 authorships per paper. After excluding the authorships with only initials, unisex names, or given names that did not match the GenderChecker database, we obtained 60,642 (88.1%) items with known gender (male, female) (henceforth, the percentages of female or male authorships will always refer to the known-gender total). Importantly, the total number of authorships with known gender included 40,464 (66.7%) authorships corresponding to men and 20,178 (33.3%) corresponding to women. Taking all the data together, it is obvious that the gender inequality is considerable. Applying the Chi square test, the difference between male and female authorships was statistically significant, $X^2(df=1) = 3490.70$, $p < 0.0001$, Cramer's $V = .170$ (effect size).

Data were further segregated according to the journals and the author order in the paper by-lines Table S1 present the gender distribution of the authorships (overall, as first author, and as last author).

Table S1. Gender distribution of Authorships of articles published in *The Lancet* Journals during the years 2014-17. Gender percentages refer to gender-identified values. M: Males; F: Females.

	Articles	Author ships	Gender Identified	Gender of the Authorships											
				Overall				First Author				Last Author			
				M	%	F	%	M	%	F	%	M	%	F	%
THE LANCET	1047	31822	28481	19417	68.2	9064	31.8	682	71.2	276	28.8	751	77.8	214	22.2
THE LANCET DIABETES & ENDOCR	215	3794	3329	2195	65.9	1134	34.1	126	66.7	63	33.3	153	78.5	42	21.5
THE LANCET GLOBAL HEALTH	234	3203	2680	1614	60.2	1066	39.8	113	57.1	85	42.9	142	64.3	79	35.7
THE LANCET HAEMATOLOGY	147	2352	2103	1364	64.9	739	35.1	83	65.9	43	34.1	111	84.1	21	15.9
THE LANCET HIV	144	2341	2116	1296	61.2	820	38.8	81	62.3	49	37.7	84	62.2	51	37.8
THE LANCET INFECTIOUS DISEASES	368	5401	4601	2947	64.1	1654	35.9	221	67.6	106	32.4	263	78.5	72	21.5
THE LANCET NEUROLOGY	217	4029	3610	2474	68.5	1136	31.5	132	67.3	64	32.7	166	83.0	34	17.0
THE LANCET ONCOLOGY	544	10546	8875	6167	69.5	2708	30.5	371	74.6	126	25.4	407	79.5	105	20.5
THE LANCET PSYCHIATRY	202	1779	1607	880	54.8	727	45.2	90	48.9	94	51.1	133	69.3	59	30.7
THE LANCET RESPIRATORY-MED.	203	3579	3240	2110	65.1	1130	34.9	135	71.4	54	28.6	162	84.4	30	15.6
TOTAL:	3321	68846	60642	40464	66.7	20178	33.3	2034	67.9	960	32.1	2372	77.0	707	23.0

References

- ¹ Mansour AM, Shields CL, Maalouf FC, et al. (2012). Five-decade profile of women in leadership positions at ophthalmic publications. *Arch Ophthalmol*. 2012;130(11):1441-1446.
- ² Yun EJ, Yoon DY, Kim B, et al. Closing the Gender Gap: Increased Female Authorship in AJR and Radiology. *AJR Am J Roentgenol*. 2015;205(2):237-241.
- ³ Carnahan S, Kryscynski D, Olson D. When does corporate social responsibility reduce employee turnover? Evidence from attorneys before and after 9/11. *Acad Management J*. 2017;60(5):1932-1962.
- ⁴ González-Alvarez J, Palomar-García M. Twenty-two years of psychological science in *Psychological Science*. *Int J Psych*. 2014;49(1):56-60.
- ⁵ Larivière V, Ni C, Gingras Y, Cronin B, Sugimoto CR. Bibliometrics: global gender disparities in science. *Nature*. 2013;504(7479):211-213.
- ⁶ González-Alvarez J, Cervera-Crespo T. Research production in high-impact journals of contemporary neuroscience: a gender analysis. *J Informetrics*. 2017;11(1):232-243.
- ⁷ González-Alvarez J, Cervera-Crespo T. Contemporary psychology and women: A gender analysis of the scientific production. *Int J Psych*. 2017. doi:10.1002/ijop.12433 (First online version).

⁸ Persson OD, Danell R, Wiborg-Schneider J. How to use Bibexcel for various types of bibliometric analysis. In Astrom F, Danell R, Larsen B, Schneider J. Celebrating scholarly communication studies: A Festschrift for Olle Persson at his 60th Birthday (pp. 9–24). 2009. Leuven, Belgium: International Society for Scientometrics and Informetrics. Retrieved from <http://www8.umu.se/inforsk/Bibexcel/>.