Diversity in Cochrane Reviews' authorship: Protocol for a meta-research study

Ahmad Sofi-Mahmudi

Seqiz Health Network, Kurdistan University of Medical Sciences, Seqiz, Iran.

Cochrane Iran Associate Centre, National Institute for Medical Research Development (NIMAD), Tehran, Iran.

Corresponding author: Ahmad Sofi-Mahmudi; **Address**: 41, Shanaz St., Seqiz, Kurdistan, Iran; **Telephone**: +98 87 362 34834; **Email**: a.sofimahmudi@gmail.com.

Abstract

Objective and aim: To examine the level of country, region, language, and gender diversity in Cochrane Reviews' authorship.

Methods: I will search for all Cochrane Reviews available through Cochrane Library website (cochranelibrary.com). After extracting the list of reviews, I will extract the affiliations for all authors. Next, I will extract affiliations for first authors, corresponding authors, and last authors. Then, I will assess the proportion of authors from all countries and World Bank income regions as well as developed English-speaking countries. Furthermore, I will assess gender diversity by calculating male-to-female ratio. I will provide yearly trend for each of the abovementioned proportions. I will use R for searches, data handling, analysis and reporting.

Keywords: Meta-research; Cochrane; Cochrane Reviews; Systematic Review; Gender Equity; Authorship.

Aims

The aim of this meta-research study is to examine the level of country, region, language, and gender diversity in Cochrane Reviews' authorship via a fully-automated reproducible approach.

Methods

Data sources and study selection

I will search for all Cochrane Reviews available through Cochrane Library website (cochranelibrary.com) in .csv format.

Data extraction and synthesis

After extracting the list of reviews, I will extract the affiliations for all authors from the Cochrane Library website using web crawling technique. The web address I will use follows this pattern:

"https://www.cochranelibrary.com/cdsr/doi/"

+

DOI

+

"/information"

I will search for "authors" html node and then extract each item and put it in a list. Then, I will merge this list with .csv file that contains information regarding each Review.

Next, I will extract affiliations for first authors, corresponding authors, and last authors. Then, I will assess the proportion of authors from all countries and World Bank income regions (databank.worldbank.org) as well as developed English-speaking countries. Developed English-speaking countries are:

- United States;
- United Kingdom;
- Canada;
- Australia;
- Ireland;
- New Zealand.

Furthermore, I will assess gender diversity by calculating male-to-female ratio. The database that I will use is World Gender Name Dictionary (1).

I will provide yearly trend for each of the abovementioned proportions.

Data analysis

I will use R (2) for searches, data handling, analysis and reporting. Web scrapping will be done using *rvest* package (3). Trends over time will be reported using descriptive tabulations and graphical illustrations, for instance using the *ggplot2* package (4).

The protocol of this study will be published beforehand on the Open Science Framework website (OSF, osf.io).

Limitations

This study includes only papers shown on the Cochrane Library website Search section. This section does not show previous versions of the Reviews. However, as the majority of these previous versions are authored by the same authors as the latest version, this will have small effect on the accuracy of the results of this study.

References

1. Raffo J. WGND 2.0 [Internet]. Harvard Dataverse; 2021. Available from: https://dataverse.harvard.edu/citation?persistentId=doi:10.7910/DVN/MSEGSJ
2. R Core Team. R: A language and environment for statistical computing. 2022; Available from: https://www.R-project.org/
3. Wickham H. Rvest: Easily harvest (scrape) web pages. 2021; Available from: https://CRAN.R-project.org/package=rvest
4. Wickham H. ggplot2: Elegant graphics for data analysis. 2016; Available from: https://ggplot2.tidyverse.org