

Addressing Bias in Health Applications: A Path Towards Health Equity

Keywords: Digital health, Bias, Health data, Health equity

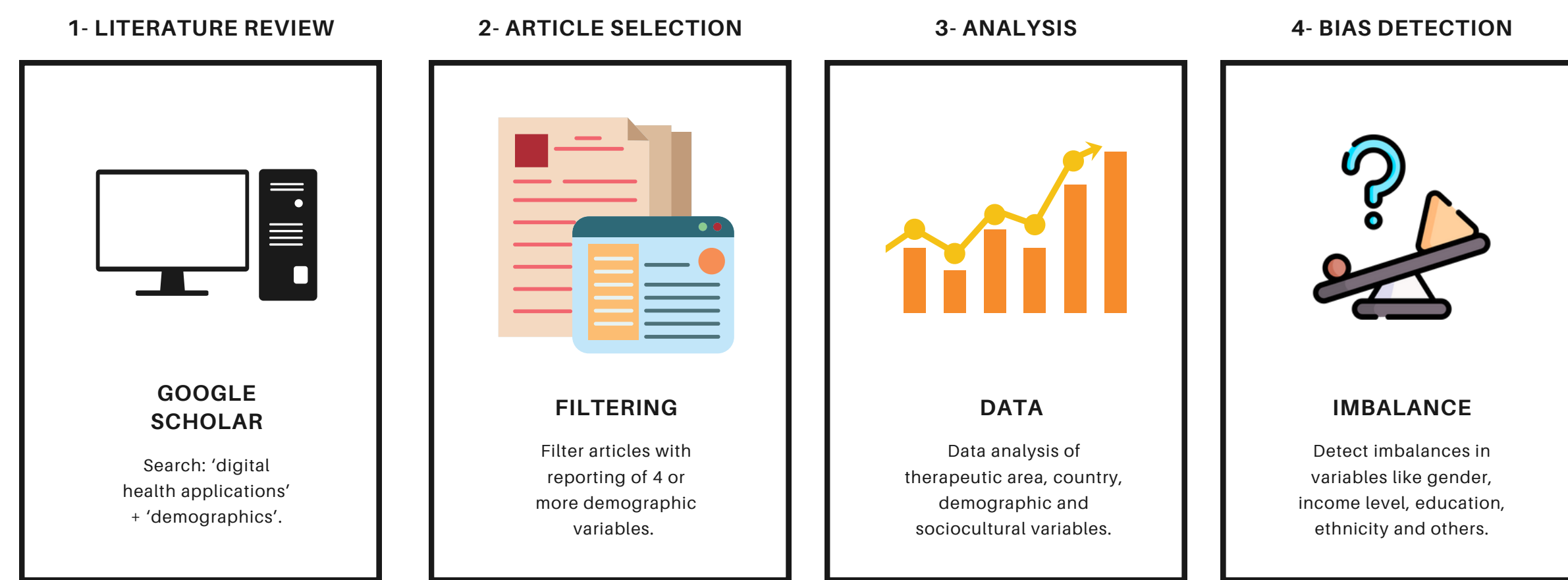


Figure 1. Illustration of the Literature Review Process for Bias Detection. This representation outlines the stages of article search, data analysis, and bias identification in articles describing the data collection methods of health apps.

01. Introduction

In the present era, our everyday use of health applications involves the collection of sensitive personal data. This data is subsequently utilized to train algorithms capable of shaping healthcare decisions. Should this collected information contain any bias, the resulting models could perpetuate these biases, greatly affecting individuals misrepresented in these datasets (1, 2). Predictive biases within algorithms can harm patients and, in turn, disrupt the equity of healthcare systems. Such inequity in healthcare can diminish quality of life, erode patient trust, and escalate healthcare costs.

02. Objective

The promise of big data has raised great expectations in healthcare research for drug discovery, as well as for personalized patient treatment and monitoring. This study aims to examine digital health applications that collect patient data in order to identify potential sources of bias in the resulting databases.

References

- Ledford, H. (2019). Millions affected by racial bias in health-care algorithm. *Nature*, 574(31), 2.
- Norori, N., Hu, Q., Aellen, F. M., Faraci, F. D., & Tzovara, A. (2021). Addressing bias in big data and AI for health care: A call for open science. *Patterns*, 2(10).

03. Methods

A literature search (Fig. 1) was conducted on Google Scholar using the keywords "digital health applications" and "demographics." Variables investigated included the surveyed population (country), sample size, and application type. Demographic variables analyzed within the surveyed population encompassed income level, education level, age, gender, disease status (ongoing diseases or reported healthy), and presence of ethnic disparities (including the number of analyzed ethnic groups, if applicable).

04. Results

Out of the 250 articles reviewed, 34 met the inclusion criteria, necessitating the presence of at least 4 demographic variables. These chosen articles predominantly explored healthy lifestyle (8/34) and diabetes care (5/34), with the United States taking the lead, contributing 10 out of the 34 articles (Fig. 2 and 3). The user profiles in these studies (Fig. 4) often mirrored individuals with significant purchasing power (15/34 vs. 2/34), advanced education (27/34 vs. 1/34), age under 40 (16/34 vs. 10/34), female gender (19/34 vs. 9/34), and self-reported unhealthy status (14/34 vs. 8/34). Remarkably, articles providing ethnicity data (12/34 vs. 23/34) highlighted disparities in minority group representation. Furthermore, only 7 out of 34 articles disclosed the number of ethnic groups considered when addressing ethnicity in their databases, indicating a clear misrepresentation of diverse groups in the collected data (Fig. 4).

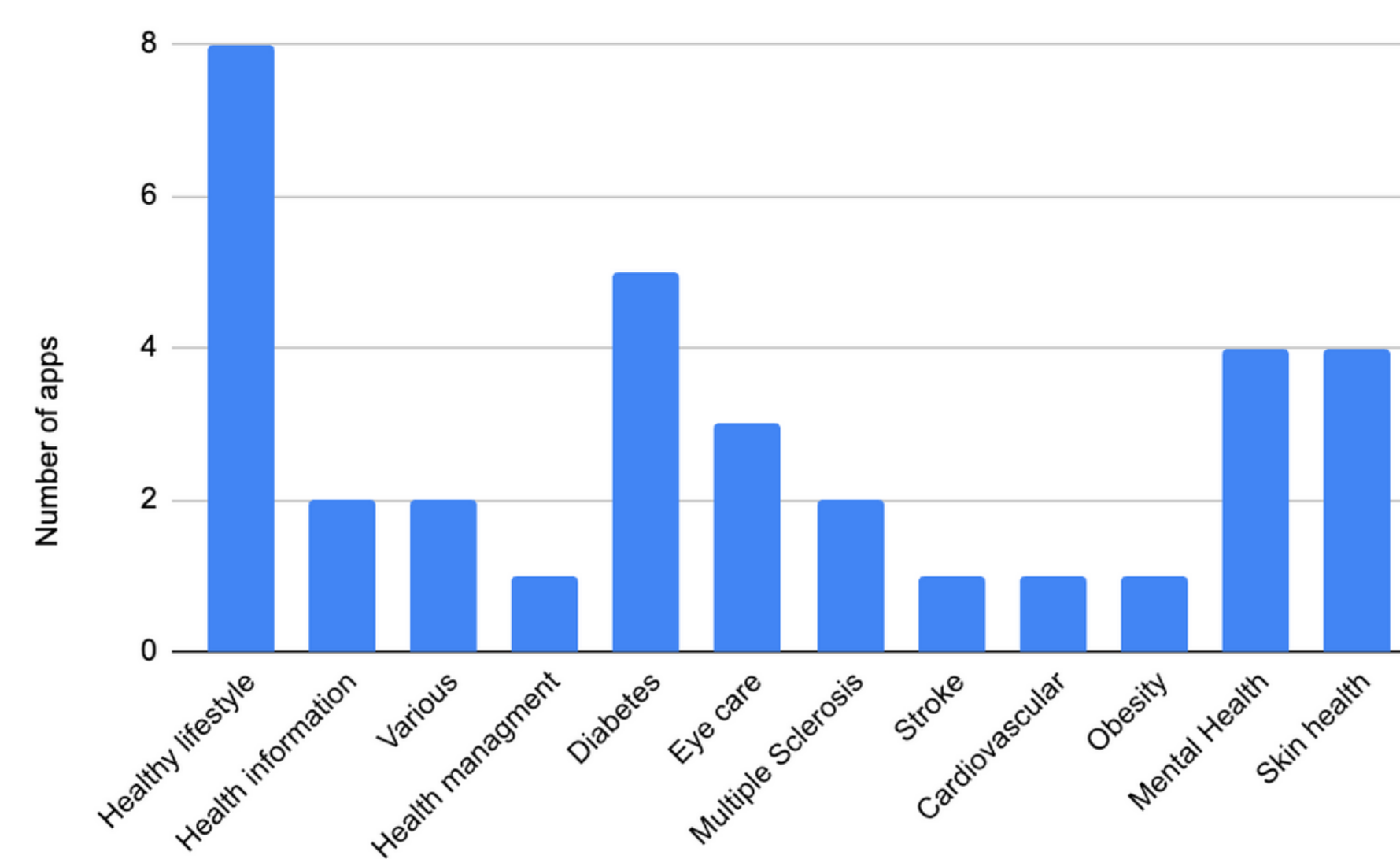


Figure 2. Distribution of Apps by Therapeutic Area. Category various encompass surveys of various apps across diverse therapeutic areas.

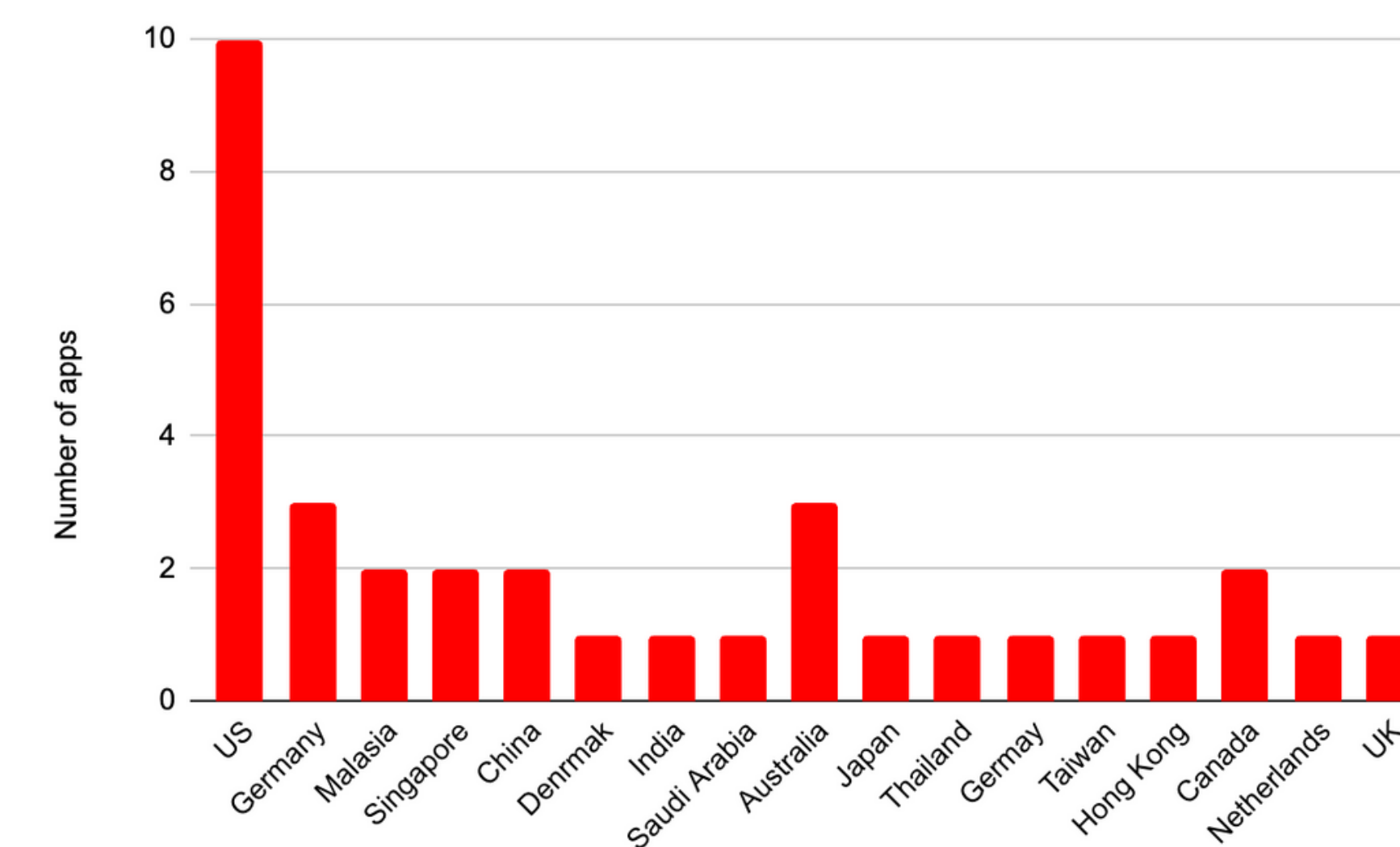


Figure 3. Distribution of Apps Based on Data Collection Country. This figure illustrates the population imbalance in the datasets across the world.

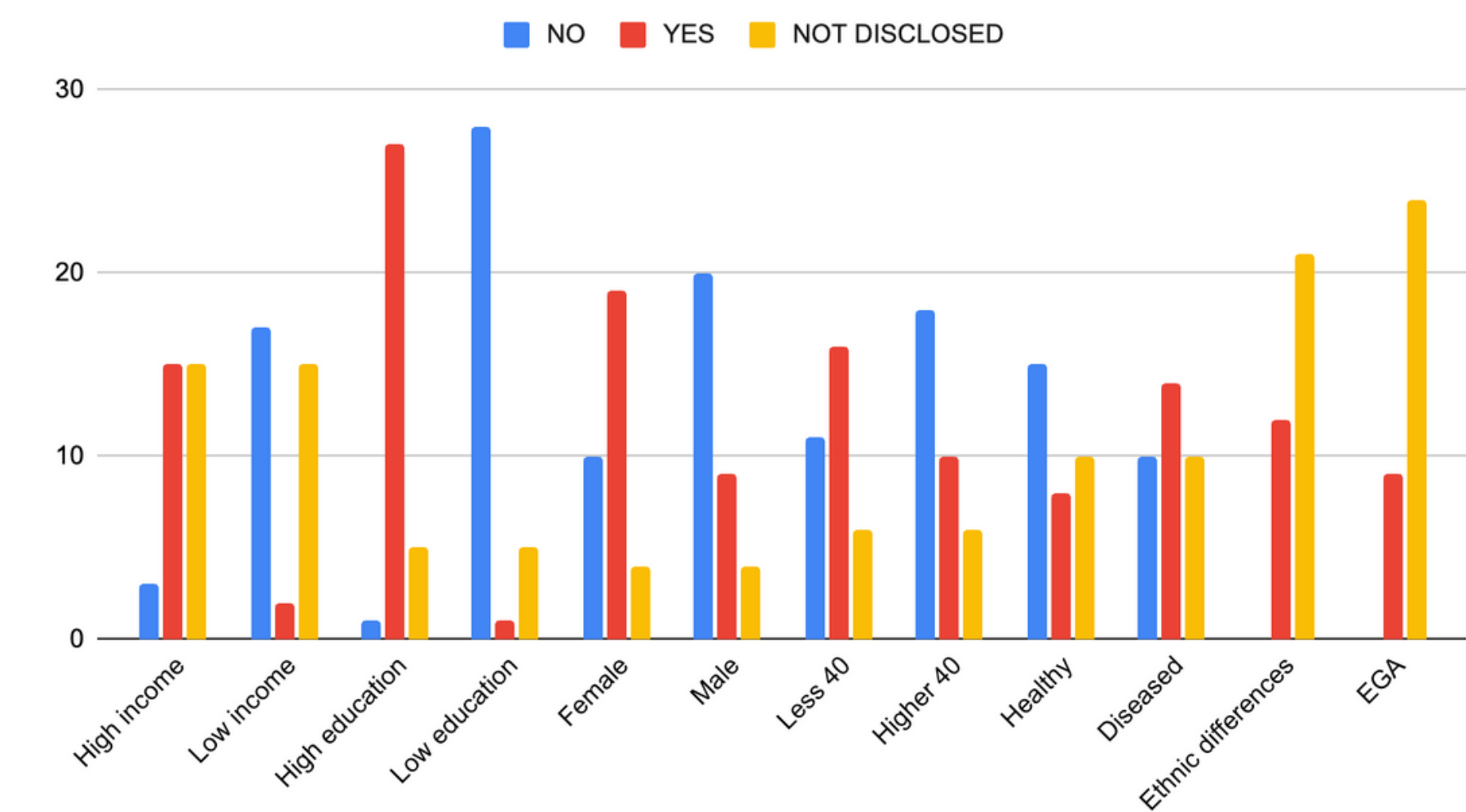


Figure 4. Demographic and Socioeconomic Information of Registered App Users in Publications. 'No' indicates apps lacking this trait, 'Yes' indicates apps presenting this data, and 'Not Disclosed' refers to apps omitting this information from their retrieval. ECA signifies details of assessed ethnic groups.

05. Conclusion

This study highlights the presence of various biases in databases constructed from health applications, which may negatively impact the development of secure and inclusive algorithms, consequently affecting health equity and the efficiency of health systems. Further research is necessary to expand the range of reviewed applications and encompass a broader spectrum of potential sources of bias, fostering a more equitable healthcare landscape and enhancing the efficiency of health systems.