Analyzing Diversity in Healthcare LLM Research: A Scientometric Perspective

David Restrepo[1*], Chenwei Wu[2], Constanza Vásquez-Venegas[3, 4], João Matos[1], Jack Gallifant[1], Luis Filipe Nakayama[1]

[1]Laboratory for Computational Physiology, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.
[2]Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, Michigan, United States of America.
[3]Scientific Image Analysis Lab, Faculty of Medicine, Universidad de Chile, Santiago, RM, Chile.
[4]Department of Computer Science, Faculty of Engineering, Universidad de Concepción, Concepción, Biobio, Chile.

*Corresponding author(s). E-mail(s): davidres@mit.edu;
Contributing authors: chenweiw@umich.edu; covasquezv@inf.udec.cl; jcmatos@mit.edu; jgally@mit.edu; luisnaka@mit.edu;

Abstract

The deployment of large language models (LLMs) in healthcare has demonstrated substantial potential for enhancing clinical decision-making, administrative efficiency, and patient outcomes. However, the underrepresentation of diverse groups in the development and application of these models can perpetuate biases, leading to inequitable healthcare delivery. This paper presents a comprehensive scientometric analysis of LLM research for healthcare, including data from January 1, 2021, to June 16, 2024. By analyzing metadata from PubMed and Dimensions, including author affiliations, countries, and funding sources, we assess the diversity of contributors to LLM research. Our findings highlight significant gender and geographic disparities, with a predominance of male authors and contributions primarily from high-income countries (HICs). We introduce a novel journal diversity index based on Gini impurity to measure the inclusiveness of scientific publications. Our results underscore the necessity for greater representation in order to ensure the equitable application of LLMs in healthcare. We propose actionable strategies to enhance diversity and inclusivity in artificial intelligence research, with the ultimate goal of fostering a more inclusive and equitable future in healthcare innovation.

Keywords: Large Language Models, Scientometrics, Review, Diversity

1 Introduction

The advent of large language models (LLMs) represents a significant breakthrough in artificial intelligence (AI), revolutionizing a wide range of applications across numerous fields[1–5]. In the context of healthcare, LLMs have demonstrated considerable potential in enhancing clinical decision-making, streamlining administrative processes, and improving patient outcomes through natural language processing tasks such as medical record analysis, automated diagnosis, and personalized treatment recommendations [6]. These advanced models, trained on vast corpora of text

data, possess the ability to understand, interpret, and generate human language with remarkable accuracy, thus offering promising solutions to some of the most pressing challenges in modern healthcare [7, 8].

However, the deployment of LLMs in healthcare is not without its challenges, particularly concerning the issue of representation within the machine learning and healthcare communities [9]. Underrepresentation of diverse groups in the development and application of LLMs can lead to the perpetuation or amplification of biases, stereotypes, and negative perceptions of minorities in society [10–13]. Such biases can adversely affect the quality and equity of healthcare delivery, potentially exacerbating existing disparities [14]. For instance, if LLMs are predominantly developed and trained by researchers from high-income countries (HICs), the resulting models may not adequately address or understand the healthcare needs of populations in low- and middle-income countries (LMICs). This lack of representation can lead to biased algorithms that fail to generalize across different demographic groups, thus limiting the global applicability and fairness of AI-driven healthcare solutions.

Addressing these challenges requires a concerted effort to promote diversity and inclusion in AI research and development [15]. One effective approach to highlight and address the lack of representation is through scientometric analysis — a method that involves the quantitative study of science, technology, and innovation [16]. By analyzing publication patterns, authorship demographics, and funding sources, scientometric analyses can provide critical insights into the evolving landscape of scientific research and its global impact [17]. This, in turn, can help identify gaps and biases in the current research landscape and inform strategies to foster greater diversity in the development of LLMs [18].

In this paper, we present a scientometric review of LLM research in healthcare, utilizing data extracted from PubMed and Dimensions API covering the period from January 1, 2021 to June 16, 2024. Our study involved extracting metadata such as authors' affiliations, countries, names, and funding information, and estimating the diversity of journals and research in the field of LLMs applied to healthcare. Through this analysis, we aim to shed light on the diversity of contributors to LLM research in healthcare and underscore the need for broader representation to ensure the equitable and effective application of LLMs worldwide. By evidencing the current state of representation and proposing actionable solutions, our study seeks to contribute to the ongoing discourse on diversity in AI and healthcare, ultimately advocating for a more inclusive and equitable future in health technology innovation.

2 Methods

2.1 Cohort Selection

The cohort for this study was defined through a systematic search of PubMed [19], a premier biomedical literature database, to identify research articles focused on the application of LLMs in healthcare. PubMed was chosen due to its comprehensive coverage of biomedical and life sciences literature, ensuring an extensive collection of relevant studies.

To construct the search query, we used specific terms to capture the essence of LLMs and their application in healthcare. The search terms included:

> (Large Language Models[Title/Abstract] OR LLMs[Title/Abstract] OR Language Models[Title/Abstract]) AND (Health care[Title/Abstract] OR Healthcare[Title/Abstract] OR Medicine[Title/Abstract] OR Medical[Title/Abstract] OR Clinical[Title/Abstract])

These terms were designed to be inclusive enough to cover various aspects and terminologies related to LLMs in healthcare. The search was restricted to articles published between January 1, 2021, and June 16, 2024, to capture recent advancements and trends in this rapidly evolving field. The retrieved articles formed the initial dataset for our scientometric analysis. The details of each article, including titles, abstracts, and publication dates, were extracted to facilitate further examination and classification. This initial extraction served as the foundation for subsequent steps in our analysis, where we aimed to gather comprehensive metadata about the authors and their affiliations to assess the diversity of research contributions in this domain.

2.2 Metadata Extraction and Preprocessing

Following the initial identification of relevant articles through PubMed, the next step involved the extraction of detailed metadata from the Dimensions database [20]. The PubMed ID (PMID) of each article was used to retrieve comprehensive information about the publications, including details on the authors, their affiliations, and funding sources. This extraction process utilized the Dimensions API, which allowed for efficient retrieval of metadata by querying the database with the specific PMIDs of the selected articles.

The data extracted from Dimensions included key information such as the name of the author, their respective affiliations, the position they held in the authoring process, and the countries in which they were based. Additionally, details about the funding organizations and the countries from which they hailed was included. The dataset enabled us to perform a thorough analysis of the geographic and demographic distribution of research contributions. To facilitate this analysis, we performed several preprocessing steps:

- **Gender Identification**: The Genderize.io API [21] was employed to infer the gender of the authors based on their first names. This API uses a large database of names to provide probabilistic estimates of gender, allowing us to classify authors as male or female using the maximum probability.
- **Country Classification**: Authors' countries were classified according to the World Bank's 2024 income classifications, which categorize countries into HIC or LMIC [22]. This classification enabled us to assess the representation of different income groups in LLM research.
- **Continent Grouping**: For a comprehensive geographic analysis, countries were grouped into continents. The continents used for this categorization were Africa, Asia, Europe, North America, Oceania, and South America. This categorization facilitated an understanding of the regional distribution of LLM research and the identification of any geographical disparities.

The preprocessing of the extracted data entailed the cleansing and organization of the data to ensure its accuracy and consistency. This included the handling of missing values, the standardizing country names, and the conversion of country codes to their respective ISO alpha-3 formats, which was achieved through the use of the 'pycountry' library. Additionally, the authors' names and affiliations were parsed to extract relevant information, and the data was structured to facilitate subsequent analysis.

2.3 Data Analysis

The data analysis conducted in this study was aimed at evaluating the diversity and representation in LLM research in healthcare, focusing on gender distributions, geographic distributions by income and continent, funder analysis, and journal diversity.

Gender Distributions: A gender estimation was conducted using the Genderize.io API, which classified genders according the methodology described in the previous section. Using the gender proportion, the number of female authors across the dataset and the position in the paper was used to calculate the overall proportions.

Continent and Income Group Distributions: From the Authors' information extracted in the previous section grouped by continents (Africa, Asia, Europe, North America, Oceania, South America) and classified into HIC and LMIC, The distribution of authors by continent and income group was analyzed to highlight geographic disparities in LLM research contributions.

Funding Source Analysis: The funding source analysis involved the identification of the countries from which the funding organizations resided, and categorizing them into HICs and LMICs. The proportions of publications funded by organizations from each income group were calculated in order to identify potential biases in research funding.

Distribution of Publications per Country: The distribution of publications per country was analyzed by aggregating the number of publications attributed to authors from each country. A choropleth map was employed to illustrate the geographic distribution of research contributions.

Journal Diversity and Index Testing: Journal diversity was assessed using a novel index based on Gini impurity, measuring the conditional probabilities of an author being of a certain gender, from a specific country, and from a particular income group, given that they are an author in a specific journal. The diversity index D_j for a journal j is calculated as:

$$D_j = 1 - \sum_{i=1}^{n} [P(G = i \mid J = j) \cdot P(C = i \mid J = j) \cdot P(I = i \mid J = j)]^2$$

where:

- $P(G = i \mid J = j)$ is the probability of an author being of gender i given they are an author in journal j.
- $P(C = i \mid J = j)$ is the probability of an author being from country i given they are an author in journal j.
- $P(I = i \mid J = j)$ is the probability of an author being from income group i given they are an author in journal j.

The index ranges from 0 (less diverse) to 1 (more diverse), with higher values indicating greater diversity. The conditional probabilities were computed using the formula:

$$P(X = i \mid J = j) = \frac{N_{ij}}{N_j}$$

where N_{ij} is the number of authors in category i for journal j, and N_j is the total number of authors for journal j.

The index was validated on journals with more than 100 authors to ensure robustness. This metric is a key contribution of our analysis, providing a quantitative measure of diversity in scientific publishing.

Overall, the data analysis yielded a comprehensive assessment of diversity and representation in LLM research in healthcare, identifying areas for improvement and proposing measures to enhance inclusivity in this critical field.

3 Results

3.1 Obtained Cohort

After running the search query on PubMed, we obtained a total of 1,274 articles.

3.2 Gender Distributions

The results showed disparities in gender representation across different author positions (overall, first authors, and last authors), and these distributions were further categorized by HICs and LMICs. Tables 1, 2, and 3 summarize the different gender distributions.

Table 1 Overall Gender Distribution of authors in the Dataset.

Author Position	Male (%)	Female (%)
Overall	76.29	23.71
First Authors	79.61	20.39
Last Authors	83.40	16.60

Table 2 Gender Distribution of atuhors in HIC papers

Author Position	Male (%)	Female (%)
Overall	76.55	23.45
First Authors	79.31	20.69
Last Authors	84.38	15.62

Table 3 Gender Distribution of authors in LMIC papers

Author Position	Male (%)	Female (%)
Overall	74.87	25.13
First Authors	81.48	18.52
Last Authors	76.67	23.33

3.3 Distribution of Papers per Continent for LMICs and HICs

The distribution of papers was analyzed by continent and income group (HICs and LMICs). The results are summarized in the Table 4. Additionally, a world map illustrating the distribution of research contributions across continents provides a geographic perspective on paper diversity. (Figure 1)

Table 4 Distribution of Papers by Continent and Country Income Group

Continent	Income Group	N	Proportion (%)
Africa	LMIC	26	0.29
Asia	HIC	703	7.77
Asia	LMIC	1066	11.79
Europe	HIC	1808	19.99
Europe	LMIC	21	0.23
North America	HIC	5197	57.46
Oceania	HIC	176	1.95
South America	HIC	10	0.11
South America	LMIC	38	0.42

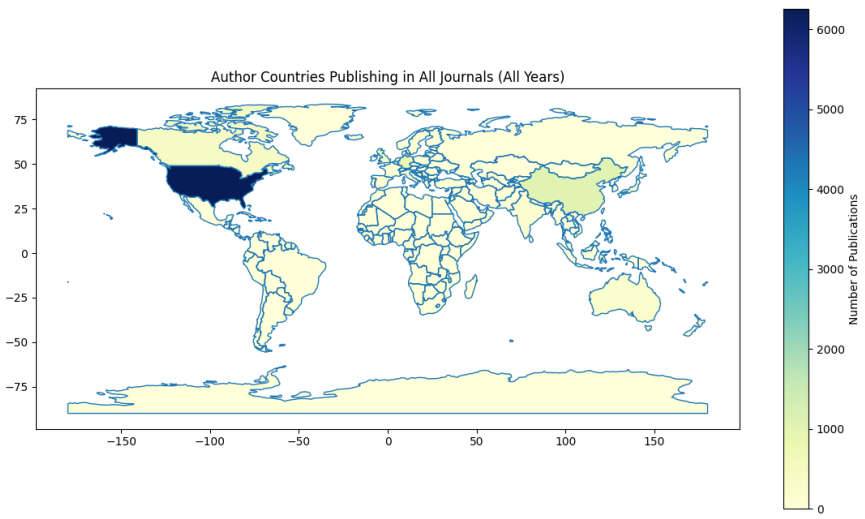

Fig. 1 World Map Distribution of Research Contributions

3.4 Funding Source Analysis

The analysis of funding sources was categorized by continent and income group. The results indicate potential biases in the distribution of funding across different regions. Table 5 summarizes these findings.

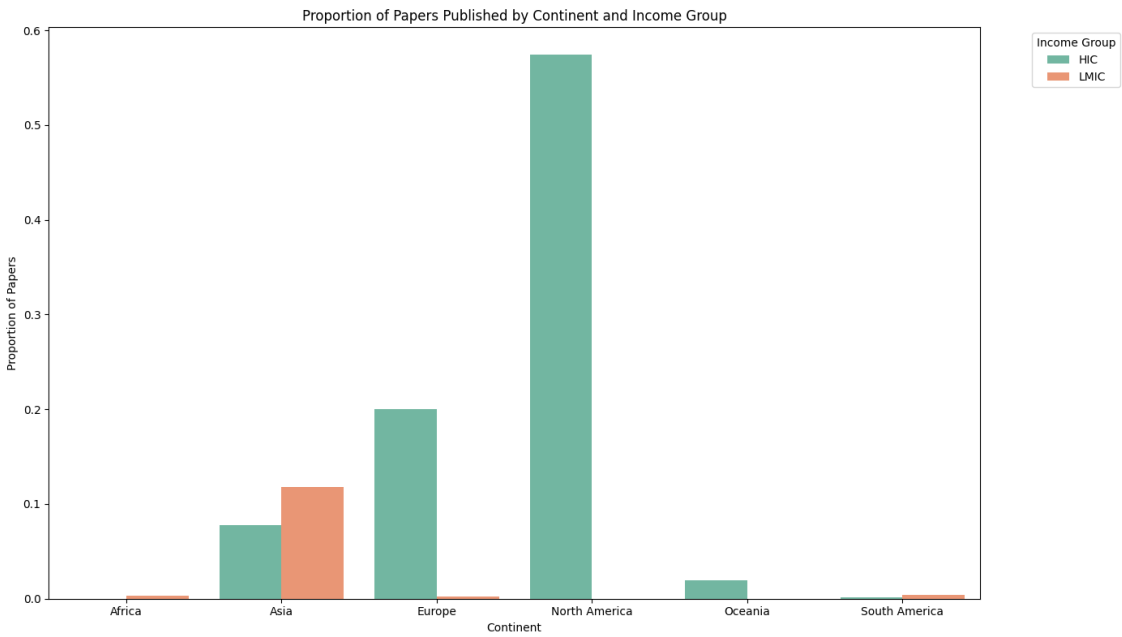

Fig. 2 Distribution of Papers Published by Continent and Income Group

Table 5 Distribution of Funding Countries by Continent and Income Group

Continent	Income Group	N	Proportion (%)
Africa	LMIC	1	0.21
Asia	HIC	39	8.25
Asia	LMIC	52	10.99
Europe	HIC	88	18.61
North America	HIC	286	60.47
Oceania	HIC	4	0.85
South America	HIC	1	0.21
South America	LMIC	2	0.43

3.5 Journal Diversity Index

The diversity of journals assessed using our diversity index based on Gini impurity calculated for journals with more than 100 authors was represented in a table and a map. A summary of the diversity index for selected journals is presented in Table 6. The results in Table 6 re[resent the diversity index having values close to 1 as more diverse, and values close to 0 as less diverse.

4 Discussion

The findings of our scientometric analysis reveal significant disparities in gender and geographic representation in the research on LLMs applications to healthcare. These disparities have important implications for the development and deployment of LLMs in healthcare, particularly with regard to their generalizability, usability, and fairness.

4.1 Gender Representation

Our analysis shows that the representation of female authors in LLM research is consistently low across all datasets. Overall, females constitute only 23.71% of the authors, and their representation further diminishes in key author positions, such as first and last authors, with proportions of 20.39% and 16.60%, respectively. This underrepresentation is observed in both HICs and LMICs, though the extent of the disparity varies. In HICs, female authors make up 23.45% overall, 20.69% of first authors, and 15.62% of last authors. In LMICs, the proportions are slightly higher overall (25.13%) but lower for first authors (18.52%) and last authors (23.33%).

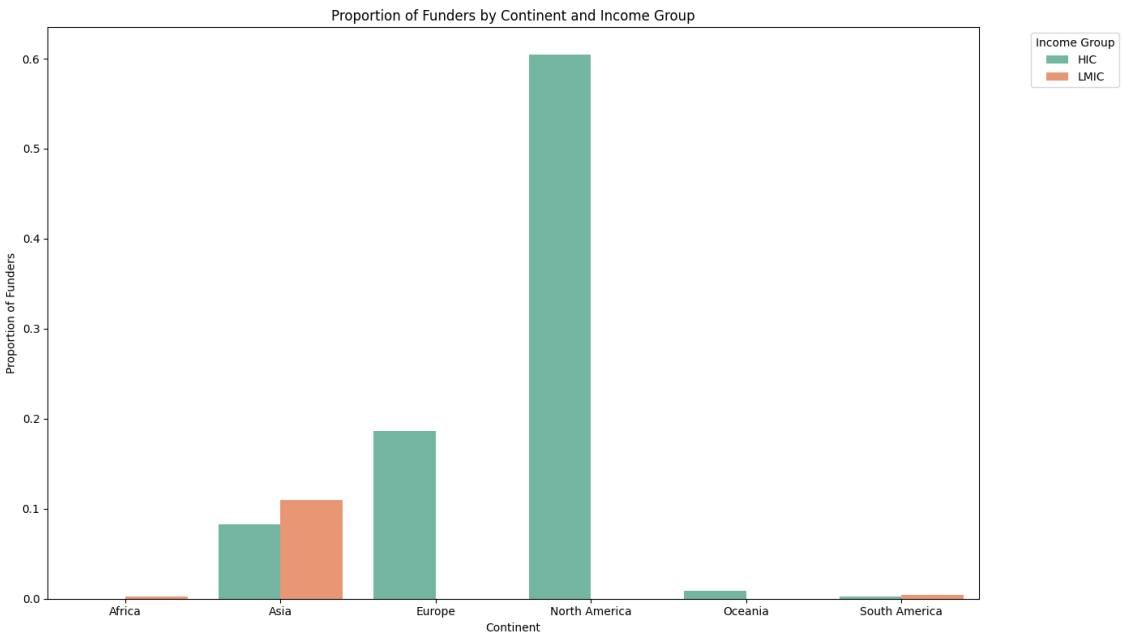

Fig. 3 Proportion of Funding Source by Continent and Income Group

Table 6 Diversity Index for Selected Journals (Ordered from Higher to Lower)

Journal Title	Diversity Index
Artificial Intelligence in Medicine	0.823
Journal of Medical Internet Research	0.822
PLOS Digital Health	0.821
JMIR Medical Education	0.808
Scientific Reports	0.795
Cureus	0.771
Journal of Medical Systems	0.765
JAMA Network Open	0.735
The Lancet Digital Health	0.716
JMIR Medical Informatics	0.707
BMC Medical Informatics and Decision Making	0.697
Journal of Biomedical Informatics	0.685
Computers in Biology and Medicine	0.644
Nucleic Acids Research	0.639
AMIA Annual Symposium Proceedings	0.610
Journal of the American Medical Informatics Association	0.597
European Archives of Oto-Rhino-Laryngology	0.580
bioRxiv	0.503
npj Digital Medicine	0.502
Research Square	0.499
Journal of the American Academy of Dermatology	0.440
medRxiv	0.421
AMIA Joint Summits on Translational Science Proceedings	0.411
Journal of the American College of Radiology	0.331
Nature	0.302

The lower representation of women, particularly in prominent author positions, could limit the diversity of perspectives and expertise in LLM research, potentially leading to biased outcomes in healthcare applications. It is crucial to address these gender disparities to ensure that the development of LLMs benefits from a wide range of viewpoints and experiences, thereby enhancing the models' fairness and applicability across diverse populations.

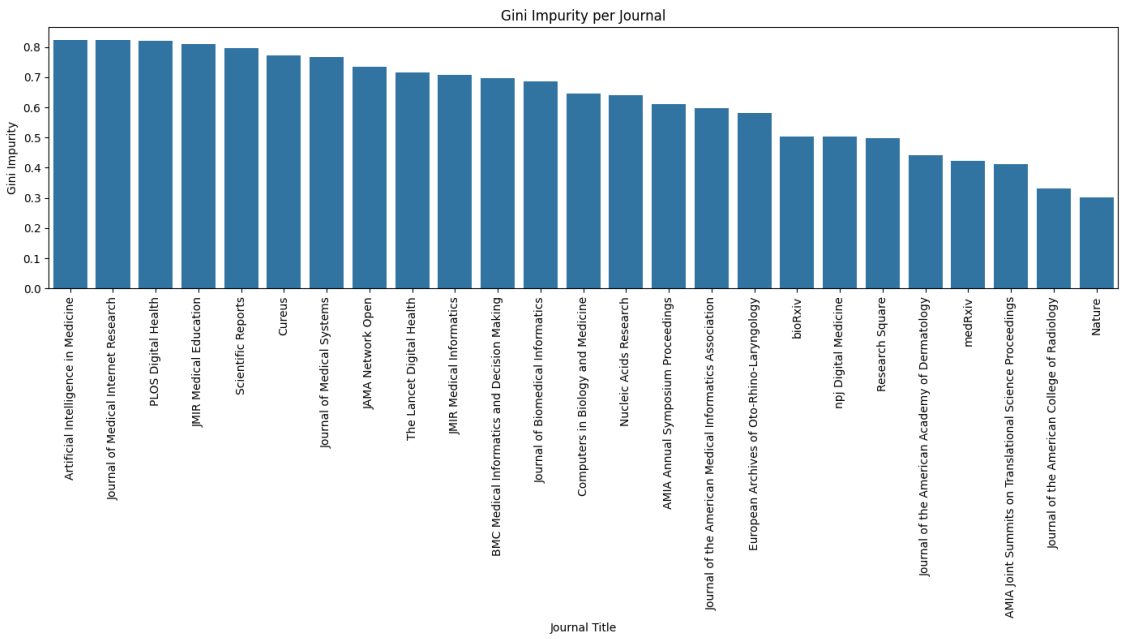

Fig. 4 Diversity Metric per Journal

4.2 Geographic and Income Group Representation

The geographic distribution of LLM research contributions highlights a significant imbalance favoring authors from HICs. Our data indicate that North America (HIC) dominates the research landscape, accounting for 57.46% of the publications, followed by Europe (HIC) with 19.99%. In contrast, contributions from LMICs are markedly lower. For instance, Asia (LMIC) contributes 11.79%, and Africa (LMIC) a mere 0.29%. This skewed distribution raises concerns about the global applicability of LLMs, as models predominantly developed in and for HIC contexts may not address the unique healthcare challenges and needs of LMIC populations.

The disparity is also evident in the funding sources. The majority of research funding comes from organizations based in HICs, with North America (HIC) again leading at 60.47%, and Europe (HIC) at 18.60%. LMIC funding contributions are minimal, with Asia (LMIC) at 10.99% and Africa (LMIC) at just 0.21%. This funding imbalance can further exacerbate the inequities in research outputs, limiting the ability of LMICs to conduct and contribute to high-quality LLM research. The lack of representation from LMICs, particularly in Latin America confirming the lack of resources for research in the region [23]. These challenges represent a significant risk to the generalizability and usability of LLMs in these regions, potentially perpetuating healthcare inequities.

4.3 Journal Diversity Index

The novel journal diversity index based on Gini impurity provides a quantitative measure of diversity within scientific publishing. Our analysis reveals considerable variability in diversity across journals. Journals such as "Artificial Intelligence in Medicine" (0.823901), "Journal of Medical Internet Research" (0.822026), and "PLOS Digital Health" (0.821324) exhibit high diversity, indicating a more balanced representation of authors from different genders, countries, and income groups. Conversely, journals like "Nature" (0.302608) and "Journal of the American College of Radiology" (0.331631) show lower diversity, highlighting the need for greater inclusivity.

Promoting diversity in scientific journals is essential for democratizing access to healthcare and technologies like LLMs, particularly in low-resource settings. Initiatives aimed at enhancing diversity should be implemented across journals to ensure that research outputs reflect a broad spectrum of perspectives and are applicable to diverse populations worldwide. This can help mitigate biases and enhance the generalizability and fairness of LLMs, ultimately contributing to more equitable healthcare outcomes.

4.4 Recommendations

Based on our findings, we recommend several actions to address the disparities in gender and geographic representation in LLM research in healthcare.

Institutional policies and initiatives to support women in research: Academic institutions and research organizations should implement targeted policies and initiatives to encourage and support the participation of women in AI and healthcare research. This can include establishing mentorship programs that pair female researchers with experienced mentors, providing dedicated funding and grants for women-led research projects, and ensuring gender balance in research teams and leadership positions [24]. These initiatives should be accompanied by efforts to address systemic barriers and biases that hinder women's career advancement in research [25].

Equitable distribution of research funds: Funding bodies and agencies should prioritize the equitable distribution of research funds, supporting projects that involve LMIC researchers and address healthcare challenges pertinent to low-resource settings. This can be achieved through targeted funding calls, quotas for LMIC participation, and the inclusion of LMIC researchers in grant review panels. Funding agencies should also consider the unique constraints and challenges faced by LMIC researchers, such as limited access to computational resources and data that are particularly common in LLM research, and provide appropriate support and accommodations[26].

Collaborative research networks and partnerships: To involve researchers from LMICs in global LLM research projects, collaborative research networks and partnerships between HIC and LMIC institutions should be established and strengthened [27]. These collaborations can facilitate knowledge transfer, capacity building, and the exchange of expertise and resources. By engaging LMIC researchers as equal partners in research projects, these collaborations can help to ensure that LLM technologies are developed and applied in a manner that is relevant and beneficial to diverse healthcare contexts [28].

Diversity policies in scientific publishing: Scientific journals and publishing bodies should adopt policies that promote diversity in authorship and editorial boards. This can include implementing diversity metrics, such as tracking the gender and geographic representation of authors and reviewers, and setting targets for improvement. Journals should also encourage submissions from underrepresented regions and groups, potentially through special issues or dedicated submission tracks, such as IEEE ICHI 2024's first workshop of LLM applications in LMIC. Editorial boards should strive for diverse representation and actively seek out expertise from LMICs. Journals, editors, reviewers, and grantors can mandate that author teams disclose their goals for achieving diversity in their research groups and collaborations [29]. This practice, similar to the National Institutes of Health's (NIH) approach for conference grant applications, would promote transparency, accountability, and prioritization of diversity and inclusion in research.

Provide formal training on equity and diversity: Incorporating formal training on equity issues and the importance of diversity in the research process into the syllabi of academic institutions can help educate the next generation of researchers [30–32]. This training should highlight the benefits of diversity in working groups and published work, such as increased scientific impact and more effective complex problem-solving.

4.5 Limitations

While our study provides valuable insights into the diversity of LLM research in healthcare, it is important to acknowledge several limitations. The gender of authors was inferred using the Genderize.io API based on their first names. This method may introduce biases and inaccuracies, particularly for names that are unisex, culturally specific, or less common. Additionally, it does not account for non-binary and other gender identities. Country information was collected based on authors' affiliations, which may not accurately reflect their nationality or cultural background. This can lead to misclassification, particularly for authors with affiliations in multiple countries or those working abroad.

Our analysis was based on data extracted from PubMed [33], which primarily indexes biomedical literature. Consequently, research articles published in journals not indexed by PubMed were not included in our study. This exclusion may result in an incomplete representation of the global research landscape. The study focused on articles published between January 1, 2021, and June

16, 2024. Research trends and representation patterns may evolve over time, and future studies should consider broader temporal scopes to capture these dynamics.

Despite these limitations, our findings highlight critical areas for improvement in promoting diversity and inclusivity in LLM research in healthcare. Addressing these challenges is essential for developing AI technologies that are fair, generalizable, and beneficial to diverse populations worldwide.

5 Conclusion

Our scientometric review demonstrates the necessity for greater representation and inclusivity in LLM research in healthcare. Addressing gender and geographic disparities is crucial for developing fair and generalizable AI models that can benefit diverse populations. By promoting diversity in research contributions and funding, and by applying robust measures such us the journal diversity index, we can work towards more equitable and effective healthcare innovations.

In conclusion, this study provides critical insights into the current state of diversity in LLM research and proposes actionable strategies to enhance inclusivity. Future efforts should focus on implementing these strategies to ensure that the benefits of AI-driven healthcare are accessible to all, regardless of gender, geographic location, or economic status.

References

[1] Liu, Y., Han, T., Ma, S., Zhang, J., Yang, Y., Tian, J., He, H., Li, A., He, M., Liu, Z., Wu, Z., Zhao, L., Zhu, D., Li, X., Qiang, N., Shen, D., Liu, T., Ge, B.: Summary of ChatGPT-Related research and perspective towards the future of large language models. Meta-Radiology **1**(2), 100017 (2023) https://doi.org/10.1016/j.metrad.2023.100017 . Accessed 2024-06-18

[2] Chen, Y., Liu, C., Huang, W., Cheng, S., Arcucci, R., Xiong, Z.: Generative text-guided 3d vision-language pretraining for unified medical image segmentation. arXiv preprint arXiv:2306.04811 (2023)

[3] Mei, T., Zi, Y., Cheng, X., Gao, Z., Wang, Q., Yang, H.: Efficiency optimization of large-scale language models based on deep learning in natural language processing tasks. arXiv preprint arXiv:2405.11704 (2024)

[4] Yao, Z., Lin, F., Chai, S., He, W., Dai, L., Fei, X.: Integrating medical imaging and clinical reports using multimodal deep learning for advanced disease analysis. arXiv preprint arXiv:2405.17459 (2024)

[5] Xiao, L., Li, M., Feng, Y., Wang, M., Zhu, Z., Chen, Z.: Exploration of attention mechanism-enhanced deep learning models in the mining of medical textual data. arXiv preprint arXiv:2406.00016 (2024)

[6] Thirunavukarasu, A.J., Ting, D.S.J., Elangovan, K., Gutierrez, L., Tan, T.F., Ting, D.S.W.: Large language models in medicine. Nature Medicine **29**(8), 1930–1940 (2023) https://doi.org/10.1038/s41591-023-02448-8 . Accessed 2024-06-18

[7] Moor, M., Banerjee, O., Abad, Z.S.H., Krumholz, H.M., Leskovec, J., Topol, E.J., Rajpurkar, P.: Foundation models for generalist medical artificial intelligence. Nature **616**(7956), 259–265 (2023) https://doi.org/10.1038/s41586-023-05881-4

[8] Khurana, D., Koli, A., Khatter, K., Singh, S.: Natural language processing: state of the art, current trends and challenges. Multimedia Tools and Applications **82**(3), 3713–3744 (2023) https://doi.org/10.1007/s11042-022-13428-4

[9] Raiaan, M.A.K., Mukta, M.S.H., Fatema, K., Fahad, N.M., Sakib, S., Mim, M.M.J., Ahmad, J., Ali, M.E., Azam, S.: A Review on Large Language Models: Architectures, Applications, Taxonomies, Open Issues and Challenges. IEEE Access **12**, 26839–26874 (2024) https://doi.org/10.1109/ACCESS.2024.3365742 . Accessed 2024-06-18

[10] Dychiao, R.G.K., Alberto, I.R.I., Artiaga, J.C.M., Salongcay, R.P., Celi, L.A.: Large language model integration in Philippine ophthalmology: early challenges and steps forward. The Lancet Digital Health **6**(5), 308 (2024) https://doi.org/10.1016/S2589-7500(24)00064-5 . Accessed 2024-06-18

[11] Bender, E.M., Gebru, T., McMillan-Major, A., Shmitchell, S.: On the dangers of stochastic parrots: Can language models be too big? In: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, pp. 610–623 (2021)

[12] Blodgett, S.L., Barocas, S., Daumé III, H., Wallach, H.: Language (technology) is power: A critical survey of" bias" in nlp. arXiv preprint arXiv:2005.14050 (2020)

[13] Smith, E.M., Hall, M., Kambadur, M., Presani, E., Williams, A.: " i'm sorry to hear that": Finding new biases in language models with a holistic descriptor dataset. arXiv preprint arXiv:2205.09209 (2022)

[14] Rajkomar, A., Hardt, M., Howell, M.D., Corrado, G., Chin, M.H.: Ensuring Fairness in Machine Learning to Advance Health Equity. Annals of Internal Medicine **169**(12), 866 (2018) https://doi.org/10.7326/M18-1990 . Accessed 2024-06-18

[15] Swartz, T.H., Palermo, A.-G.S., Masur, S.K., Aberg, J.A.: The Science and Value of Diversity: Closing the Gaps in Our Understanding of Inclusion and Diversity. The Journal of Infectious Diseases **220**(Supplement_2), 33–41 (2019) https://doi.org/10.1093/infdis/jiz174 . Accessed 2024-06-18

[16] Mingers, J., Leydesdorff, L.: A review of theory and practice in scientometrics. European Journal of Operational Research **246**(1), 1–19 (2015) https://doi.org/10.1016/j.ejor.2015.04.002 . Accessed 2024-06-18

[17] Charpignon, M.-L., Matos, J., Nakayama, L., Gallifant, J., Alfonso, P.G.I., Cobanaj, M., Fiske, A., Gates, A.J., Ho, F.D.V., Jain, U., Kashkooli, M., McCoy, L.G., Shaffer, J., Link Woite, N., Celi, L.A.: Does diversity beget diversity? a scientometric analysis of over 150,000 studies and 49,000 authors published in high-impact medical journals between 2007 and 2022. medRxiv [Preprint] **2024**(2024.03.21.24304695) (2024) https://doi.org/10.1101/2024.03.21.24304695

[18] Adebayo, S.O., Saini, M.: A scientometric study for scientific research publication on gender inequality. Quality & Quantity **57**(6), 5107–5135 (2023) https://doi.org/10.1007/s11135-022-01604-5 . Accessed 2024-06-18

[19] National Center for Biotechnology Information (NCBI): PubMed. https://pubmed.ncbi.nlm.nih.gov/. Accessed: Accessed: June 17, 2024

[20] Dimensions. https://app.dimensions.ai/auth/base/landing?redirect=%2Fdiscover%2Fpublication Accessed 2024-06-18

[21] Genderize.io: Genderize.io. https://genderize.io/. Accessed: Accessed: June 17, 2024

[22] World Bank Country and Lending Groups – World Bank Data Help Desk. https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups Accessed 2024-06-18

[23] Restrepo, D., Quion, J., Vásquez-Venegas, C., Villanueva, C., Anthony Celi, L., Nakayama, L.F.: A scoping review of the landscape of health-related open datasets in latin america. PLOS Digital Health **2**(10), 0000368 (2023)

[24] Ceci, S.J., Williams, W.M.: Are claims of fairness toward women in the academy "manufactured"? the risk of basing arguments on incomplete data. Sexuality & Culture **28**(1), 1–20 (2024)

[25] Sato, S., Gygax, P.M., Randall, J., Schmid Mast, M.: The leaky pipeline in research grant peer review and funding decisions: challenges and future directions. Higher Education **82**(1), 145–162 (2021)

[26] Charani, E., Abimbola, S., Pai, M., Adeyi, O., Mendelson, M., Laxminarayan, R., Rasheed, M.A.: Funders: The missing link in equitable global health research? PLOS global public health **2**(6), 0000583 (2022)

[27] Bowsher, G., Papamichail, A., El Achi, N., Ekzayez, A., Roberts, B., Sullivan, R., Patel, P.: A narrative review of health research capacity strengthening in low and middle-income countries: lessons for conflict-affected areas. Globalization and health **15**, 1–13 (2019)

[28] Meissner, P., Wulf, T.: The effect of cognitive diversity on the illusion of control bias in strategic decisions: An experimental investigation. European Management Journal **35**(4), 430–439 (2017)

[29] Jones, P.L., Sauma, S., Bernard, M.A.: Perspective: scientific workforce diversity and its impact on aging research. Oxford University Press US (2022)

[30] Koolen, C., Van Cranenburgh, A.: These are not the stereotypes you are looking for: Bias and fairness in authorial gender attribution. In: Proceedings of the First Ethics in NLP Workshop, pp. 12–22 (2017). Association for Computational Linguistics (ACL)

[31] Zhang, F., Shuai, Z., Kuang, K., Wu, F., Zhuang, Y., Xiao, J.: Unified fair federated learning for digital healthcare. Patterns **5**(1) (2024)

[32] Beddoes, K., Schimpf, C.: What's wrong with fairness? how discourses in higher education literature support gender inequalities. Discourse: Studies in the Cultural Politics of Education **39**(1), 31–40 (2018)

[33] Canese, K., Weis, S.: Pubmed: the bibliographic database. The NCBI handbook **2**(1) (2013)