MedCrave
Step into the World of Research
Journal of Historical Archaeology & Anthropological Sciences

Research Article

Open Access

Artificial intelligence training and population bias: underrepresentation of Latin America in gender, bioethical and clinical decision algorithms

Abstract

Background: Artificial intelligence systems are increasingly integrated into clinical and decision-making processes, yet their performance is strongly influenced by the representativeness of training data. Concerns have emerged regarding algorithmic bias affecting gender, ethnicity, and geographic populations, particularly in underrepresented regions such as Latin America. Methods: A mini-review was conducted using the Google Scholar database to identify review articles published between 2024 and 2026 addressing artificial intelligence, bias, clinical applications, and population representation. A total of 2,820 records were identified, from which the 20 most relevant articles were screened according to predefined search terms and eligibility criteria. Following exclusion procedures, 18 review articles were included and categorized based on the presence and combination of search terms in their titles. Results: No articles simultaneously included all predefined search terms. The most frequent category involved studies addressing artificial intelligence and bias, whereas fewer articles incorporated clinical or algorithmic dimensions. Only a limited proportion explicitly referenced Latin America, and none integrated this regional focus with bias, gender, and clinical decision-making variables. The distribution of terms revealed a fragmented literature landscape with limited multidimensional integration. Conclusion: The findings suggest the existence of an important gap in the current literature, indicating that the limited representation of Latin American populations may contribute to persistent gender, bioethical, and clinical biases within artificial intelligence decision algorithms. These results highlight the importance of developing more inclusive and context-sensitive research frameworks.

Keywords: artificial intelligence, algorithmic bias, Latin American population, clinical decision algorithms, bioethics, gender bias, data underrepresentation, healthcare equity

Volume 11 Issue 2 - 2026

Pedro Hecht-López
Department of Biomedical Sciences, Faculty of Medicine, University of Tarapacá, Arica, Chile

Correspondence: Pedro Hecht-López, Anatomy Laboratory, Faculty of Medicine, University of Tarapacá, Arica, Chile, Tel +56582385777

Received: April 25, 2026 | **Published:** May 28, 2026

Abbreviations: AI, artificial intelligence

Introduction

Artificial intelligence (AI) has become a central component in contemporary biomedical, clinical, and decision-support systems, fundamentally transforming the ways in which data are processed, interpreted, and operationalized across healthcare environments. Machine learning algorithms, particularly those trained on large-scale datasets, are increasingly deployed to assist in diagnosis, prognosis, risk stratification, and therapeutic decision-making. These systems are often perceived as objective and data-driven; however, their performance and outputs are inherently dependent on the characteristics, quality, and representativeness of the data used during the training phase. Consequently, structural imbalances in training datasets may translate into systematic biases embedded within algorithmic architectures, thereby reproducing or amplifying pre-existing social, demographic, and geographic inequities.[1]

Recent advances in the field have emphasized the presence of algorithmic bias across multiple domains, particularly in relation to gender, race, and ethnicity. Evidence indicates that AI systems frequently exhibit differential performance when applied to heterogeneous populations, often demonstrating reduced accuracy and increased error rates among underrepresented groups. This phenomenon is especially critical in clinical contexts, where biased outputs may directly influence patient outcomes and healthcare delivery. The integration of AI into clinical decision algorithms has therefore raised substantial concerns regarding fairness, transparency, and accountability, particularly when such systems are trained on datasets that do not adequately reflect global population diversity.[2,3] Within this context, the underrepresentation of Latin American populations constitutes a significant and insufficiently addressed dimension of algorithmic bias. Despite representing a substantial proportion of the global population, Latin America remains marginally represented in many biomedical datasets and AI training cohorts. This disparity is not only quantitative but also qualitative, encompassing variations in genetic backgrounds, epidemiological profiles, socioeconomic determinants, and healthcare system structures. The absence of such contextual diversity in training data limits the generalizability of AI models and may result in suboptimal or inequitable clinical recommendations when these systems are applied to Latin American populations.[2,3]

The current state of the literature demonstrates a growing awareness of bias in AI systems; however, most research has been concentrated in high-income countries, particularly in North America, Europe, and parts of Asia. Studies analyzing racial and ethnic bias in AI frequently focus on comparisons between Black and White populations within the United States, with limited attention to other ethnic groups or geographic regions. Similarly, investigations into gender bias have predominantly examined binary classifications without adequately addressing intersectional dimensions that incorporate ethnicity, socioeconomic status, and regional disparities. As a result, the specific challenges associated with Latin American

underrepresentation remain largely underexplored, both in empirical research and in theoretical frameworks addressing AI ethics.[4] Recent reports focused on Latin America and the Caribbean have further emphasized the relevance of gender bias within AI systems and the unequal regional implications of emerging technologies. Clavell highlighted that AI-driven decision systems increasingly influence social, economic, and healthcare dynamics in the region, while women remain particularly vulnerable to structural forms of algorithmic discrimination due to existing socioeconomic and digital inequalities. The report also emphasized that the limited participation of Latin American populations in AI development and governance may contribute to the persistence of demographic and gender-related biases, reinforcing the need for region-specific ethical and regulatory frameworks.[5]

Recent analyses examining AI and healthcare disparities have also emphasized the potential risks associated with algorithmic bias in medical decision-making systems. Williams noted that although AI technologies may improve diagnostic accuracy, reduce healthcare costs, and optimize patient outcomes, these benefits may not be equally distributed across populations when training datasets insufficiently represent racial, ethnic, and socially vulnerable groups. The author further highlighted that structural inequities, social determinants of health, and historical patterns of exclusion may contribute to biased algorithmic outputs, thereby limiting the capacity of AI systems to reduce healthcare disparities.[6] Furthermore, bioethical considerations related to AI deployment have gained prominence, particularly in relation to principles such as justice, beneficence, and non-maleficence. The use of biased algorithms in clinical settings raises critical ethical questions regarding the distribution of risks and benefits across populations. If AI systems systematically underperform in certain demographic groups due to insufficient representation during training, their implementation may inadvertently reinforce health disparities rather than mitigate them. This concern is particularly relevant in Latin America, where existing inequalities in healthcare access and outcomes may be exacerbated by the adoption of inadequately validated AI technologies.[7]

Recent policy-oriented analyses have also emphasized that algorithmic bias in healthcare AI should be approached not only as a technical limitation but also as a governance and health systems challenge.[8] highlighted that biased datasets may compromise the accuracy and generalizability of AI systems across diverse populations, particularly in low- and middle-income settings within the Americas. The authors further argued that representation bias, deployment bias, and structural inequities may reinforce disparities in healthcare delivery if fairness and inclusivity are not incorporated as measurable components throughout the AI lifecycle, including validation, implementation, and postmarket evaluation.[8] Another important dimension involves the transparency and reporting practices associated with AI model development. A significant proportion of studies do not disclose key demographic characteristics of the datasets used for training, including gender distribution, racial composition, and geographic origin. This lack of transparency hinders the ability to evaluate the external validity and ethical implications of AI systems, as well as to identify potential sources of bias. Moreover, the limited use of publicly available datasets further restricts opportunities for independent validation and reproducibility, thereby complicating efforts to ensure equitable model performance across diverse populations.[2]

In addition to data-related limitations, methodological heterogeneity across studies contributes to inconsistencies in the identification and mitigation of bias. Variations in study design, inclusion criteria, and analytical approaches complicate the synthesis of evidence and the development of standardized guidelines. While some studies propose technical solutions, such as reweighting algorithms or incorporating fairness metrics, these approaches often yield context-dependent results and may not fully address the underlying structural issues related to data representation. Consequently, there remains a need for comprehensive strategies that integrate technical, ethical, and sociocultural perspectives in the development and evaluation of AI systems.[9] The scarcity of research explicitly addressing Latin American populations in AI further underscores the existence of a critical knowledge gap. Preliminary analyses indicate that only a minimal number of studies incorporate Latin American data or explicitly examine the implications of AI within this regional context. This lack of representation is evident not only in primary research but also in systematic reviews, where Latin America is frequently absent from comparative analyses. The implications of this omission extend beyond regional relevance, as it limits the ability to develop globally applicable AI systems and undermines efforts to achieve equity in technological innovation.[10,11]

Ethical analyses centered on Latin America have additionally emphasized that the regional implications of AI extend beyond technical performance and involve broader sociocultural and structural considerations. Mancilla-Caceres and Estrada-Villalta argued that ethical concerns in Latin America are intensified by asymmetries between AI developers and local populations, limited access to technological education, and the insufficient representation of culturally diverse communities within global AI systems. The authors further highlighted that the underrepresentation of Latin American perspectives in AI governance and development may contribute to the persistence of exclusionary practices and reduce the capacity of AI systems to address region-specific social and healthcare realities.[12] In light of these considerations, the present study aims to critically examine the extent to which Latin American populations are represented in the scientific literature on AI, with a particular focus on gender bias, bioethical implications, and clinical decision algorithms. By conducting a structured review of recent publications, this work seeks to identify patterns of inclusion and exclusion in AI-related research, as well as to evaluate the alignment between current evidence and the ethical requirements for equitable AI deployment. Through this approach, the study intends to contribute to the ongoing discourse on algorithmic fairness and to highlight the necessity of incorporating diverse populations in the development and validation of AI systems.[3,11]

Material and methods

Search strategy and study design

A mini-review design was implemented to systematically identify and analyze recent literature addressing AI, bias, clinical applications, and population representation. The search was conducted exclusively using the Google Scholar database, selected due to its extensive indexing capacity and multidisciplinary coverage. Google Scholar was selected because of its broad multidisciplinary coverage and its ability to index peer-reviewed publications from diverse academic and regional sources, including literature that may be less visible in conventional biomedical databases. This characteristic was considered particularly relevant for identifying publications related to Latin American contexts and interdisciplinary discussions involving AI, ethics, and healthcare. The search process was performed on April 2026. The temporal filter was restricted to publications between 2024 and 2026 to ensure the inclusion of the most recent evidence. No language restrictions were applied. Additionally, only review-

type articles were considered eligible, including systematic reviews, scoping reviews, and narrative reviews, in order to capture synthesized evidence rather than primary experimental studies. The decision to focus exclusively on review articles was intended to evaluate how the existing body of synthesized literature conceptualizes and represents issues related to AI bias, regional underrepresentation, and clinical implications, rather than to assess the outcomes of individual experimental studies. The search query was constructed using the following predefined terms in English: "underrepresentation," "Latin American population," "artificial intelligence," "training," "bias," "clinical," and "decision algorithms." These terms were selected to align with the conceptual framework of the study, focusing on demographic representation and algorithmic bias in clinical AI systems. Boolean operators were applied implicitly within the Google Scholar search engine to maximize retrieval sensitivity. The criterion for identifying a "perfect match" was defined as the presence of all search terms within the title of the article.

Eligibility criteria and article selection

The initial search yielded a total of 2,820 records. From this dataset, a relevance-based selection approach was applied by considering the first 20 results ranked by Google Scholar's relevance algorithm. This approach was adopted to reflect real-world literature consultation practices and to prioritize highly cited and contextually aligned publications. This strategy was adopted to maintain a manageable and analytically consistent sample within the scope of a mini-review design, while reflecting the visibility and accessibility patterns commonly encountered during exploratory literature searches.

Articles were included if they met the following criteria: (1) classified as review articles, (2) published within the defined time frame (2024–2026), and (3) containing at least one of the predefined search terms in the title. Articles were excluded if they represented duplicate entries, non-review formats (e.g., original research, editorials), or if their thematic focus was unrelated to clinical, bioethical, or algorithmic aspects of AI (e.g., business decision-making applications). One duplicate article in Portuguese and one article focused on business decision-making were excluded during the screening process. Following the application of these criteria, a final sample of 18 review articles was included for analysis. The selection process was conducted manually, and titles were screened for the presence of predefined search terms. No automated screening tools were used.

Data extraction and classification

Data extraction was performed manually from the titles and abstracts of the selected articles. Although the primary classification strategy was based on the presence of predefined search terms in article titles, abstracts were also reviewed to verify thematic relevance and confirm whether each article addressed AI bias, clinical implications, demographic representation, or Latin American context. This approach allowed the analysis to preserve a consistent and reproducible title-based categorization while using abstracts as a complementary source for eligibility confirmation and contextual interpretation. Full-text analysis was not performed because the objective of the mini-review was to evaluate the visibility and explicit framing of the topic in the most accessible bibliographic elements, particularly titles and abstracts, rather than to conduct an exhaustive systematic review of article content. The primary variables extracted included: (1) presence of predefined search terms in the title, (2) combination of terms identified, (3) authorship and year of publication, and (4) thematic focus related to AI bias, clinical application, or population representation. Each article was classified according to the combination of search terms present in its title. Categories included, but were not limited to: "AI and bias," "AI only," "AI, bias, clinical, and algorithm," "algorithm and AI," "bias and algorithm," "Latin America," "AI and Latin America," and "bias, medical, AI, clinical, and decision." This categorical classification enabled the quantification of thematic representation across the selected literature.

Quantitative analysis

A descriptive quantitative analysis was conducted to determine the frequency and percentage distribution of articles according to the identified categories. The number of articles corresponding to each category was calculated, and relative percentages were derived based on the total sample size (n = 18). These results were subsequently organized into a summary table (Table 1), presenting the distribution of search term combinations alongside corresponding authors and publication years.

Table 1 Distribution of review articles (n = 18) according to search term combinations identified in article titles.

Number of articles	Percentage	Search term combinations identified in the title	Authors and year
6	33.33%	AI and Bias	Cau et al.[15]; Prados et al.[16]; Li et al.[14]; Ramchandi et al.. 2025; van Assen et al., 2024; Muralidharan et al. [18]
5	27.77%	AI	Bondok et al.[1] ; Amaya-Santos et al.[9] ; Osonuga et al.[7] ; Vasquez et al., 2025 ; Murphy et al.[13]
2	11.10%	Bias, AI, Clinical and Algorithm	Otokiti et al.[4] ; Goldstein et al., 2024
1	5.56%	Algorithm and AI	Haider et al.[17]
1	5.56%	Bias and Algorithm	Hussain et al.[3]
1	5.56%	Latin America	Alcocer et al., 2025
1	5.56%	AI and Latin America	Hecht-López et al.[11]
1	5.56%	Bias, Medical, AI, Clinical and Decision	Cross et al., 2024

Data validation and reproducibility considerations

To ensure consistency in data extraction, the search parameters, inclusion criteria, and classification framework were predefined prior to analysis. However, the reliance on a single database and a relevance-based selection strategy may introduce selection bias. The absence of multiple independent reviewers and the lack of automated screening tools were methodological constraints inherent to the mini-review design. Nevertheless, the standardized application of search criteria and categorical classification aimed to enhance internal consistency.

The full dataset of the 20 initially retrieved articles, including abstracts and bibliographic details, was compiled and preserved for reference and transparency. This dataset served as the primary source for both inclusion decisions and subsequent analytical categorization.

Results

Distribution of search term representation in selected articles

From the initial pool of 2,820 records retrieved through Google Scholar, a relevance-based selection yielded 20 articles, of which 18 met the predefined inclusion criteria and were retained for analysis. The excluded records consisted of one duplicated article in a different language version and one article with a thematic focus unrelated to clinical or bioethical aspects of AI. The final dataset (n = 18) was subjected to categorical classification based on the presence and combination of predefined search terms within article titles.

No articles (0%) were identified that simultaneously included all predefined search terms ("underrepresentation," "Latin American population," "artificial intelligence," "training," "bias," "clinical," and "decision algorithms") in their titles. This absence indicates a complete lack of direct alignment between the search criteria and the current review literature within the specified timeframe. The most frequently identified category corresponded to articles containing both "artificial intelligence" and "bias," comprising 6 out of 18 articles (33.33%). These studies included works by Cau et al.[15], Prados et al.[16], Li et al.[14], Ramchandani et al. (2025), van Assen et al, and Muralidharan et al.[18]. Titles within this category reflected a primary focus on algorithmic bias without explicit reference to geographic representation or Latin American populations. A second category included articles containing only the term "artificial intelligence," accounting for 5 out of 18 articles (27.77%). This group comprised studies by Bondok et al.[2], Amaya-Santos et al.[9], Osonuga et al.[7], Vasquez et al, and Murphy et al.[13] These articles addressed AI broadly, without explicit inclusion of bias, clinical decision-making, or regional representation in their titles.

A smaller subset of 2 articles (11.10%) included a combination of "bias," "artificial intelligence," "clinical," and "algorithm." These were identified as Otokiti et al.[4] and Goldstein et al. Titles in this category demonstrated a more integrative focus on clinical applications of AI and the presence of bias within algorithmic systems. Four additional categories were each represented by a single article (5.56% per category). One article Haider et al.[17] included the terms "algorithm" and "artificial intelligence," while another Hussain et al.[3] contained "bias" and "algorithm." One article Alcocer et al.[10] included only "Latin America," and one Hecht-López et al.[11] combined "artificial intelligence" with "Latin America." An additional article (Cross et al., 2024) included a broader combination of terms: "bias," "medical," "artificial intelligence," "clinical," and "decision," representing 5.56% of the sample. The distribution of these categories is summarized in Table I, which presents both absolute frequencies and percentage values for each combination of search terms, alongside corresponding authors and years of publication.

Absence of comprehensive thematic integration

Across the analyzed sample, no article demonstrated comprehensive thematic integration of all key variables defined in the search strategy. Specifically, no titles simultaneously addressed Latin American population representation, AI training processes, algorithmic bias, gender considerations, bioethical implications, and clinical decision-making. The absence of such integrative studies was consistent across all reviewed articles.

Limited representation of latin american context

Only 2 out of 18 articles (11.10%) included explicit reference to Latin America within their titles. One article focused exclusively on Latin America Alcocer et al.,[10] while another combined Latin America with AI Hecht-López et al.[11] No articles incorporated Latin America in conjunction with bias, clinical algorithms, or decision-making within the same title. This indicates a minimal presence of Latin American contextualization in the current review literature.

Prevalence of bias-centered studies without regional specification

A substantial proportion of articles (44.43%) included the term "bias" in combination with other variables such as artificial intelligence, clinical application, or algorithms. However, these studies did not specify geographic or regional populations in their titles. The focus remained primarily on generalized or unspecified populations, without explicit differentiation of underrepresented regions such as Latin America.

Heterogeneity in term combinations

The analyzed articles demonstrated considerable heterogeneity in the combinations of search terms present in their titles. While some studies focused narrowly on single variables (e.g., artificial intelligence alone), others incorporated multiple dimensions such as bias, clinical application, and algorithmic processes. However, this heterogeneity did not translate into comprehensive coverage of all relevant variables, as no single article achieved full alignment with the defined search criteria.

Summary of quantitative findings

The categorical distribution revealed a concentration of literature around AI and bias, with decreasing representation as additional variables were introduced. The inclusion of clinical and algorithmic terms was limited, and the incorporation of Latin American population context was minimal. The absence of articles meeting all predefined criteria highlights a gap in the current literature, as evidenced by the complete lack of titles encompassing the full spectrum of search terms defined in the study design.

Discussion

The present analysis suggests a limited integration of approaches addressing the underrepresentation of Latin American populations within AI training, particularly in relation to gender bias, bioethical considerations, and clinical decision algorithms. In the analyzed sample, no articles incorporated all predefined variables simultaneously, indicating that current review-based literature may remain fragmented and compartmentalized. This fragmentation suggests that, although individual components such as algorithmic bias or clinical applications are widely studied, their intersection with geographic underrepresentation, specifically within Latin America, has not yet been systematically addressed.[2,3] The novelty of these findings resides in the explicit identification of a structural disconnect between global AI research priorities and the inclusion of region-specific populations. While prior literature has extensively documented racial and ethnic disparities in AI systems, these analyses have predominantly focused on populations within high-income countries. The current results support the interpretation that Latin

America remains insufficiently represented in the conceptual framing of the analyzed review articles. This observation suggests that the issue may extend beyond data collection practices and may also involve the research agendas that guide AI development and evaluation.[3,13]

The predominance of studies addressing "artificial intelligence" and "bias" without geographic specification suggests a generalized treatment of bias that may obscure context-specific disparities. Algorithmic bias is often conceptualized as a universal phenomenon; however, the mechanisms through which bias manifests are inherently dependent on local demographic, socioeconomic, and healthcare system variables. The lack of explicit reference to Latin American populations implies that existing models and mitigation strategies may be derived from datasets that do not capture the heterogeneity of this region. Consequently, the transferability of such models to Latin American clinical settings should be considered with caution and may require additional regional validation.[2,3] The limited representation of Latin America within the analyzed literature further reflects disparities in scientific production and visibility. Research originating from or focused on Latin American contexts may be underrepresented in major databases, less frequently cited, or published in journals with lower international impact. This dynamic contributes to a feedback loop in which the absence of regional data reinforces its continued exclusion from global AI development. Moreover, the identification of only a single article explicitly combining artificial intelligence with Latin America underscores the marginal positioning of this region within the broader AI research landscape.[11]

From a bioethical perspective, these findings highlight critical concerns regarding justice and equity in the deployment of AI systems. The principle of justice requires that technological advancements benefit all populations equitably; however, the absence of representative data undermines this objective. AI models trained on non-representative datasets may yield systematically biased outputs, particularly in clinical decision-making contexts where accuracy and fairness are essential. The observed lack of integration between bioethical discourse and regional representation suggests that ethical frameworks may not be sufficiently operationalized in current research practices.[7] Gender bias, although frequently addressed in the literature, is rarely examined in conjunction with regional underrepresentation. The intersection between gender and geographic context remains largely unexplored, despite evidence that gender-related disparities in healthcare are influenced by cultural, economic, and institutional factors that vary across regions. The absence of studies incorporating both gender and Latin American variables indicates a limitation in the application of intersectional approaches within AI research. As a result, existing analyses may fail to capture the compounded effects of multiple forms of bias, leading to incomplete assessments of algorithmic fairness.[11]

The heterogeneity observed in the combination of search terms across articles reflects variability in research focus and methodological approaches. While some studies adopt a comprehensive perspective on bias within clinical AI systems, others remain narrowly focused on technical or domain-specific aspects. This variability complicates the synthesis of evidence and suggests a lack of standardized frameworks for evaluating bias and representation. The absence of articles integrating all key variables further indicates that current methodologies may not be designed to address multidimensional research questions.[14] Several explanations may account for the observed gaps. First, the reliance on existing datasets, which are often derived from high-income countries, limits the availability of data representing Latin American populations. Second, institutional and infrastructural barriers within Latin America may constrain the generation and dissemination of AI-related research. Third, editorial and publication biases may influence which studies are accepted and indexed in widely used databases. These factors collectively contribute to the systemic underrepresentation identified in the results.[14,15]

The methodological approach employed in this study also introduces certain limitations that must be acknowledged. The use of a single database (Google Scholar) and a relevance-based selection of the first 20 results may have restricted the diversity of the sample. Additionally, the focus on review articles excludes primary studies that may contain relevant data on Latin American populations. The manual screening process and the absence of multiple reviewers may also introduce subjective bias in article selection and classification. Furthermore, the reliance on title-based search term identification may underestimate the presence of relevant variables discussed within the full text of articles.[16] Despite these limitations, the findings provide a structured overview of current trends in the literature and identify a clear gap in the integration of key variables related to AI bias and population representation. The absence of comprehensive studies addressing Latin American underrepresentation within AI training and clinical decision-making algorithms underscores the need for a reorientation of research priorities. Future investigations should adopt multidimensional frameworks that incorporate geographic, gender, and ethical considerations simultaneously, as well as promote the inclusion of diverse populations in both data collection and model development processes.[17]

Within the Latin American context, Chile has initiated region-specific efforts to develop artificial intelligence models trained on locally representative datasets. This approach has contributed to the emergence of Latam-GPT, a model designed within Latin America to mitigate biases associated with systems trained predominantly on data from high-income countries with distinct demographic profiles. These initiatives aim to enhance contextual relevance and reduce disparities in algorithmic performance across populations. Furthermore, Chile's national supercomputing infrastructure has provided critical support for these developments, functioning as a technological platform for AI innovation with an explicit emphasis on ethical considerations and the reduction of regional technological asymmetries.[18,19,20]

Conclusion

The present study suggests that the current body of review-based literature on AI has limited comprehensive integration of key variables related to underrepresentation, particularly concerning Latin American populations.[2,3] The most relevant conclusion derived from this analysis is that Latin American populations remain minimally represented both in the conceptual framing and in the evaluative scope of AI research, which may contribute to the persistence of gender, bioethical, and clinical biases within decision-making algorithms. This underrepresentation is not only quantitative but also epistemological, as it reflects a lack of integration of regional perspectives into the development and assessment of AI systems.[11] nThe importance of these findings lies in their implications for the equitable deployment of AI in healthcare and related domains. The absence of representative data and context-specific analysis compromises the external validity and fairness of AI-driven decision-making systems. This limitation is particularly critical in clinical environments, where biased outputs may influence diagnostic accuracy, treatment recommendations, and overall patient outcomes. Furthermore, the insufficient incorporation of bioethical principles into AI research, particularly in relation to justice and inclusivity, raises concerns about the responsible implementation of these technologies across heterogeneous populations.[16]

Citation: Hecht-López P. Artificial intelligence training and population bias: underrepresentation of Latin America in gender, bioethical and clinical decision algorithms. *J His Arch & Anthropol Sci.* 2026;11(2):90–95. DOI: 10.15406/jhaas.2026.11.00344

Future perspectives should focus on the development of integrative research frameworks that simultaneously address demographic diversity, gender dimensions, and ethical considerations within artificial intelligence. There is a need to promote the inclusion of Latin American populations in training datasets and to encourage region-specific studies that evaluate algorithmic performance in diverse sociocultural and clinical contexts. This approach is exemplified by recent initiatives in Chile, including the development of the Latam-GPT model and the support provided by the National Laboratory for High-Performance Computing,[19,20] which collectively represent early efforts to generate locally trained AI systems aimed at reducing structural biases associated with non-representative data. Additionally, the establishment of standardized reporting guidelines for demographic variables and the expansion of collaborative research networks involving underrepresented regions are essential to improve transparency, reproducibility, and equity in AI development. Advancing these directions may contribute to the creation of more robust, inclusive, and ethically aligned AI systems capable of supporting equitable decision-making processes.[18]

Acknowledgments

None.

Conflicts of interest

None.

References

1. Cross JL, Choma MA, Onofrey JA. Bias in medical AI: implications for clinical decision-making. *PLoS Digit Health*. 2024;3(11):e0000651.
2. Bondok M, Selvakumar R, Asdo A, et al. Sociodemographic reporting in artificial intelligence studies of retinal diseases: a critical appraisal of the literature. *Ophthalmol Retina*. 2025;9(12):1133–1139.
3. Hussain SA, Bresnahan M, Zhuang J. The bias algorithm: how AI in healthcare exacerbates ethnic and racial disparities–a scoping review. *Ethn Health*. 2025;30(2):197–214.
4. Otokiti AU, Shih HJ, Williams KS. Gender and racial bias unveiled: clinical artificial intelligence (AI) and machine learning (ML) algorithms are fanning the flames of inequity. *Oxford Open Digit Health*. 2025;3:oqaf027.
5. Clavell GG. Gender Bias in AI: Risks and Opportunities for Latin America and the Caribbean. 2024.
6. Williams NH. Artificial Intelligence and Healthcare: The Impact of Algorithmic Bias on Health Disparities. *Springer Nature*. 2024;107.
7. Osonuga A, Osonuga AA, Fidelis SC, et al. Bridging the digital divide: artificial intelligence as a catalyst for health equity in primary care settings. Int J Med Inform. 2025:204:106051.
8. da Silva Jr JB, Birminghamm M, Cinnamond AR, et al. Bias-Mitigated AI as a foundation for resilient and effective health systems. *JMIR Public Health Surveill*. 2026;12(1):e88457.
9. Amaya-Santos S, Vargas C, Bermúdez-Tamayo C. Frameworks encompassing intersectional perspective of artificial intelligence in healthcare. *Public Health Pract*. 2025;11:100713.
10. Alcocer L, Schiffrin EL, Fink G, et al. Racial and Gender Discrimination When Tailoring Medical Management to Hypertension Treatment in Latin America. *Am J Hypertens*. 2025;38(10):739–747.
11. Hecht-López P, Parra-Villegas E. Artificial intelligence, bioethics, and gender in medical training: a critical review of the current landscape in Latin America. *J His Arch Anthropol Sci*. 2026;11(1):31–34.
12. Mancilla-Caceres JF, Estrada-Villalta S. The ethical considerations of AI in Latin America. *Digital Soc*. 2022;1(2):16.
13. Murphy A, Bowen K, Naqa IME, et al. Bridging health disparities in the data-driven world of artificial intelligence: a narrative review. *J Racial Ethn Health Disparities*. 2025;12(4):2367–2379.
14. Li H, Moon JT, Shankar V, et al. Health inequities, bias, and artificial intelligence. *Tech Vasc Interv Radiol*. 2024;27(3):100990.
15. Cau R, Pisu F, Suri JS, et al. Addressing hidden risks: Systematic review of artificial intelligence biases across racial and ethnic groups in cardiovascular diseases. *Eur J Radiol*. 2025;183:111867.
16. Prados JSF, Fernández YC, de Haro MJT. Racial/Ethnic Bias in AI Systems: A PRISMA Systematic Review of Magnitude, Domains, and Mitigation Strategies (2014–2025). *Rev Int Estud Migr*. 2025;15(1):57–81.
17. Haider SA, Borna S, Gomez-Cabello CA, et al. The algorithmic divide: a systematic review on AI-driven racial disparities in healthcare. *J Racial Ethn Health Disparities*. 2026;13(1):188–217.
18. Muralidharan V, Schamroth J, Youssef A, et al. Applied artificial intelligence for global child health: Addressing biases and barriers. *PLoS Digit Health*. 2024;3(8):e0000583.
19. Gobierno de Chile. Lanzamiento Latam-GPT: primera inteligencia artificial de Latinoamérica y el Caribe. 2026.
20. National Laboratory for High Performance Computing (NLHPC). National Laboratory for High Performance *Computing Chile*. 2026.