Perspective article

Virosa Journal of AI in Science and Healthcare

The Female Body Is Not Standardized: Why AI Must Catch Up

HerConversation Editorial Team [a]

[a] *Girl Power Talk, 11th Floor, ITC-11, Adjoining World Tech, Sector 67, Mohali Punjab, India.*

ABSTRACT

Artificial intelligence (AI) is increasingly positioned as a transformative force in modern healthcare. From diagnostic imaging and predictive analytics to personalized treatment recommendations, AI-driven systems promise efficiency, accuracy, and scalability. However, underneath this guise of technological progress, the problem of treating the male anatomy as standard in healthcare continues to pose a challenge.

This paper explores how AI in healthcare reinforces historical patterns of gender bias by excluding women from medical datasets and male-centered model design. It discusses the negative impacts of deploying male-standardized AI systems in women's health and calls for the development of women-centered AI. The author's recommendations include the use of diverse clinical data, creating algorithms suitable for female physiology, and setting up ethical oversight to achieve equity and accuracy in healthcare in the age of AI.

Literature review

Gender bias in healthcare AI cannot be understood without examining the history of biomedical research. Extensive scholarship has established that women have been consistently left out of clinical trials and large-scale medical studies over the past decades, creating the foundations of a biased system that lasted for centuries. One of the major reasons for the exclusion was the concern about the effects of hormonal fluctuations, pregnancy, etc., on research outcomes. This 'precaution' led to the exclusion of critical data that affects women and girls every day, leading to the subsequent knowledge bases being largely male-dominated (National Academies of Sciences, Engineering, and Medicine, 2010) [1].

Research demonstrates that the historical practice of excluding women from clinical trials has resulted in the development of medical datasets that favor male study subjects (National Academies of Sciences).

The research findings have established that existing diagnostic criteria and treatment standards have created a risk model that inadequately accounts for endometriosis, PCOS, autoimmune disorders, and cardiovascular disease. The AI systems that use these datasets have learned to replicate existing prejudiced patterns. Further studies confirm this bias. A UCL study on liver disease screening found AI missed diagnoses in women almost twice as often as men. A Nature Medicine review reported similar under diagnosis in imaging AI. Historical stereotypes about women's symptoms continue to influence clinical documentation tools, which under represent their medical symptoms. Most healthcare AI systems lack sex-disaggregated validation and monitoring according to algorithmic fairness research, which results in ongoing gender disparities and postponed medical treatment for female patients [2].

Even though the importance of gender bias in the field of healthcare AI is starting to be openly discussed, the existing research remains small-scale and limited in its scope.

The majority of the existing literature, including the works on liver disease screening and medical imaging is focused on documenting the difference in the diagnosis accuracy between genders. Although these findings hold relevance in establishing the fact that bias is a reality, it is also worth bearing in mind that much of the research conducted does not go beyond descriptive studies to provide systematic and implementable measures.

These gaps indicate the need for a shift from problem identification toward structural solutions. Addressing gender bias in healthcare AI requires rethinking data practices, algorithm design, validation standards, and governance frameworks through an explicitly women-centered lens: an approach this paper seeks to advance.

Introduction

Artificial intelligence (AI) is becoming progressively integrated in healthcare systems to enhance the quality, efficiency, and accessibility of medical care. AI-powered software is now applied in diagnostics, predictive modeling, clinical decision support, and individual treatment planning. Existing technologies are proposed as transformative to the field of healthcare, making it possible to detect the disease earlier, optimize treatment regimes, and assist clinicians during the complicated decision-making process.

Though promising, AI in healthcare faces a significant hurdle: because women had traditionally been excluded from clinical research, a significant gap in gender data was formed. The majority of medical datasets that are typically used to train AI models are male-oriented i.e., based on symptom patterns, disease progression, and responses to treatment that are more common in men. Consequently, AI-based systems tend to miss the differences in female physiology, hence leading to under- or misdiagnosis and treatment delays. Conditions and diseases where women are overrepresented, including autoimmune disorders, polycystic ovary syndrome (PCOS), endometriosis, and even cardiovascular disease, are often poorly represented or characterized according to the female autonomy in such datasets [3].

This healthcare AI gender bias is not an accidental technical mistake; it embodies the structural inequities in biomedical research and healthcare clinical practice. The diagnostic AI tools tend to miss classifying symptoms that women commonly experience adequately, such as fatigue, nausea, and general pain. AI-assisted clinical documentation systems show a tendency to downplay health issues that women report because they rely on historical stereotypes that depict women as less credible and more emotional about their health problems.

The very notion that female and male differences in biology are so minor that they can be overlooked or treated statistically is a reinforcement of the existing medical biases; only this time,

the biases are encoded into the systems of algorithmic decision-making. This study explores how AI in healthcare reinforces the historical patterns of gender bias by excluding women from medical datasets and male-centered model design.

The objectives of this study, in particular are:

- Discuss the history and structural foundation of gender bias in biomedical research and data.
- Explore the gender bias in the AI applications in the fields of diagnostics, predictive modeling, and clinical decision support.
- Critically analyze how AI models based on males affect clinical outcomes, patient safety, and confidence in healthcare systems.
- Suggest measures to build a fair and women-focused AI that can be implemented in data gathering, algorithm development, clinical testing, and appropriate ethical regulation.

Historical and Structural Roots of Gender Bias in Medical Data

a) Biomedical Research Bias

As the literature review established, for much of the twentieth century, male bodies have been treated as the default biological standard. This resultant (and continuous) bias continues to have repercussions for women's health today. The initial criteria for diagnosis, the standard for treatment, and the models for risk continue to be based on incomplete data that do not accurately reflect women's physiology. Gender-specific conditions such as endometriosis, polycystic ovary syndrome, and autoimmune disorders, etc. are still poorly researched, diagnosed, and not represented at all in the large datasets. Even in the case of diseases common across all sexes, women's symptoms and the progress of the disease in female bodies have been largely neglected. For instance, ADHD in girls frequently appears as inattentiveness, anxiety, or emotional dysregulation rather than hyperactivity, leading to delayed or missed diagnoses. In heart disease, women may experience nausea, fatigue, or breathlessness instead of chest pain. Similarly, autoimmune diseases such as lupus disproportionately affect women, yet vague symptoms such as pain and fatigue are often dismissed as stress [4].

These significant gaps in medical knowledge are caused by the unavailability of sex-disaggregated data on a large scale. AI systems that rely on historical datasets to learn clinical patterns are already widening these gaps.

b) How Biased Biomedical Data Feeds Into AI Systems

A study conducted in Argentina showed how biomedical AI models tend to give less accurate diagnoses for female users when gender isn't sufficiently accounted for in the dataset that trains the AI system. This inequity is more than just an error in computer programming; it is but an outcome of its historically skewed dataset structure.

AI systems do not generate knowledge on their own accord, but rather learn logical connections from the existing data. So, when the training datasets consist of historical exclusions and imbalances, the algorithms can hardly help but integrate these biases. Male-biased datasets train AI systems to regard male symptom profiles as standard and are unable to provide accurate results for female and other minority groups that are less represented in the research sets [5].

Consequently, the issue of gender bias in AI systems is not accidental; it is inherent. If this issue is not addressed, then AI systems will only continue perpetuating inequities and restricting access to quality healthcare and information for women.

Evidence of Gender Bias in Healthcare AI Applications

a) *Diagnostic Accuracy and Predictive Tools*
AI diagnostic and predictive tools may seem very precise, but they are often unable to capture the variations in the presentation of symptoms between genders and other physiologically diverse groups, resulting in misdiagnosis and delays in treatment. For instance, a UCL-led study on AI tools used for liver disease screening found that these systems were almost twice as likely to miss the disease in women as compared to men, despite showing high overall accuracy due to the imbalances in the data used.

Tools for prediction that are based on earlier research variables give more weight to the symptom profiles of male patients due to the larger evidence available. Symptoms that are more often present in women, such as fatigue, diffuse pain, dizziness, or nausea, are prone to being misunderstood or receiving lower clinical urgency. This scenario is particularly pronounced in cardiovascular care, where women's symptoms are regarded as odd in male-styled diagnostic methods and risk delaying diagnosis in women. A Nature Medicine review of AI medical imaging similarly found higher under diagnosis rates for women and other underrepresented groups when sex-disaggregated performance was examined.

Moreover, performance metrics can be misleading and might act as a blindfold, hiding the difference between the groups in terms of their outcomes. If an AI model shows high average accuracy, it may still deliver decreased effectiveness for those underrepresented in training. The gender bias, hence, makes its way into the systems that seem to be overtly neutral and efficient [6].

b) *AI Summaries and Clinical Decision Support Tools*
Healthcare bias is not just apparent in medical diagnosis AI; it is also found in systems applied in clinical documentation and decision support systems. One common example of this is automatic summarization systems that are increasingly being used in organizing medical information from patients.

Studies indicate that in practical applications, such methods tend to trivialize women's complaints regarding physical and mental health while summarizing medical records. Often, descriptions of pain, mental anguish, or chronic symptoms are abbreviated, trivialized, or relegated to a lower priority when the patient in question is a woman.

When the symptoms experienced by women are constructed as being of less severe proportions or urgency in records, the assessments made based upon them have the potential to reinforce historical beliefs related to women's experiences being driven by their emotions a notion that has long been used to discredit women's claims and to suggest that they experience pain differently than men. Although this may not seem discriminatory in individual cases, repeated minimization across consultations accumulates over time, contributing to delayed diagnoses, inadequate treatment, and systemic neglect.

Broader Algorithmic Fairness Issues

The issue of gender bias in healthcare AI is a reflection of the larger issue of algorithmic fairness. Most AI systems are developed and validated against evaluation frameworks that often prioritize average accuracy rather than equitable performance amongst different demographic groups. When gender differences are not paid attention to, the discrepancies in error rates will remain undetected and are amplified through deployment and user learning in a stratified society.

A comprehensive review found that most clinical AI tools lack sex or gender-specific validation, allowing performance gaps to persist undetected both before and after deployment. These gaps have the potential to become even more harmful when other factors, such as race, class, and geographic access, are added to gender-based discrimination [7].

This situation creates a paradox that is at the core of healthcare AI: the technologies that are meant to enhance efficiency, standardization, and access might end up exacerbating the existing inequalities. Instead of reducing the systemic bias, such systems amplify past exclusions into their automated decision-making processes.

Discussions

a) *Consequences of Male-Centric AI Models for Women's Health*
The effects of male-dominated models of AI in the healthcare of women are not only mistakes in individual diagnosis, but are also reflected in the continuum of health offered. If AI fails in identifying symptoms for women, women will likely be directed towards uncertain, delayed, or even incorrect diagnoses.

These tendencies increase their vulnerability to preventable complications. Inadequate evaluation of risk factors can also

lead to diagnosis only at the progressed stages of diseases. Further, inadequate preventive measures or appropriate Treatments can also result in additional health risks. It is essential to understand that these consequences are not individual issues, but rather a systematic failure to see the female anatomy and symptoms as 'standard.'

These errors in diagnosis will be exacerbated for women who find themselves at the nexus of gender, race, class, or geographic disadvantage. The system also restricts users from accessing second opinions, because AI-driven healthcare tools are often positioned as authoritative or final decision-makers. This, in turn, leads to unaddressed mistakes that result in postponed medical treatment and permanent health issues [8].

b) Health Equity and Trust in Healthcare Systems

When women keep getting unfairly labeled as overreacting, are pushed aside, or not attended to adequately, it chips away at their credibility. If systems always downplay women's health issues, it makes their concerns seem less real. Such issues with AI in healthcare show consequences in a patient's trust. When 'objective' AI healthcare systems appear unbiased and trustworthy, but results do not align with the realities that women experience, the distance between the patient and the healthcare system can widen. Women might not seek care in time, not share symptoms freely, or ignore preventive measures, especially in cases where their experiences with the systems were defined by rejection of their issues or delay in care.

On a societal scale, these patterns magnify gender inequalities in how healthcare is provided and planned. When public health strategies rely on biased AI results, they risk putting resources in the wrong places, underestimating how diseases affect women, and reaffirming unfair practices in preventive care. In this way, AI that excludes female-specific data doesn't just impact one-on-one encounters but also shapes what institutions prioritize and the direction of health policies.

We can't achieve fairness in healthcare if the technology guiding decisions keeps pushing aside women's bodies, experiences, and needs.

Recommendations: Designing Women-Centered AI in Healthcare

The review of currently existing healthcare AI solutions demonstrates that models usually use training datasets that are not diverse enough in terms of sex, age, ethnicity, and health status of women. Menstruation, pregnancy, menopause, and post-menopause are some of the most common life stages that are underrepresented, and therefore, they fail to capture the biologically important variation in the datasets. Consequently, AIs are more likely to follow male-biased norms instead of being trained on a wide range of manifestations of disease among women [9].

The development of AI systems that meet women's healthcare needs requires rigorous checks in the data collection stages and algorithm development processes. Reliability improves when AI systems incorporate methods that treat gender and life stages as distinct categories, and when these systems are validated across diverse populations.

a) Data Collection and Representation

Ensuring equitable representation in AI results and accuracy has to begin with quality data. Sex-disaggregated datasets with women from varying ages, ethnic groups, and health statuses are required to inform AI systems and learning. Data should be gathered with different stages of life in mind, such as menstruation, pregnancy, menopause, and post-menopause- biologically significant and unique aspects that women experience that male-focused or aggregated data will not account for.

The creation of standards of balanced training data will enable technology to learn from diverse presentations of disease rather than being trained from a biased norm based on males [10].

b) Fairness-Aware Algorithms and Evaluation Metrics

In developing a model, making use of fairness metrics is essential to spot disparities in performance for female and male users. Specificity, sensitivity, and error rates can be analyzed separately for women and men.

Equitable techniques for auditing an algorithm and bias reduction need to be integrated throughout the life cycle of an AI solution, right from design and learning stages through deployment and updates.

c) Clinical Validation and Monitoring

Rigorous validation of AI tools before deployment with female sub-groups needs to be prioritized. Gender-sensitive monitoring of AI tools after deployment can also help determine any unforeseen gender bias outcomes. Feedback loops allow developers, as well as practitioners, to improve systems and eliminate bias.

d) Ethical, Regulatory, and Policy Considerations

The ethical, regulatory, and policy considerations for the use of AI in the healthcare sector should shift from the current idealistic framework of automated objectivity to a framework with clear equity standards that embrace gender and the full range of biological and physiological diversity. Standardized metrics to test for biases and inaccuracies by demographic groups, along with third-party audits of systems, should be put in place to eliminate algorithmic discrimination [11].

In terms of governance, gender equity frameworks should be embedded in AI policies, guiding data gathering, development, and outcome analysis. Medical technologies need to treat female variation as an essential principle in research and technological advancement.

Effective regulation in this vein will require interdisciplinary collaboration. An interoperable system where clinicians inform the complexity of diagnosis and patient practice, ethicists analyze assumptions about justice in algorithmic systems, technologists map the translation of equity principles into technical solutions, and women's health practitioners contextualize non-standardized realities of women's bodies through their experiential knowledge's is required to ensure the benefits of technological progress are reaped equitably by all.

When these viewpoints come together, they can help make sure AI is used in ways that align the integration of new tools with the unique realities of women's bodies [12].

Conclusion

Healthcare AI must move beyond the illusion of neutrality and confront the long-standing assumption of the male body as the default medical standard. For innovation to promote equity, interdisciplinary interaction between data science and gender-informed clinical practice is essential to ensure that data systems mirror the diversity of life. Strengthening future research will require a re-evaluation of the standards developed for evaluating how AI tools function in and for female health.

Fair and unbiased AI in medicine requires active approaches to AI system development, data, and governance that center on gender diversity and equity. The female body is not standardized, and healthcare AI must evolve to reflect this reality.

References

1. Biddle C, et al. Gender differences in symptom misattribution for coronary heart disease symptoms and intentions to seek health care. *Women Health*. 2020;60(4):367-381.
2. Chinta SV, et al. AI-driven healthcare: fairness in AI healthcare: a survey. *PLOS Digit Health*. 2025;4(5):e0000864
3. Chinta SV, et al. AI-driven healthcare: fairness in AI healthcare: a survey. *PLOS Digit Health*. 2025;4(5):e0000864
4. Harrilal-Maharaj K. Gender bias and diagnostic delays in young women: a narrative review. *Cureus*. 2025;17(12):e100004.
5. Larrazabal AJ, Nieto N, Peterson V, Milone DH, Ferrante E. Gender imbalance in medical imaging datasets produces biased classifiers for computer-aided diagnosis. *Proc Natl Acad Sci U S A*. 2020;117(23):12592-12594.
6. Liu M, Ning Y, Teixayavong S, et al. A scoping review and evidence gap analysis of clinical AI fairness. *NPJ Digit Med*. 2025;8:360.
7. National Academies of Sciences, Engineering, and Medicine. *Women's health research: progress, pitfalls, and promise.* Washington (DC): National Academies Press; 2010.
8. Skogli EW, Teicher MH, Andersen PN, et al. ADHD in girls and boys - gender differences in co-existing symptoms and executive function measures. *BMC Psychiatry*. 2013;13:298.
9. Seyyed-Kalantari L, Zhang H, McDermott MBA, et al. Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations. *Nat Med*. 2021;27:2176-2182.
10. Straw I, Wu H. Investigating for bias in healthcare algorithms: a sex-stratified analysis of supervised machine learning models in liver disease prediction. *BMJ Health Care Inform*. 2022;29(1):e100457.
11. Gender bias revealed in AI tools screening for liver disease. *UCL News*. 2022 Jul 11. Available from: https://www.ucl.ac.uk/news/2022/jul/gender-bias-revealed-ai-tools-screening-liver-disease.
12. Zhang L, Reynolds Losin EA, Ashar YK, Koban L, Wager TD. Gender biases in estimation of others' pain. *Pain*. 2021;22(9):1048-1059.