PAN AMERICAN JOURNAL OF PUBLIC HEALTH

Special report

AI's ongoing impact: Implications of AI's effects on health equity for women's healthcare providers

Suman Vadlamani[1] and Elizabeth Wachira[2]

Suggested citation Vadlamani S, Wachira E. AI's ongoing impact: Implications of AI's effects on health equity for women's healthcare providers. Rev Panam Salud Publica. 2025;49:e19. https://doi.org/10.26633/RPSP.2025.19

ABSTRACT

Objective. To assess the effects of the current use of artificial intelligence (AI) in women's health on health equity, specifically in primary and secondary prevention efforts among women.

Methods. Two databases, Scopus and PubMed, were used to conduct this narrative review. The keywords included "artificial intelligence," "machine learning," "women's health," "screen," "risk factor," and "prevent," and papers were filtered only to include those about AI models that general practitioners may use.

Results. Of the 18 articles reviewed, 8 articles focused on risk factor modeling under primary prevention, and 10 articles focused on screening tools under secondary prevention. Gaps were found in the ability of AI models to train using large, diverse datasets that were reflective of the population it is intended for. Lack of these datasets was frequently identified as a limitation in the papers reviewed (n = 7).

Conclusions. Minority, low-income women have poor access to health care and are, therefore, not well represented in the datasets AI uses to train, which risks introducing bias in its output. To mitigate this, more datasets should be developed to validate AI models, and AI in women's health should expand to include conditions that affect men and women to provide a gendered lens on these conditions. Public health, medical, and technology entities need to collaborate to regulate the development and use of AI in health care at a standard that reduces bias.

Keywords Artificial intelligence; women's health; primary prevention; secondary prevention; ethics.

Artificial intelligence (AI), an increasingly prominent topic in every field, refers to the computer science discipline where computers can learn, adapt to new information, and carry out executive tasks, allowing computers to think, learn, and work like humans (1). In medicine, the application of AI has centered on disease pathophysiology research, risk factor characterization, and disease diagnosis and management (1). For example, for Pap smears AI can be used to classify pathologies across risk levels for cervical cancer, like a cytologist would. A survey published in 2024 found that 40% of physicians were ready to integrate AI technology into their clinical practice, and over 60% of physicians had become more amenable to the use of AI in health care (2). Although application varies across healthcare sectors, the basic functions of AI included in most models proposed include feature extraction models such as deep neural networks and convolutional neural networks that enable AI to identify key characteristics in the data being analyzed (3). The classifying models like support vector machines help categorize data across different risk levels and are used to determine next steps for patients based on their risk level for a disease (3).

In the clinical setting, clinicians can use AI to assess patient data and guide timely decisions. An AI model's strength depends on the data sources it is trained in, and if these are not representative of the intended population, the model may be inaccurate in its assessments to clinicians (1). Therefore, as AI use expands in health care, it is important to address the

[1] University of Texas Health Science Center at Houston, Houston, TX, United States of America ✉ Suman Vadlamani, suman.vadlamani@utsouthwestern.edu

[2] East Texas A&M University, Commerce, TX, United States of America

lack of data representation for underserved populations, who historically have higher rates of morbidity and mortality, as a potential limitation. These health disparities are shaped by the social determinants of health (SDOH), the conditions in which people live daily and which include aspects of economic stability, access to quality education, and access to quality health care (4). Women experiencing social and economic disadvantages are more likely to have poor health outcomes, which can improve with better disease prevention and management efforts (5). Therefore, integrating AI in primary and secondary prevention measures introduces more opportunities to improve health outcomes.

Primary prevention refers to the ability of the AI model to identify risk factors among populations for certain diseases. Secondary prevention refers to AI's ability to look for early signs of diseases that clinicians can intervene on early in the disease process. Concerns regarding health equity are brought up as AI is increasingly present in health care. More specifically, as SDOH impact health outcomes, AI technology can magnify these negative impacts due to the influence SDOH have on the training and implementation of AI technology in healthcare settings. Women are more likely to have poor health outcomes with unique healthcare needs for conditions that exclusively affect women, like cervical cancer, and conditions that affect women differently from men, like heart disease. This article is written as a narrative review, as the field of AI in health care is an emerging topic and presents an intersectionality across multiple field domains of technology, business, public health, and medicine. Providing a perspective on this topic as a narrative review is valuable to the field, as it can integrate the multiple perspectives of the topic (6). While there has been commentary published on the implications of AI on health equity, the purpose of this review is to assess effects of the current use of AI in women's health on health equity, specifically as it is used in primary and secondary prevention efforts among women.

METHODS

This narrative review was conducted using two databases: Scopus and PubMed. Keywords for the search included "artificial intelligence," "machine learning," "women's health," "screen," "risk factor," and "prevent." The search was restricted to articles published in English between 2014 and 2024, to ensure representation of recent advancements of AI integration in women's health prevention methods. There were no exclusions for papers published outside of the United States of America. Results included 71 articles from Scopus (n = 23) and PubMed (n = 48) with 9 duplicate abstracts eliminated. The remaining 62 abstracts were reviewed to include only articles where AI was used in primary and secondary prevention methods as a clinical decision-making tool for conditions regularly screened for among women, such as cervical cancer, breast cancer, osteoporosis, fall risk, postpartum depression, and intimate partner violence. One paper about cardiovascular disease (CVD) risk factor modeling was included, as CVD is a significant cause of morbidity and mortality for women and it is general practice to check for risk factors for CVD for all patients. Papers that described diseases not regularly screened for, like Alzheimer disease, or AI models that were used to investigate associations between different clinical factors were excluded. This review focused on AI models that can be used by clinicians in their general practice, as these would be the models that would affect the most patients – virtually the entire population of women. Lastly, review articles, editorials, and articles where full text was not provided were also eliminated.

RESULTS

Thematic trends

A total of 18 articles met the search criteria and were included in this review. Areas of focus include cancers (n = 8), osteoporosis (n = 2), mental health (n = 2), antepartum care (n = 3), and general issues including falls (n = 1), cardiovascular disease (n = 1), and intimate partner violence (n = 1). Risk factor modeling and AI as a screening tool were the two major themes noted where AI can be used in primary and secondary prevention, respectively.

Risk factor modeling refers to how AI can determine a patient's risk for developing a specific health condition using relevant medical history, physical exam findings, and imaging information. The United States Preventive Services Task Force (USPSTF) is the organization that sets standardized recommendations for screening tests that clinicians should perform for their patients (7). Since screening recommendations are made for their target populations, risk factor modeling may inform clinicians to screen for certain conditions earlier and more often if a patient is at a higher risk for developing the condition. Screening tools are important in ensuring early disease diagnosis and treatment, which significantly improves morbidity and mortality outcomes. AI can consolidate information obtained from a screening test and assist clinicians' ability to interpret results easily. Table 1 provides an overview of the AI models used and what area of women's health the AI model targeted. Table 2 outlines the various metrics used to describe their findings. The studies were inconsistent in which AI metrics they used to assess their AI models, as seen in Table 1.

Role of AI in risk factor assessment

As outlined in Table 1, about half of the papers from the literature review looked at the use of AI in risk factor modeling for conditions that have significant effects on women's health. These models were developed for conditions that USPSTF recommends regular screening for, including osteoporosis, breast cancer, intimate partner violence, postpartum depression, and falls (8–14). While not covered by USPSTF guidelines, other screenings in women's health considered as good practice are risk factors for fetal health anomalies and severe maternal morbidity among pregnant patients and cardiovascular health (15–18).

AI application for these screenings can improve healthcare providers' ability to provide individualized patient care. For example, in cases where a patient has a higher risk for developing osteoporosis, screening at an earlier age is beneficial. In the paper of Zeitlin et al., the average age of patients who self-reported a diagnosis of osteoporosis was 46.9 years of age, both younger than the average age of menopause, which is 52, and the recommended age to begin screening for osteoporosis, at 65 years old (13). These are cases where early risk factor characterization and appropriate intervention initiation can improve the morbidity of these conditions.

TABLE 1. Summary results of the literature review, 18 articles

Article title	First author	Year published	Prevention type	Type of AI used	Health service area	Performance measures	Limitations noted
Fetal health status prediction based on maternal clinical history using machine learning techniques	Akbulut, A.	2018	Primary	Decision forest model (best performing) compared to ultrasonography report	Antepartum fetal status	Accuracy, F-1 score, area under the curve (AUC)	Trained on a small dataset of 96 patients
Learning to identify severe maternal morbidity from electronic health records	Gao, C.	2019	Primary	Ridge logistic regression compared to CDC criteria	Antepartum care	AUC	Developers preselected candidate features and fed the machine learning models the features based on frequency in the population (instead of having a machine learning model perform the task) and only looked at diagnosis and procedural features
Ensemble learning-based feature engineering to analyze maternal health during pregnancy and health risk prediction	Raza, A.	2022	Primary	DT-BiLTCN made up of decision trees, bi-directional long short-term memory (BiLSTM), and temporal convolutional network (TCN) compared to other current AI models	Antepartum care	Precision, recall, accuracy, F-1 score, AUC	None listed by authors
Medical breast ultrasound image segmentation by machine learning	Xu, Y.	2018	Secondary	Convolutional neural networks compared to manual evaluation of images by experienced radiologists	Breast cancer	Accuracy, precision, recall, F-1 score	Noted that it can be time-consuming
A benign and malignant breast tumor classification method via efficiently combining texture and morphological features on ultrasound images	Wei, M.	2020	Secondary	Texture and morphological feature extraction and support vector machine (texture) and negative Bayes classifier (morphological) combined compared to other proposed AI models	Breast cancer	Accuracy, sensitivity, specificity	None listed by authors
Artificial intelligence medical ultrasound equipment: Application of breast lesions detection	Zhang, X.	2020	Secondary	Lightweight neural network compared to high precision neural network	Breast cancer	Sensitivity, specificity, average precision	Lightweight neural network did not perform as well as the high-precision neural network
External validation of a mammography-derived ai-based risk model in a U.S. breast cancer screening cohort of White and Black women	Gastounioti, A.	2022	Secondary	ProFound AI Risk compared to European results of where the AI model was originally developed and Gail 5-year risk score	Breast cancer	Age-adjusted area under the receiver operating characteristic curve (AUC)	Validation was only performed with data from one site in the United States
A clinical risk model for personalized screening and prevention of breast cancer	Eriksson, M.	2023	Secondary	Novel lifestyle-imaging AI model compared to an imaging-only AI model and Tyrer–Cuzick v8 model	Breast cancer	Positive predictive value, AUC	Missing data availability in the dataset, and the dataset used was from a Swedish population
Efficient breast cancer diagnosis from complex mammographic images using deep convolutional neural network	Rahman, H.	2023	Secondary	ResNet-50 convolutional neural network (CNN) compared to the reported data of the dataset and other AI models	Breast cancer	Confusion matrix, accuracy, sensitivity, specificity, error rate, F-1 score, AUC	None listed by authors
Dual-path network with synergistic grouping loss and evidence driven risk stratification for whole slide cervical image analysis	Lin, H.	2021	Secondary	Deep CNN with dual-path (DP) encoder and rule-based risk stratification compared to reported data of the dataset	Cervical cancer	Sensitivity, specificity	High computational costs

(Continued)

TABLE 1. (Cont.)

Article title	First author	Year published	Prevention type	Type of AI used	Health service area	Performance measures	Limitations noted
Differentiating single cervical cells by mitochondrial fluorescence imaging and deep learning-based label-free light scattering with multi-modal static cytometry	Liu, S.	2022	Secondary	CNN compared to other AI models	Cervical cancer	Accuracy, sensitivity, specificity, precision	None listed by authors
Accelerometer-based predictive models of fall risk in older women: a pilot study	Hua, A.	2018	Primary	Random Forest compared to results of short physical performance battery and self-report of falls	Falls	Accuracy, precision, AUC, recall, F-1 score	Small sample size, gait characteristics may be affected by observer bias, data from past falls vs. prospective fall data
Predicting incident heart failure in women with machine learning: The women's health initiative cohort	Tison, G.H.	2022	Primary	Least absolute shrinkage and selection operator (LASSO), classification and regression trees (CART) compared to known reported heart failure diagnosis	Heart failure	C-statistic (derived from receiver under curve analysis)	Variable availability to apply to other populations
Help seeking behavior by women experiencing intimate partner violence in India: a machine learning approach to identifying risk factors	Dehingia, N.	2022	Primary	LASSO and L2 regularized models compared to self-reports from the women	Intimate partner violence	AUC, balanced error rate	Recall bias due to self-reporting
A clinical prediction model for 10-year risk of self-reported osteoporosis diagnosis in pre-and perimenopausal women	Zeitlin, J.	2023	Secondary	Generalized additive model with pair-wise interactions (GA2M) compared to self-report of a diagnosis of osteoporosis	Osteoporosis	Area under the receiver operating characteristic curve, Brier score, likelihood ratio, sensitivity, specificity	The model does not consider time to diagnosis; the model is dependent on self-reports of diagnosis; the model was dependent on a 10-year follow-up, so lack of consideration of those without a 10-year follow-up; the model also does not include femoral neck BMD, which is included in the FRAX model, limited data availability on many patients in the dataset
Enhanced osteoporotic fracture prediction in postmenopausal women using Bayesian optimization of machine learning models with genetic risk score	Wu, Q.	2024	Primary	GRS-integrated XGBoost with Bayesian optimization (best performing compared to known reported history of osteoporotic fracture)	Osteoporosis	Accuracy, weighted F-1 score, area under the precision-recall curve (PRAUC), area under the receiver operating characteristic curve (AUC) for binary fracture predictions, C-index, Brier score, dynamic mean AUC over a 10-year follow-up period for fracture risk predictions	Lack of independent data validation and additional datasets
Prevalence and risk factor analysis of postpartum depression at early stage of using hybrid deep learning model	Lilhore, U.K.	2024	Primary	CNN and hybrid framework with bi-directional long short-term memory (Bi-LSTM) with transfer learning (TL compared to existing AI models)	Postpartum depression	Precision, recall, accuracy, F-1 score	Data privacy ethics
Social media mining for postpartum depression prediction	Trifan, A.	2020	Secondary	Not specified; AI models compared had different estimators and configurations	Postpartum depression	Precision, recall, accuracy, F1 score	None listed by authors

Source: Prepared by the authors based on the study data.

TABLE 2. Metrics used to describe AI models

Metric	Definition
Accuracy	Proportion of the AI model's identification of those with (true positives) and those without (true negatives) the disease out of all the predictions
Precision [a]	Proportion of true positives to all predicted positives the AI model identifies
Recall	Proportion of positive instances that were correctly predicted to be positive
F-1 score	Derived from the precision and recall measures and calculated as the harmonic mean between the two: F1-score = 2 x (precision x recall) / (precision + recall)
Area under the curve [b]	Used to determine how well an AI model can distinguish between positive and negative instances
Area under the precision-recall curve	Used to determine how well an AI model can distinguish between positive and negative instances for imbalanced sets where the number of negative instances is very different from the number of positive instances in the data set that is used to train the AI mode
Brier score [c]	Mean squared error between what is predicted and what the actual outcomes are
C incidence	Measures the concordance between what was predicted and what the actual outcomes were

Notes:
[a] Precision and recall differ where precision is used to minimize false positives and recall is used to minimize false negatives.
[b] Same value as a C-statistic.
[c] The lower the score, the better calibrated the model is.
Source: Prepared by the authors.

Role of AI in screening methodologies

Only four papers addressed official USPSTF-recommended screening processes, likely due to the novelty of AI technology and its inability to go through the validation needed to lend itself as an alternative to a gold standard test available for screening (10, 19–21). As seen in Table 1, the literature review highlighted that AI's use in screening practices is limited to cervical cancer screening and breast cancer screening (19–25). The proposed models for cervical cancer screening identify and classify high-risk specimens that would warrant further diagnostic testing. A model from the paper of Liu et al. demonstrates the ability of the AI technology to distinguish transformed cells from malignant cells to catch cancers earlier than normal (20). These models, however, require a specific type of imaging technique in addition to the computer capabilities of supporting the AI technology.

The clinical significance of the contributions of an AI model's input is an important consideration for whether it is worth implementing. In their paper on breast cancer screening, Eriksson et al. demonstrated the ability of their AI model to integrate imaging and lifestyle factors that would better personalize breast cancer screening compared to an imaging-only model and the Tyrer–Cuzick risk calculator which is currently used (22). Both models performed better than the Tyrer–Cuzick risk calculator. The lifestyle-imaging model had a higher proportion of patients identified as high risk at 22%, which is 3% higher than the imaging-only model's. There is a question of whether that difference is clinically significant enough to justify the cost of the implementation of one model versus the other.

AI model performance

All AI models demonstrated a high level of accuracy in whichever metrics were reported. For example, Hua et al.'s model for fall risk factor assessment showed that their best model had an accuracy of 79.3%, precision of 84.6%, sensitivity of 88.1%, and area under the curve (AUC) of 0.834 compared to their control, which had an accuracy of 69%, precision of 75%, sensitivity 0.873, and AUC of 0.545 (8). AUC, as defined in Table 2, refers to the ability of an AI model to correctly distinguish between a positive and negative instance. From here, the ability of the model to be implemented in medicine will be assessed. Most of the papers from the literature review did not include information about the cost of the infrastructure needed to support the AI model, which is important for accessibility. AI's existence in primary and secondary prevention practices shows that it can benefit clinicians in regular practice.

DISCUSSION

The purpose of this review was to assess AI usage in primary and secondary prevention efforts in women's health. Using AI in risk factor modeling can increase risk factor awareness and inform health promotion initiatives, including universal screening recommendations by USPSTF. One challenge to its impact is the lack of women's representation in clinical trials due to recruitment and retention barriers in medical research (26). Review findings show that most of the AI models developed for women's health applications are relevant to risk factor modeling for conditions unique to women, such as osteoporosis and breast and cervical cancers.

Conditions that affect both sexes should also be included, such as CVD – the leading cause of death for women – diabetes, and chronic respiratory diseases like chronic obstructive pulmonary disease and asthma (27). Although the prevalence of CVD among women is significant at 44%, there still is a lack of AI utilization in this space, as only one paper in our review discussed female-specific risk factors (27). Tison et al. looked at female-specific CVD risk factors, like obstetric history and exposure to hormone therapy, that can shape heart health outcomes (17). Therefore, as AI developers consider technology for conditions that affect both sexes, input from clinicians can inform what additional risk factors should be integrated into the risk assessment models. AI models also can identify risk factors that may not have been well described for certain conditions.

Challenges minority women face in health care – and how AI is affected

Women from ethnic minorities, compared to White women, are less likely to utilize health care due to socioeconomic barriers linked to insurance coverage and healthcare availability (28). As such they have poor access to quality health care and resources, resulting, for example, in higher rates of maternal mortality and CVD (29, 30). Despite higher rates of morbidity and mortality, Black and Hispanic women are not well represented in CVD studies due to distrust in the medical system and systemic racism, and language barriers (29, 31). These factors or disadvantages ultimately shape their access to quality and continuous care, resulting in poor representation in health research and data models.

One of the major critiques of AI usage in medicine is the potential for bias and, subsequently, its ability to magnify healthcare disparities. This concern is linked to the issue on the availability and representativeness of data used in AI models, which can impact the ability of AI integration in primary or secondary prevention measures. Therefore, AI developers should target underrepresented minority women in data collection or curation. This can be accomplished by validating AI technology against external datasets to provide a gendered racial lens and ensuring that data from minority-serving healthcare settings, county hospitals, free clinics, and rural clinics are included (32).

AI's role in reducing healthcare barriers

Access to quality care is limited by barriers associated with transportation or appointment scheduling. AI's ability to be integrated into mobile phone applications reduces these barriers, especially given the high likelihood that many patients have mobile phones that support this technology (8). Providers and patients can virtually coordinate the type and frequency of follow-up appointments needed based on specific risk factor characterization, although specific considerations for low-income women must be assessed.

AI can also be used as an alternative to currently recommended screening practices in low-resource settings. For breast cancer screening, it is recommended that patients undergo mammography with follow-up MRI for suspicious lesions noted on the screening mammographic images. The availability and cost of an MRI scan makes this imaging process expensive and difficult, yet AI technology provides a cheaper and accessible alternative method by evaluating breast ultrasound images and stratifying them into risk categories for breast cancer (23–25). Although not part of the standard algorithm for breast cancer screening, whole breast ultrasounds have been suggested to be viable as a supplemental tool for mammograms, especially for women with denser breast tissue, a characteristic that is known to be of higher prevalence among Black women (33). These ultrasound-based models have shown to be accurate, with appropriate level of sensitivity and specificity for a screening methodology. Also, since most clinics have ultrasound machines, integrating this AI technology makes a good alternative for breast cancer screening, as proposed by Zhang et al. (24). Therefore, as AI use increases in these settings, its accuracy for these specific populations would be important, as that reflects the databases on which it is trained.

Barriers to AI implementation in health care

AI offers numerous benefits in health care, yet barriers to full integration exist, ranging from cost to infrastructure needed for effectiveness. The inclusion of AI models in clinical practice should outweigh the implementation costs, which include money, time, and staffing (34). The potential for AI to lead to healthcare cost savings have been highlighted by both academic and commercial fields; however, these savings are not seen until long-term implementation of the technology (35). An important barrier to the implementation of AI in health care is the initial integration of the technology, as noted by Forbes (36). Although they note that open access models are a feasible solution to this barrier, consideration of health data privacy management with open access models complicates this solution.

Despite the cost-effectiveness of AI models, the disruption to the original workflow process is an additional barrier, as training would be required for clinicians and their support staff. Therefore, part of the AI models' validation should include simulation or beta-testing modes for different healthcare settings to evaluate AI model impact on the existing flow. Lastly, the need for expensive infrastructure to implement and maintain AI models can introduce difficulties in AI's ability to be integrated in a healthcare setting. Nonetheless, some AI technology can be integrated into simpler interfaces with clinicians and patients to reduce some of the barriers to its implementation. Therefore, AI developers should also be transparent about the upfront infrastructure and costs needed to run their program.

Dataset representation and availability

Health care's history of introducing racial biases as correctional factors for medical calculations has led to poorer health outcomes for minority women. AI must be well trained to not allow the repeat of these mistakes in health care. More databases are needed to effectively train AI models, and there must be consistency across patient data recorded. Data availability of patient-specific variables and datasets composed of a large, diverse population were commonly cited limitations of the articles included in this review. Many of the larger datasets described were from other countries, and other United States-based datasets were small, failing to include important groups such as a Hispanic category, which represents a significant proportion of the United States population (37). The most-used large dataset is from the Women's Health Initiative, which started recruitment in the 1990s, reflecting a population different from ours today (38). As databases of patients are developed by healthcare institutions across the United States, more data will be available for various regions of the country, settings (urban versus rural), and demographics, including sets specific to social factors.

Logistical considerations regarding data collection and privacy need to be considered in developing these datasets. Generation of systems to develop datasets, data-sharing policies between healthcare systems and AI companies, protection of private health information, and consent of individuals whose data will be used in the dataset are areas of potential in the healthcare technology industry. This will be reliant on partnerships between technology companies developing the AI models and healthcare institutions planning to use the models. Patients' trust in their healthcare providers heavily impacts the movement toward developing these datasets. The history of medicine taking advantage of minority groups, specifically Black patients like Henrietta Lacks, makes it more important that this trust is honed across healthcare settings in a unified front (39).

Challenges with AI integration among clinicians

As AI continues to build its presence in health care, it is essential to consider the ability of the user to fully understand the technology. Multiple health organizations including the World Health Organization have pushed for better transparency in the discussion of new AI technology in health care (1). A key limitation of AI-related literature is the use of heavy technical jargon, which hinders the ability to effectively communicate the methodologies and results to the public and intended users. Public health professionals can continue to advocate for policies that standardize how AI technology is presented to healthcare professionals to limit technical jargon and standardize which AI metrics are used to report the effectiveness of different AI models, to allow comparison between models. Including education about the technicalities of AI technology in continuing medical education curricula would also help address this challenge.

As AI becomes more heavily integrated into healthcare practices, a new field of research emerges that can investigate whether the implementation of AI on health outcomes will lead to improved health outcomes as disease processes may be caught earlier on. This would be a useful regulatory service to ensure that AI is contributing to the betterment of health care overall as the technology evolves and updates.

AI in health care also represents a crossover between the technology and healthcare industries in novel way. Given this, it is important to consider that the healthcare industry may not be the only interface that exists between healthcare-related AI technology and patients, and literature on the subject may not be available in medical journals. This makes it important for healthcare providers to remain up to date on the state of AI in health care, and to continue to call for the reduction of bias and encouragement of transparency in AI technology.

Study limitations

There is a possibility that this review did not assess all the articles related to AI in women's health. This could be because of the search terms used. Terms like "women's health" may not have been applied to conditions like CVD that affect both sexes. Selection bias may have occurred, as medical databases were exclusively used although this topic intersects technology and medicine. Technology-based databases may have also had articles addressing AI technology use in women's health. Despite these limitations, this study is important in showcasing the usefulness of AI integration to inform women's health prevention efforts. These limitations can be reduced with increased efforts toward interdisciplinary collaboration across public health, medicine, and technology.

Conclusion

AI holds significant promise in health care, particularly in identifying risk factors and supporting disease prevention. As comorbidities rise, especially among minority women, integrating AI into clinical workflows can enhance early diagnosis and treatment. AI can streamline decision-making, tailoring screening and follow-up care. However, challenges like bias and data representation must be addressed through model validation and transparent development processes. Facilitating interdisciplinary collaboration can ensure clinicians or those without an AI technology background can understand and thus inform development and implementation processes. In addition, efforts to incorporate the basics of AI technology in medical training and continuing education efforts can ensure clinicians have a basic understanding of AI and the role they can play to mitigate issues related to equity.

Recommendations

Public health experts should be aware of the presence of AI in women's health and the increased risk for bias to be present in this space due to health disparities. To reduce bias, public health officials should continue to work toward collaborating with AI technology developers and healthcare providers to develop comprehensive patient datasets that include patients who are low users of health care. There is also a need for public health experts to develop standardized methods for reporting AI effectiveness so that models can be compared to each other. Part of this standardization should include standards for the way AI development is presented to promote transparency so that potential areas of bias are evaluated for and addressed. It is also recommended that AI technology specific to women's health branch out to health conditions that affect both men and women that require a gendered lens. This is important, as these conditions, including cardiovascular disease, pose a threat to population health, and AI technology has the potential to improve the understanding and treatment of the way these diseases affect women specifically.

Author contributions. SV contributed to data curation, analysis, interpretation, and writing of the drafts of the manuscript. EW contributed to study conception and design, data interpretation, review of drafts of the manuscript, and writing of the final draft. Both authors reviewed and approved the final version of the manuscript.

Acknowledgments. The authors would like to acknowledge the support of Dr. Bijal Balasubramanian, M.B.B.S., Ph.D. of the Department of Epidemiology, UTHealth School of Public Health for her support and guidance of this work.

Conflict of interest. None declared.

Disclaimer. The opinions expressed in this manuscript are solely the authors' responsibility and do not necessarily reflect the views or policies of the *RPSP/PAJPH* or the Pan American Health Organization (PAHO).

REFERENCES

1. World Health Organization. Ethics and governance of artificial intelligence for health: WHO guidance. Geneva: WHO; 2021. Available from: https://iris.who.int//handle/10665/341996.
2. Wolters Kluwer. Over two-thirds of U.S. physicians have changed their mind, now viewing GenAI as beneficial in healthcare. [Alphen aan den Rijn]: Wolters Kluwer; 2024 Apr 16 [cited 2024 Nov 6]. Available from: https://www.wolterskluwer.com/en/news/gen-ai-clincian-survey-press-release.
3. IBM. What is machine learning (ML)?. [Armonk, NY]: IBM; [2024?] [cited 2024 Sept 23]. Available from: https://www.ibm.com/topics/machine-learning.

4. U.S. Centers for Disease Control and Prevention. Social determinants of health (SDOH) Atlanta, GA: CDC; 2024 [cited 2024 Sept 23]. Available from: https://www.cdc.gov/about/priorities/why-is-addressing-sdoh-important.html.
5. American Psychological Association. Fact sheet: Women & socioeconomic status. Washington, D.C.: APA; 2010 [cited 2024 Sept 15]. Available from: https://www.apa.org/pi/ses/resources/publications/women.
6. Sukhera J. Narrative reviews: flexible, rigorous, and practical. J Grad Med Educ. 2022;14(4):414–417. https://doi.org/10.4300/JGME-D-22-00480.1.
7. United States Preventive Services Task Force. Home page. Rockville, MD: USPSTF; 2017 [cited 2024 Sept 15]. Available from: https://www.uspreventiveservicestaskforce.org/uspstf/.
8. Hua A, Quicksall Z, Di C, Motl R, LaCroix AZ, Schatz B, et al. Accelerometer-based predictive models of fall risk in older women: a pilot study. npj Digital Med. 2018;1:25. https://doi.org/10.1038/s41746-018-0033-5.
9. Dehingia N, Dey AK, McDougal L, McAuley J, Singh A, Raj A. Help seeking behavior by women experiencing intimate partner violence in India: A machine learning approach to identifying risk factors. PLoS ONE. 2022;17(2):e0262538. https://doi.org/10.1371/journal.pone.0262538.
10. Gastounioti A, Eriksson M, Cohen EA, Mankowski W, Pantalone L, Ehsan S, et al. External validation of a mammography-derived AI-based risk model in a U.S. breast cancer screening cohort of white and black women. Cancers. 2022;14(19):4803. https://doi.org/10.3390/cancers14194803.
11. Lilhore UK, Dalal S, Varshney N, Sharma YK, Rao KBVB, Rao VVRM, et al. Prevalence and risk factors analysis of postpartum depression at early stage using hybrid deep learning model. Sci Rep. 2024;14(1):4533. https://doi.org/10.1038/s41598-024-54927-8.
12. Wu Q, Dai J. Enhanced osteoporotic fracture prediction in postmenopausal women using Bayesian optimization of machine learning models with genetic risk score. J Bone Miner Res. 2024;39(4):462–472. https://doi.org/10.1002/jbmr.4663.
13. Zeitlin J, Parides MK, Lane JM, Russell LA, Kunze KN. A clinical prediction model for 10-year risk of self-reported osteoporosis diagnosis in pre- and perimenopausal women. Arch Osteoporos. 2023;18(1):78. https://doi.org/10.1007/s11657-023-01292-0.
14. Trifan A, Semeraro D, Drake J, Bukowski R, Oliveira JL. Social media mining for postpartum depression prediction. Stud Health Technol Inform. 2020;270:1391–1392. https://doi.org/10.3233/SHTI200457.
15. Akbulut A, Ertugrul E, Topcu V. Fetal health status prediction based on maternal clinical history using machine learning techniques. Comput Methods Programs Biomed. 2018;163:87–100. https://doi.org/10.1016/j.cmpb.2018.06.010.
16. Gao C, Osmundson S, Yan X, Edwards DV, Malin BA, Chen Y. Learning to identify severe maternal morbidity from electronic health records. Stud Health Technol Inform. 2019;264:143–147. https://doi.org/10.3233/SHTI190200.
17. Tison GH, Avram R, Nah G, Klein L, Howard BV, Allison MA, et al. Predicting incident heart failure in women with Machine Learning: The Women's Health Initiative cohort. Can J Cardiol. 2021;37(11):1708–1714. https://doi.org/10.1016/j.cjca.2021.08.006.
18. Raza A, Siddiqui HUR, Munir K, Almutairi M, Rustam F, Ashraf I. Ensemble learning-based feature engineering to analyze maternal health during pregnancy and health risk prediction. PLoS ONE. 2022;17(11):e0276525. https://doi.org/10.1371/journal.pone.0276525.
19. Lin H, Chen H, Wang X, Wang Q, Wang L, Heng P-A. Dual-path network with synergistic grouping loss and evidence driven risk stratification for whole slide cervical image computation. Med Image Anal. 2021;69:101955. https://doi.org/10.1016/j.media.2021.101955.
20. Liu S, Chu R, Xie J, Song K, Su X. Differentiating single cervical cells by mitochondrial fluorescence imaging and deep learning-based label-free light scattering with multi-modal static cytometry. Cytometry A. 2022;103(3):240–250. https://doi.org/10.1002/cyto.a.24684.
21. Rahman H, Naik Bukht TF, Ahmad R, Almadhor A, Javed AR. Efficient breast cancer diagnosis from complex mammographic images using deep convolutional neural network. Comput Intell Neurosci. 2023;2023:7717712. https://doi.org/10.1155/2023/7717712.
22. Eriksson M, Czene K, Vachon C, Conant EF, Hall P. A clinical risk model for personalized screening and prevention of breast cancer. Cancers (Basel). 2023;15(12):3246. https://doi.org/10.3390/cancers15123246.
23. Wei M, Du Y, Wu X, Su Q, Zhu J, Zheng L, et al. A benign and malignant breast tumor classification method via efficiently combining texture and morphological features on ultrasound images. Comput Math Methods Med. 2020;2020:5894010. https://doi.org/10.1155/2020/5894010.
24. Zhang X, Lin X, Zhang Z, Dong L, Sun X, Sun D, et al. Artificial intelligence medical ultrasound equipment: application of breast lesions detection. Ultrason Imaging. 2020;42(4–5):191–202. https://doi.org/10.1177/0161734620928453.
25. Xu Y, Wang Y, Yuan J, Cheng Q, Wang X, Carson PL. Medical breast ultrasound image segmentation by machine learning. Ultrasonics. 2019;91:1–9. https://doi.org/10.1016/j.ultras.2018.07.006.
26. Balch B. Why we know so little about women's health. Washington, D.C.: Association of American Medical Colleges; 2024 [cited 2024 Sept 15]. Available from: https://www.aamc.org/news/why-we-know-so-little-about-women-s-health.
27. U.S. Centers for Disease Control and Prevention. About Women and Heart Disease. Atlanta, GA: CDC; 2024 [cited 2024 Sept 15]. Available from: https://www.cdc.gov/heart-disease/about/women-and-heart-disease.html.
28. Sutton MY, Anachebe NF, Lee R, Skanes H. Racial and ethnic disparities in reproductive health services and outcomes, 2020. Obstet Gynecol. 2021;137(2):225–233. https://doi.org/10.1097/aog.0000000000004224.
29. Chinn JJ, Martin IK, Redmond N. Health equity among black women in the United States. J Womens Health (Larchmt). 2021;30(2):212–219. https://doi.org/10.1089/jwh.2020.8868.
30. U.S. Department of Health and Human Services. Heart Disease and African Americans. Rockville, MD: HHS; 2023 [cited 2024 Sept 15]. Available from: https://minorityhealth.hhs.gov/heart-disease-and-african-americans.
31. Dreyfus B, Kuri L, Ferri M, Doykos P, Fazeli MS, Hofer K, et al. Understanding Hispanic/Latino participation in clinical trials and observational studies, and strategies to increase participation: a targeted literature review. J Health Care Poor Underserved. 2023;34(1):399–424. https://doi.org/10.1353/hpu.2023.0026.
32. Ramjewan A. Exploring the future of artificial intelligence in cardiovascular medicine: A discussion with Mayo Clinic's Dr. Demilade Adedinsewo. Rochester, MN: Mayo Foundation for Medical Education and Research; 2022 [cited 2024 Sept 15]. Available from: https://innovationexchange.mayoclinic.org/future-of-artificial-intelligence-in-cardiovascular-medicine-with-mayo-clinics-dr-demilade-adedinsewo/.
33. McCarthy AM, Keller BM, Pantalone LM, Hsieh MK, Synnestvedt M, Conant EF, et al. Racial differences in quantitative measures of area and volumetric breast density. J Natl Cancer Inst. 2016;108(10):djw104. https://doi.org/10.1093/jnci/djw104.
34. Ahmed MI, Spooner B, Isherwood J, Lane M, Orrock E, Dennison A. A systematic review of the barriers to the implementation of artificial intelligence in healthcare. Cureus. 2023;15(10):e46454. https://doi.org/10.7759/cureus.46454.
35. Khanna NN, Maindarkar MA, Viswanathan V, Fernandes JFE, Paul S, Bhagawati M, et al. Economics of artificial intelligence in healthcare: diagnosis vs. treatment. Healthcare (Basel). 2022;10(12):2493. https://doi.org/10.3390/healthcare10122493.
36. Kasyanau A. Balancing the cost of AI in healthcare: future savings vs current spending. [Jersey City, NJ]: Forbes; 2024 [cited 2024 Nov 7]. Available from: https://www.forbes.com/councils/forbestechcouncil/2024/04/17/balancing-the-cost-of-ai-in-healthcare-future-savings-vs-current-spending/.
37. U.S. Department of Health and Human Services. Hispanic/Latino Health. Rockville, MD: HHS; 2024 [cited 2024 Sept 15]. Available from: https://minorityhealth.hhs.gov/hispaniclatino-health.
38. National Heart, Lung, and Blood Institute, National Institutes of Health. Women's Health Initiative (WHI). Bethesda, MD: NHLBI; [date unknown] [cited 2024 Sept 15]. Available from: https://www.nhlbi.nih.gov/science/womens-health-initiative-whi.
39. Skloot R. The immortal life of Henrietta Lacks. New York: Crown Publishing Group; 2010.

Manuscript received on 23 September 2024. Revised version accepted for publication on 22 November 2024.

El impacto actual de la inteligencia artificial: Implicaciones para quienes prestan servicios de salud a las mujeres de los efectos de la inteligencia artificial en la equidad en la salud

RESUMEN

Objetivo. Evaluar los efectos que el uso actual de la inteligencia artificial (IA) en la salud de las mujeres tiene sobre la equidad en la salud, específicamente en las actividades de prevención primaria y secundaria en las mujeres.

Método. Para realizar esta revisión narrativa se utilizaron dos bases de datos, Scopus y PubMed. En la búsqueda se utilizó el equivalente en inglés de algunas palabras clave como "inteligencia artificial", "aprendizaje automático", "salud de la mujer", "tamizaje", "factor de riesgo" y "prevenir", y los artículos solo se filtraron para incluir los que trataban sobre modelos de IA que los médicos generales podrían utilizar.

Resultados. De los 18 artículos examinados, 8 se centraron en la modelización de factores de riesgo en el marco de la prevención primaria y 10 se centraron en las herramientas de tamizaje en el marco de la prevención secundaria. Se encontraron brechas en la capacidad para entrenar a los modelos de IA con conjuntos de datos amplios y diversos que reflejen la población a la que están destinados. La falta de estos conjuntos de datos se detectó con frecuencia como una limitación en los artículos examinados (n = 7).

Conclusiones. Las mujeres pertenecientes a grupos minoritarios y de ingresos bajos tienen poco acceso a la atención de salud y, por lo tanto, no están bien representadas en los conjuntos de datos que se utilizan para entrenar los modelos de AI, lo que podría introducir sesgos en sus resultados. Para mitigar esto, deben crearse más conjuntos de datos para validar los modelos de IA, y la IA en la salud de las mujeres debe ampliarse para incluir las afecciones que afectan a hombres y mujeres, a fin de proporcionar una perspectiva de género al respecto. Las entidades de salud pública, medicina y tecnología deben colaborar para regular el desarrollo y el uso de la IA en la atención de salud de una manera estandarizada que reduzca los sesgos.

Palabras clave Inteligencia artificial; salud de la mujer; prevención primaria; prevención secundaria; ética.

O impacto contínuo da IA: repercussões dos efeitos da IA na equidade em saúde para profissionais de saúde da mulher

RESUMO

Objetivo. Avaliar os efeitos do atual uso da inteligência artificial (IA) na área de saúde da mulher sobre a equidade em saúde, especificamente em atividades de prevenção primária e secundária direcionadas para mulheres.

Métodos. Revisão narrativa de artigos indexados em duas bases de dados, Scopus e PubMed. As palavras-chave incluíram "artificial intelligence", "machine learning", "women's health", "screen", "risk factor" e "prevent", e os artigos foram filtrados de modo a incluir somente artigos sobre modelos de IA para uso por médicos generalistas.

Resultados. Dos 18 artigos examinados, 8 se concentraram na modelagem de fatores de risco na prevenção primária e 10, em ferramentas de rastreamento na prevenção secundária. Foram constatadas lacunas na capacidade de treinar os modelos de IA com conjuntos de dados grandes e diversificados que reflitam as populações às quais se destinam. A falta de tais conjuntos de dados foi frequentemente identificada como uma limitação nos artigos examinados (n = 7).

Conclusões. Mulheres minoritárias e de baixa renda têm acesso limitado à atenção à saúde e, portanto, estão sub-representadas nos conjuntos de dados utilizados para treinamento de modelos de IA, o que gera o risco da introdução de viés nos resultados. Para mitigar isso, é preciso desenvolver mais conjuntos de dados para validar os modelos de IA. Além disso, o uso da IA em saúde da mulher deve ser expandido para incluir afecções que afetem homens e mulheres, de modo a proporcionar uma perspectiva de gênero sobre essas afecções. As entidades de saúde pública, medicina e tecnologia precisam colaborar para regulamentar o desenvolvimento e o uso da IA na atenção à saúde de maneira a reduzir o viés.

Palavras-chave Inteligência artificial; saúde da mulher; prevenção primária; prevenção secundária; ética.