SJIF Impact Factor (2026): 8.863| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 11 | Issue: 3 | March 2026 - Peer Reviewed Journal

EQUITABLE AI-BIOSTATISTICAL FRAMEWORKS FOR BIAS MITIGATION IN CANCER PREDICTION AND OUTCOME ANALYTICS: A NARRATIVE REVIEW

Robert Amevor[1]
[1]*University of South Carolina, USA*

Article DOI: https://doi.org/10.36713/epra26770
DOI No: 10.36713/epra26770

ABSTRACT

Background: Cancer prediction AI models hold the potential for precision oncology but reinforce inequity with lifecycle bias. This narrative review summarizes recent findings on the sources of bias and the effectiveness of mitigation to suggest equitable biostatistical frameworks for oncology analytics.

Methods: This current paper reviewed 19 primary studies (2020-2025), identified through PubMed, Scopus, ResearchGate, and Google Scholar, that address AI analyses in oncology with fairness reporting. Thematic synthesis was applied across stages of the AI lifecycle, from conception to post-deployment.

Results: Bias was found to be a multi-faceted, lifecycle issue created due to the formulation of the problem, generation of data, and proxy outcomes. The proportion of models utilizing demographically diverse cohorts was only 22% with reports of race present in just 5% of papers, which comprised 88% white participants, where stated. In 11% of the models, environmental/life-course variables emerged. AI performance decreased from training (AUROC=0.93) and evaluation (AUROC=0.76) with uniform demographic disparities. Breast cancer staging models performed best in White compared to non-White patients, with older adults (≥70 years) having the worst performance despite having the highest cancer burden. Using the HEAL framework, fairness-health burden alignment was absent across age (0.0%), minimal across race/ethnicity (80.5%), and low in intersectional analyses (17.0%). Mitigation methods including resampling, adversarial learning and post-processing showed mixed and inconsistent effectiveness, with no method significantly improving fairness across all datasets.

Conclusion: AI in oncology perpetuates structural inequity through upstream design decisions and insufficient representativeness. Fair biostatistical models require burden-concordant fairness assessment, integration of social determinants, participatory formation of problems, and lifecycle regulation and constant sub-group tracking.

KEYWORDS: *Oncology AI bias; Fairness mitigation; Cancer prediction equity; Biostatistical frameworks; Lifecycle governance*

INTRODUCTION

Artificial intelligence and machine learning algorithms in cancer prediction and outcome analytics are one of the most revolutionary advancements in contemporary oncology [1]. These computational methods have shown promise for risk stratification, early diagnosis, prediction of the effect of treatment and for finding associations with patient survival across various types of cancer [2]. Yet, the potential for AI-based precision oncology is in jeopardy as evidence of algorithmic bias rises, producing biased predictions systemically across demographic subgroups according to race, ethnicity, socioeconomic status, sex and geographic regions [3, 4]. With US health systems scaling the use of AI-based clinical decision support tools, creating robust frameworks to identify, quantify and address bias has become a defining challenge at the crossroads of oncology, biostatistics and health equity [5, 6].

The origins of algorithm bias in cancer analytics are complex and sometimes even indicative of entrenched disparities within the health care system [7]. Training datasets are often biased toward minority groups, rural and low socioeconomic status individuals which results in models with poor performance for historically marginalized populations [4, 8]. In addition to representation gaps, biases can result from differential data quality where some populations are under-documented, followed up less often, or lack access to more sophisticated diagnostic technologies [9]. The selection of specific features could bias the model towards some explanatory factors which are simply proxies to protected characteristics, and the definition of outcomes can even be confounded by structural barriers to care rather than biology [10]. The intricate nature of these bias processes requires advanced model-oriented analytical methods which combine mechanisms derived from AI approaches and statistical principles based on solid biostatistical knowledge, such as causal inference or fairness theory [11].

SJIF Impact Factor (2026): 8.863| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 11 | Issue: 3 | March 2026 - Peer Reviewed Journal

It has been recently shown that there exist alarming inequities in the performance of AI models across subpopulations for medical imaging applications. AI based classifiers applied to chest radiographs have shown systematic underdiagnosis of pathologies in historically underserved demographics including female individuals, Black individuals and those with lower socioeconomic status, with even higher rates seen in intersectional groups like Hispanic females [8]. In breast cancer screening, a commercial AI algorithm for digital breast tomosynthesis exams showed increased false-positive rates in Black patients and those with dense breasts [12]. Although a few more recent deep learning models for mammography-based breast cancer risk prediction have demonstrated comparable performance in racial groups [13, 14], underreporting of race and ethnicity among AI algorithm development studies is still a significant obstacle to equity assessment [15]. These empirical results highlight an important mismatch between the technical metrics that are generally considered in AI research, and the equity-informed evaluation benchmarks needed to promote fair deployment into practice [16].

Classical biostatistics has long prioritized principles essential to equity-oriented inference such as strong study design, transparent variable selection, controlling confounding, quantifying uncertainty and performing subgroup analysis [4]. These characteristics may seem at odds with many modern AI models which are rated based on their predictive performance and scalability, but not interpretability, fairness evaluation or context quality [17]. Although tools for explainable AI (XAI), algorithmic fairness metrics, and bias mitigation have recently gained significant interest, their adoption in oncology is disjointed, methodologically disarrayed, and often devoid of biostatistical methods [18].

Various definitions and metrics for fairness may lead to conflicting conclusions, and rectification strategies have shown inconsistent efficacy at best across cancer categories, data types and population groups [19, 20]. This methodological inconsistency exposes a fundamental gap in the present body of cancer AI research. Without such integration, bias mitigation may be limited to shallow, reactive countermeasures that are not technologically aligned with real-world health disparities [21, 22]. Additionally, the lack of standardized reporting on demographics, subgroup performance, and data source precludes reproducibility, comparison across studies or regulatory scrutiny [23]. As AI-based frameworks become closer to deployment in clinical practice, the challenges of these deficiencies are becoming significant for patient safety, trust and health equity.

Therefore, a comprehensive synthesis of the literature is warranted to assess how equity concerns are currently being addressed within AI-based cancer prediction and outcome analytics. This includes but is not limited to critically enumerating sources of bias throughout the AI lifecycle, evaluating demographic representativeness in research and development, and observing differences in performance across sub-groups, as well as assessing existing strategies for mitigating bias. This narrative review seeks to summarize the available evidence on AI-based cancer prediction and outcome modeling in the context of equity and bias. It discusses the source, quantification and mitigation of bias with emphasis on how AI methods can be integrated within classical biostatistical frameworks. By synthesizing findings across tumor types, data sources, and analytic strategies, the review aims to guide efforts to formulate fair AI-biostatistical guidelines that may facilitate more equitable, robust, and clinically relevant cancer analytics.

METHOD

This narrative review synthesized current evidence on bias in oncology AI prediction models. Literature was searched through PubMed, Scopus, ResearchGate and Google Scholar. Nineteen primary studies published in the last 5 years (2020-2025) were selected from 347 records, based on direct relevance to bias identification, mitigation strategies and equity evaluation within cancer prediction and outcome analytics. Theoretical and other supporting literature also offered fairness frameworks, governance criteria, and biostatistical techniques. Narrative review principles were used to thematically synthesize studies by the AI lifecycle stage including conceptualization, data generation, model development, evaluation, deployment and post-deployment. Quantitative results including AUROCs, fairness metrics and demographic representation were extracted and descriptively summarized.

FINDINGS

Conceptualization, Taxonomies, and Lifecycle Origins of Bias in Oncology AI

In all reviewed articles, bias in AI-based cancer prediction and outcome analytics was viewed as a complex and lifecycle-wide phenomenon, which arises in the conceptualization of the problem, the creation of data, the creation of the model, the evaluation of its predictions, its implementation, and the post-implementation utilization [20-22, 24-26]. A single coherent conceptual model divided bias into evidence or research bias, expertise or provider bias, exclusion or embedded data bias, environmental and life-course exposure bias, and empathy-related bias which represented poor contextualization of patient information [21].

SJIF Impact Factor (2026): 8.863| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 11 | Issue: 3 | March 2026 - Peer Reviewed Journal

There were frequent reports of provider-level bias, such as bias in diagnostic reasoning, stereotyping in clinical decision-making, and inaccuracies or omissions in electronic health record (EHR) documentation that are carried forward to downstream predictive models [21]. Another mechanism of bias was the formulation of problems, especially where proxy outcomes like healthcare utilization or cost or length of stay were being modeled. These proxies were related to structural biases in care access as opposed to clinical need, which led to systematic under-identification of high-need patients in historically underserved groups [22, 24, 25].

Algorithms studies also found the problem of embedded data bias, selective reporting of outcomes and metrics, uneven treatment of missing data, and extensive use of artificial or imputed data sets without a standardized rationale as common methodological shortcomings [27]. Measurement bias, such as instrument errors and variability based on clinicians, was recognized to be relevant to oncology diagnostics and monitoring but only a limited number of studies explicitly assessed this [27]. Additionally, biases were regarded as recurrent throughout the AI lifecycle, where insufficient mitigation at the initial levels exacerbates downstream inequity in the process of validation and deployment [22, 24-26].

Dataset Composition, Representativeness, and Demographic Reporting Practices

The composition of data and the overall demographic representation across studies was found to be very heterogeneous. A relatively high racial diversity was demonstrated in one large assessment group (n = 43,274) with 52.2% White, 18.9% Black, 9.5% Asian, and 19.2% other or unknown participants. Nonetheless, the representation of native Americans was sufficiently low (0.1%), which did not allow subgroup analysis [28]. Black patients in this cohort were also a larger proportion of the deceased subgroup than survivors (25.2% vs 18.1%), which reflected outcome disparities that existed even before AI modeling [28].

Conversely, the majority of AI studies in the field of oncology were based on datasets that were not extensively diverse in demographics. Only 22% of the reviewed models were trained on cohorts that are explicitly stated to be demographically diverse [27]. Of all published oncology AI literature from2016-2021, 47 of them reported some form of training or validation data, 58 percent of them reported the age of the population, 51 percent reported sex, and only 5 percent reported race or ethnicity [23]. In the case where race was reported, about 88 percent of the respondents were White with little granularity past broad categories [23].

The underrepresentation was spread between the sources of data, such as clinical trials, EHRs, and administrative datasets. Less than 20 percent of the trials regulated by FDA cited race-specific effects or negative outcomes [21]. The sociodemographic variables were often combined, overlooking the heterogeneity within the groups, especially between the Black, Asian, and Hispanic groups [21, 23]. Only 11% of reviewed models included environmental, occupational, and life-course exposure variables yet these variables have been demonstrated to be of relevance to cancer risk and outcomes [27].

Baseline Cancer Burden and Population Health Disparities Informing AI Equity

Population-level analyses offered essential baseline information on the assessment of AI equity. The data on global burden showed significant racial and ethnic differences in the number of years lost (YLL) with the lowest YLL (45.2) being Asian and Pacific Islander, then Hispanic (70.4), Black (131.4), and White populations (223.7) [29]. There were small differences in sex-based disability adjusted life years, with the males and females having equal DALYs (830.6 vs 832.6) [29].

The age stratified analyses repeatedly demonstrated that people aged 70 years and above had the highest cancer-related burden followed by people aged 50-69 years, and then the burden continuously decreased with the age [29]. These differences were highly linked with social determinants of health, such as socioeconomic deprivation, financial constraint, geographic accessibility to oncology care, environmental exposure, and unequal access to screening and early detection strategies [29, 30].

Model Inputs, Feature Selection, and Reporting of Variable Importance

Researchers drew on diverse inputs for their AI models, blending clinical data, lab results, imaging findings, molecular profiles, and healthcare utilization metrics. In a prominent mortality prediction effort, for instance, the model relied on 35 features. Albumin levels, inpatient admission frequency, chloride concentrations, lymphocyte percentages, and heart rate emerged as the strongest predictors [28]. Cancer stage failed to rank in the top 20. Cancer-related factors like treatment regimens carried weight but typically trailed behind broader physiological markers [28].

Studies employing algorithms such as Random Forests, support vector machines, logistic regression, XGBoost, gradient boosting, deep neural networks, and ensembles featured markedly different variable sets. This hampered direct comparisons across works [27]. Techniques like SHAP, LIME, CAM, and Grad-CAM consistently pinpointed immune cell proportions, molecular signatures, imaging

SJIF Impact Factor (2026): 8.863| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 11 | Issue: 3 | March 2026 - Peer Reviewed Journal

textures, and patient clinical details as key drivers of model outputs [16, 24-26]. A recurring insight from explainability tools was that models often captured subtle non-clinical cues such as race, site of care, or socioeconomic indicators. These exposed embedded biases [24-26]. Reports on feature importance were common, yet scant attention went to hyperparameter tuning, feature engineering logic, handling of missing values, or the rationale for sample sizes [27]. Social determinants of health, patient values, and situational factors saw little incorporation [21, 30].

Predictive Performance and Demographic Subgroup Variability

AI models generally achieved solid discriminatory power, but performance often dipped from training to testing phases. One extensive mortality study posted an evaluation AUROC of 0.76 (95% CI, 0.75-0.77), down from 0.93 in training. Sensitivity stood at 0.71 with an F1-score of 0.35 [28]. Racial subgroup AUROCs stayed relatively stable, hovering between 0.75 and 0.77 [28].

In the wider literature, metrics span 0.80 to 1.00, especially for pancreatic and breast cancer applications. Patchy validation details clouded assessments [27]. Disparities by demographics surfaced reliably across prediction tasks and data types. Deep learning for breast cancer staging delivered better accuracy, precision, recall, and F1-scores for White patients than non-White ones. No model flipped that pattern [24, 25, 31]. Non-White groups faced elevated false positives and reduced true positives. Differences reached statistical significance in several cases [24, 31].

Age gaps were stark, too. Those 70 and older shouldered the heaviest cancer load, yet the weakest AI results [29]. Sex differences varied by task. For chemotherapy toxicity, bilirubin forecasts faltered more in women, while creatinine and bilirubin predictions held up better for seniors [19]. Imaging fairness also varied spatially by body region. Some areas showed big gaps, even when overall metrics looked fine [20].

Fairness Definitions, Metrics, and Alignment with Health Burden

Most fairness assessments relied on group-stratified evaluations, slicing data by age, sex, race, ethnicity, or skin tone to gauge subgroup metrics [19, 20, 26]. A range of measures came into play: differences via subtraction, ratios, calibrated error rates, entropy indices, normalized ratios, and inequality stats like $I^2(b)$ and I^2_β [19, 20, 26]. Multiple papers showed how the same model could pass one fairness test yet fail another. This underscored clashing criteria and a no one-size-fits-all benchmark [20, 25, 26]. Chemotherapy toxicity models revealed minor but persistent inequities across cycles, never dropping to zero [19].

The HEAL framework revealed middling alignment of AI performance with actual health burdens by race/ethnicity (80.5%) and sex (92.1%), but none by age (0.0%) [29]. Cancer tasks aligned better (73.8%) than non-cancer ones (0.0%) [29]. Layering age, sex, and race together exposed even poorer matches (17.0%). This flagged high-burden groups underserved by AI [23, 29, 30].

Bias Mitigation Strategies and Empirical Effectiveness

Efforts to curb bias spanned pre-, in-, and post-processing, but results were inconsistent and often underwhelming [19, 20, 24, 26, 31]. Pre-processing tactics encompassed resampling, stratified sampling, normalization, confounder stripping, synthetic data, and outside dataset fusion [20, 26]. In-processing drew on adversarial training, fairness-aware optimization, disentangled representations, domain shifts, federated setups, and targeted fine-tuning [20, 26]. Post-processing targeted threshold tweaks per subgroup, calibration, reject options, and result trimming [20, 31].

Benchmarking across datasets found no strategy reliably boosted fairness with statistical backing [20]. Post-processing in breast cancer staging could not reliably shrink false-positive or true-positive gaps [24, 25, 31]. Multi-class tasks suffered from weak baselines. This rendered fairness tweaks moot for equity or accuracy [26, 31]. Pooling multi-hospital data did lift both general and subgroup performance in toxicity forecasting [19].

Data Quality, Missingness, Transparency, and Reproducibility

The limitation of data quality was widespread. Only a quarter of the studies provided a gap in data management policies, and none of them justified adequacy of sample size or outlined data augmentation methodology sufficiently [27]. In 63 percent of the studies, synthetic or imputed data were utilized, and the assumptions and implications of bias were frequently not disclosed [27]. Outlier, measurement error, and environmental confounding assessment were infrequent [27]. The collection of sociodemographic data was not standardized, which constrained the contextual interpretation of model results [21, 23]. Additionally, there was lack of transparency and reproducibility. Data sharing and material reproducibility were only found in 22 percent of studies, and none satisfied transparency requirements [27]. These reporting lapses limited the use of systematic assessment of bias, generalizability, and equity [23, 29].

SJIF Impact Factor (2026): 8.863| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 11 | Issue: 3 | March 2026 - Peer Reviewed Journal

Deployment, Human Oversight, Governance, and Equity-Oriented Applications

The issues of data shift, automation bias, and workflow integration were pointed out in deployment studies [19, 25, 26, 32]. High-risk patients identified by models in cancer risk stratification programs represented a disproportionate proportion of preventable acute care use, although the vast majority of patients enrolled to the program were identified by clinician referral as opposed to model output alone, which highlights the need to persist with human oversight [32]. Models based on chemotherapy toxicity proved to be robust to changes in time and extrapolation in unobserved tumor types, but subgroup unfairness grew modestly during treatment cycles [19]. Misleading AI recommendations proved to worsen clinician performance [25, 26].

The approaches to governance focused on bias as a lifecycle phenomenon that needs to be constantly monitored, developed by heterogeneous teams, has its demographics clearly reported, evaluated in subgroups, and recalibrated to accommodate concept drift [22, 24-26]. Regulatory frameworks were characterized as abstract and lacking standardized terms of operations in terms of equity evaluation [22, 24, 25]. In addition to prediction, AI systems aided equity-based cancer care infrastructure such as the identification of underserved populations, resource distribution, tele-oncology, and recruitment of clinical trials. Trial matching with the assistance of AI enhanced the identification of eligible participants belonging to underrepresented groups [22, 24, 25].

DISCUSSION

The current narrative review describe bias as a multilevel, lifecycle phenomenon and consider the effectiveness of interventions aimed at mitigating its influence. There are three key implications; bias occurs upstream of model development and thus technical solutions cannot fix current biases; health-fairness misalignment results in safety and legitimacy crisis for clinical implementation; and equitable AI necessitates structural reform based on participatory engagement, embedded social epidemiology, and sustained governance.

Bias originates at problem formulation, data generation, and outcome proxy selection, yet mitigation strategies target only algorithmic patterns [20-22, 24-26]. When researchers train models to predict healthcare utilization rather than clinical need, they mechanize structural inequities [3]. Current mitigation approaches, including resampling, adversarial learning, and fairness optimization, achieved mixed effectiveness precisely because the problem has been defined in a wrong way [33]. Equitable AI frameworks must begin with participatory problem formulation involving affected communities to define outcomes reflecting clinical need rather than system characteristics [5].

Existing fairness metrics make subgroup performance parity operational but are agnostic to burden distribution of the disease [34]. Fairness alignment was absent by age (0.0%), minimal by race/ethnicity (80.5%) and at 17.0% in intersectional analyses using the HEAL framework [29]. Paradoxically, all general performance parameters (AUROC 0.76) seemed acceptable for clinical use [28]. Older adults over 70 had the highest cancer burden and lowest AI performance, an ethical misalignment not visible in typical fairness reporting. This creates perverse incentives where developers can achieve procedural fairness while systematically failing highest need populations [29]. Only 22% of models were trained on demographically diverse populations. Race was reported in only 5% of papers and, when reported, about 88% of participants were White [23, 27]. Every homogenous cohort is a compounding equity debt, where models trained with diverse populations will inevitably be biased due to domain shift and unmeasured confounders (Hashimoto et al., 2018). Furthermore, environmental, occupational and life course factors were only included in 11% of the models even though they are known to be key overall determinants of cancer outcomes [35]. This represents an epistemic failure. Predictive accuracy emanates from statistically proximal biomedical variables, whereas equity requires the assimilation of socially proximal determinants [1]. Models need to incorporate social epidemiology and view social determinants as central drivers of cancer burden, not confounders.

Deep learning stages for breast cancer attained higher accuracy among Whites, and no model maintained superior performance for non-Whites. Error analyses showed that non-White population had relatively high false positive rates and low true positive rates [25, 31, 36]. These differences directly translate into delayed diagnosis, underestimated treatment and clinical harms. Unreliable AI suggestions caused a critical decline in clinical skills leading to overuse of automation bias, exacerbating harm under the conditions of scarce resources [26]. Performance disparities are mechanisms of clinical inequity, not acceptable as disclosed limitations [37]. In addition, only one in four papers disclosed methods for handling missing data, while 63% employed synthetic or imputed datasets with opaque rationale and less than one in five provided reproducibility materials [27]. Such gaps do not allow for external validation and community examination. For AI to be fair, transparency is a prerequisite for accountability [7]. Regulatory reform should require reproducibility standards, transparent data- handling, and justified imputation prior to publication in journals or clinical implementation [5].

Equity-focused governance necessitates diverse development teams, with patients from the highest-burden populations, and requires subgroup performance testing before and after deployment [3]. It necessitates persistent post-deployment surveillance and response-

SJIF Impact Factor (2026): 8.863| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 11 | Issue: 3 | March 2026 - Peer Reviewed Journal

triggering approaches, as well as population-level participation in governance decisions [33]. Requirements for equitable frameworks also include the operationalization of social determinants, through further data gathering, causal inference models to uncover paths, algorithms which trigger on social risk drivers, and stratified clinical guidance based on patient circumstances [38]. This will necessitate health system reform and not just statistical optimization.

Limitations and Research Gaps

The review was based on recent published oncology AI studies. Unpublished data and proprietary algorithms in clinical use which are unknown could have skewed study findings. Diverse definitions, mitigation strategies and fairness metrics precluded quantitative synthesis and meta-analysis of bias mitigation efficacy. The rigor of the studies varied, with some providing sophisticated fairness analyses and others reporting on aggregate performance. Furthermore, there were very few studies on intersectional disparities and long-term deployment outcomes, limiting the understanding of how bias arises and changes in real clinical systems.

Implications for Research, Practice, and Policy

For research, this review specifies that equitable AI-biostatistical frameworks should not be centered just in prediction accuracy but expanded toward bias origin interrogation, fairness-equity alignment verification and governance infrastructure. Prospective studies should focus on whether deployed algorithms reduce or exacerbate cancer disparities over time and participatory research where algorithms are co-designed with impacted communities. Future research should also incorporate causal inference and social epidemiological principles into AI development as well as emphasize health system implementation that evaluate the effect of equitable AI on clinical workflows, decision making, and outcomes.

In clinical practice, this involving application of equitable AI should not be passive. Requiring subgroup performance reporting, incorporating equity officers into algorithmic governance, and requiring an intermediate intervention between the delivery of a non-complaint result for excluded groups should be systemic changes to mitigate these disparities. Training courses are needed for oncologists to learn how AI is biased and how fairness of outcomes has limitations, as well as establishing an ethical framework around using AI.

For policy and regulation, this review highlights the limitations of existing AI governance approaches. Regulators should require demographic diversity and transparency as a condition of clearance, burden-concordant fairness as a regulatory standard, post-deployment surveillance for disparities with guidelines for timely action when detected. Clear standards for handling missing data, synthetic data generation and use of augmentation should also be established. Moreover, open-science requirements including reproducibility to accompany algorithms used in clinical actions should be enforced. In addition, funding agencies should also create financial incentives for responsible AI development via specific funding lines and career paths. This would signal that bias mitigation is a scientific priority rather than an ethical afterthought.

CONCLUSION

This narrative review reveals that bias in oncology AI is multi-dimensional, stemming from structural and epistemic decisions made before the model's creation, and incompletely rectified by existing mitigation strategies. Despite fairness optimization, performance imbalances persist while discrepancies between algorithmic fairness and cancer health burden incurs legitimacy and safety crisis. AI-biostatistical guidelines demand more than an incremental shift in technical performance but a structural transformation. This consists of participatory algorithm design, incorporation of social determinants and causal inference, burden-concordant fairness analysis, as well as ongoing equity measurement integrated into health system governance. Only by developing integrative, interdependent and transparent methods such as these can AI be used to support rather than mechanize cancer care disparities and realize the promise of precision medicine for all populations.

REFERENCES

1. *Topol, E.J., High-performance medicine: the convergence of human and artificial intelligence. Nature medicine, 2019.* ***25****(1): p. 44-56.*
2. *Haug, C.J. and J.M. Drazen, Artificial intelligence and machine learning in clinical medicine, 2023. New England Journal of Medicine, 2023.* ***388****(13): p. 1201-1208.*
3. *Gianfrancesco, M.A., et al., Potential biases in machine learning algorithms using electronic health record data. JAMA internal medicine, 2018.* ***178****(11): p. 1544-1547.*
4. *Rajkomar, A., et al., Scalable and accurate deep learning with electronic health records. NPJ digital medicine, 2018.* ***1****(1): p. 18.*
5. *Char, D.S., N.H. Shah, and D. Magnus, Implementing machine learning in health care – addressing ethical challenges. The New England journal of medicine, 2018.* ***378****(11): p. 981.*
6. *Parikh, R.B., S. Teeple, and A.S. Navathe, Addressing bias in artificial intelligence in health care. Jama, 2019.* ***322****(24): p. 2377-2378.*

SJIF Impact Factor (2026): 8.863| ISI I.F. Value: 1.241| Journal DOI: 10.36713/epra2016 ISSN: 2455-7838(Online)
EPRA International Journal of Research and Development (IJRD)
Volume: 11 | Issue: 3 | March 2026 - Peer Reviewed Journal

7. *Obermeyer, Z., et al., Dissecting racial bias in an algorithm used to manage the health of populations. Science, 2019.* ***366****(6464): p. 447-453.*
8. *Seyyed-Kalantari, L., et al., Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations. Nature medicine, 2021.* ***27****(12): p. 2176-2182.*
9. *Johnson, A.E., et al., MIMIC-III, a freely accessible critical care database. Scientific data, 2016.* ***3****(1): p. 1-9.*
10. *Vyas, D.A., L.G. Eisenstein, and D.S. Jones, Hidden in plain sight—reconsidering the use of race correction in clinical algorithms. 2020, Mass Medical Soc. p. 874-882.*
11. *Mitchell, S., et al., Algorithmic fairness: Choices, assumptions, and definitions. Annual review of statistics and its application, 2021.* ***8****(1): p. 141-163.*
12. *Nguyen, D.L., et al., Patient characteristics impact performance of AI algorithm in interpreting negative screening digital breast tomosynthesis studies. Radiology, 2024.* ***311****(2): p. e232286.*
13. *Yala, A., et al., A deep learning mammography-based model for improved breast cancer risk prediction. Radiology, 2019.* ***292****(1): p. 60-66.*
14. *Yala, A., et al., Toward robust mammography-based models for breast cancer risk. Science Translational Medicine, 2021.* ***13****(578): p. eaba4373.*
15. *Miyawaki, I.A., et al., Global disparities in artificial intelligence-based mammogram interpretation for breast cancer: A scientometric analysis of representation, trends, and equity. European journal of cancer (Oxford, England: 1990), 2025.* ***220****: p. 115394.*
16. *Ghasemi, A., et al., Explainable artificial intelligence in breast cancer detection and risk prediction: A systematic scoping review. Cancer Innovation, 2024.* ***3****(5): p. e136.*
17. *Mbakwe, A.B., et al., Fairness metrics for health AI: we have a long way to go. EBioMedicine, 2023.* ***90****.*
18. *Mohamed, Y.A., et al., Decoding the black box: Explainable AI (XAI) for cancer diagnosis, prognosis, and treatment planning-A state-of-the art systematic review. International Journal of Medical Informatics, 2025.* ***193****: p. 105689.*
19. *Watson, M., et al., From prediction to practice: mitigating bias and data shift in machine-learning models for chemotherapy-induced organ dysfunction across unseen cancers. BMJ oncology, 2024.* ***3****(1): p. e000430.*
20. *Xu, Z., et al., Addressing fairness issues in deep learning-based medical image analysis: a systematic review. npj Digital Medicine, 2024.* ***7****(1): p. 286.*
21. *Dankwa-Mullan, I. and D. Weeraratne, Artificial intelligence and machine learning technologies in cancer care: addressing disparities, bias, and data diversity. Cancer Discovery, 2022.* ***12****(6): p. 1423-1427.*
22. *Singh, Y., et al., Beyond the hype: navigating bias in AI-driven cancer detection. Oncotarget, 2024.* ***15****: p. 764.*
23. *D'Amiano, A.J., et al., Transparency and representation in clinical research utilizing artificial intelligence in oncology: a scoping review. Cancer medicine, 2025.* ***14****(5): p. e70728.*
24. *Chinta, S.V., et al., AI-Driven Healthcare: A Review on Ensuring Fairness and Mitigating Bias. arXiv preprint arXiv:2407.19655, 2024.*
25. *Gichoya, J.W., et al., AI pitfalls and what not to do: mitigating bias in AI. The British Journal of Radiology, 2023.* ***96****(1150): p. 20230023.*
26. *Hasanzadeh, F., et al., Bias recognition and mitigation strategies in artificial intelligence healthcare applications. NPJ Digital Medicine, 2025.* ***8****(1): p. 154.*
27. *Smiley, A., et al., Exploring artificial intelligence biases in predictive models for cancer diagnosis. Cancers, 2025.* ***17****(3): p. 407.*
28. *Ganta, T., et al., Fairness in predicting cancer mortality across racial subgroups. JAMA Network Open, 2024.* ***7****(7): p. e2421290-e2421290.*
29. *Schaekermann, M., et al., Health equity assessment of machine learning performance (HEAL): a framework and dermatology AI model case study. EClinicalMedicine, 2024.* ***70****.*
30. *Srivastav, A.K., et al., Revolutionizing Oncology Through AI: Addressing Cancer Disparities by Improving Screening, Treatment, and Survival Outcomes via Integration of Social Determinants of Health. Cancers, 2025.* ***17****(17): p. 2866.*
31. *Soltan, A. and P. Washington, Challenges in reducing bias using post-processing fairness for breast cancer stage classification with deep learning. Algorithms, 2024.* ***17****(4): p. 141.*
32. *Osterman, C.K., et al., Predictive modeling for adverse events and risk stratification programs for people receiving cancer treatment. JCO Oncology Practice, 2022.* ***18****(2): p. 127-136.*
33. *Mehrabi, N., et al., A survey on bias and fairness in machine learning. ACM computing surveys (CSUR), 2021.* ***54****(6): p. 1-35.*
34. *Carey, S., A. Pang, and M. Kamps, Fairness in AI for healthcare. Future Healthc J, 2024.* ***11****(3): p. 100177.*
35. *Hashimoto, D.A., et al., Artificial intelligence in surgery: promises and perils. Annals of surgery, 2018.* ***268****(1): p. 70-76.*
36. *Garcia-Saiso, S., et al., Artificial Intelligence as a potential catalyst to a more equitable cancer care. JMIR cancer, 2024.* ***10****: p. e57276.*
37. *Magrabi, F., et al., Artificial Intelligence in Clinical Decision Support: Challenges for Evaluating AI and Practical Implications. Yearb Med Inform, 2019.* ***28****(1): p. 128-134.*
38. *Alonge, O., Using causal models and theories to achieve equitable implementation science in global health. BMC Glob Public Health, 2025.* ***3****(1): p. 85.*