Title

Mitigating Racial And Ethnic Bias And Advancing Health Equity In Clinical Algorithms: A Scoping Review.

Journal
Health Affairs, 42(10)

Authors
Cary, Michael
Zink, Anna
Wei, Sijia
et al.

Publication Date
2023-10-01

DOI
10.1377/hlthaff.2023.00553

Peer reviewed

HHS Public Access

Author manuscript

Health Aff (Millwood). Author manuscript; available in PMC 2023 November 24.

Published in final edited form as:

Health Aff (Millwood). 2023 October ; 42(10): 1359–1368. doi:10.1377/hlthaff.2023.00553.

Mitigating Racial and Ethnic Bias and Advancing Health Equity in the Development, Evaluation, and Deployment of Clinical Algorithms in Healthcare: A Scoping Review and Implications for Policy

Michael P. Cary Jr.[1,3], **Anna Zink**[2], **Sijia Wei**[3], **Andrew Olson**[4], **Mengying Yan**[4], **Rashaud Senior**[4], **Sophia Bessias**[4], **Kais Gadhoumi**[1], **Genevieve Jean-Pierre**[5], **Demy Wang**[1], **Leila S. Ledbetter**[8], **Nicoleta J. Economou-Zavlanos**[4], **Ziad Obermeyer**[6], **Michael J. Pencina**[4,7]

[1]Duke University School of Nursing, Durham, NC

[2]University of Chicago School of Business, Chicago, IL

[3]Northwestern University School of Medicine, Chicago, IL

[4]AI Health, Duke University School of Medicine, Durham, NC

[5]Duke University Department of Computer Science, Durham NC

[6]University of California Berkeley School of Public Health, Berkeley, CA

[7]Duke Clinical Research Institute, Durham, NC

[8]Duke University Medical Center Library, Durham, NC

Abstract

In July 2022, the Department of Health and Human Services (HHS) issued a notice of proposed rulemaking prohibiting covered entities, which include health care providers and health plans, from discriminating against individuals when using clinical algorithms in decision-making, but it did not provide specific guidelines on how covered entities should prevent discrimination. We conducted a scoping review of literature published from 2011 to 2022 to identify practical strategies, frameworks, reviews and perspectives, and assessment tools that identify and mitigate bias in clinical algorithms, with a specific focus on racial and ethnic bias. Our scoping review encompassed 109 articles comprising forty-five empirical health care applications that included tools tested in health care settings, sixteen frameworks, and forty-eight reviews and perspectives. We identified a wide range of technical, operational, and systemwide bias mitigation strategies for clinical algorithms, but there was no consensus in the literature on a single best practice that covered entities could employ to meet the HHS requirements. Future research should identify optimal bias mitigation methods for various scenarios, depending on factors such as patient population, clinical setting, algorithm design, and types of bias to be addressed.

Corresponding author's contact information: Michael P. Cary, Jr., PhD, RN, Duke University School of Nursing, 307 Trent Drive, Durham, NC 27710 michael.cary@duke.edu.

In July 2022, the Department of Health and Human Services (HHS) issued a notice of proposed rulemaking that would revise the interpretation of Section 1557 of the Patient Protection and Affordable Care Act, which prohibits discrimination on the basis of race, color, national origin, sex, age, or disability.[1] This notice includes a new provision stating that covered entities, which include health care providers and health plans,[2] must not discriminate against any individual through the use of clinical algorithms in decision making —however, it does not specify what measures covered entities should take to ensure this. Instead, it solicits comments on practices to ensure that algorithms are not discriminatory and requests resources and recommendations on identifying and mitigating discrimination that results from the use of clinical algorithms.

In response to concerns about algorithmic bias, other federal agencies and nonprofit organizations have recently taken action to regulate algorithms or support governance and oversight measures aimed at making algorithms safe, fair, and transparent.[3,4] These efforts are reflected in the Food and Drug Administration's final guidance on the principles of software validation[5]; the Agency for Healthcare Research and Quality's research protocol on healthcare algorithms on racial and ethnic disparities[6]; the National Institute of Standards and Technology's artificial intelligence risk-management framework[7]; the Coalition for Health AI's "Blueprint for trustworthy AI implementation guidance and assurance for healthcare," version 1.0[8]; and the White House Office of Science and Technology Policy's "Blueprint for an AI bill of rights.[9]

These publications join a growing body of academic and professional literature in defining, describing, and providing ways to measure algorithmic bias harmonized around the articulation of principles to guide the development of algorithms, such as safety, fairness, and transparency. Although there is broad agreement on the need to remove harmful bias from clinical algorithms, there is little consensus on how to achieve this critical objective. As health care systems develop and implement these technologies, researchers, developers, and clinicians who build and deploy clinical algorithms need concrete strategies, methods, and tools that enable them to identify and mitigate bias.

Anticipating that many covered entities will actively assess whether their clinical algorithms comply with the proposed new Section 1557 provision and attempt to correct any biases they uncover, we identified practical strategies, frameworks, and tools for identifying and mitigating bias in clinical algorithms, focusing on racial and ethnic bias. Our study expanded on previous reviews[10–13] to capture the full breadth of published resources. We incorporated articles from journals focused on clinical practice, public health, ethics, law, public policy, data science, computer science, machine learning, and artificial intelligence. This comprehensive, multidisciplinary scope allowed us to summarize a full suite of mitigation approaches that can be applied to the HHS directive to prevent discrimination arising from use of algorithms in clinical decision making.

Study Data And Methods

Design

We followed the JBI scoping review methodology and are reporting the Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR).[14–16]

Information Sources

The databases searched were Medline (PubMed), Embase (Elsevier), Web of Science (Clarivate), and ProQuest Computer Science Database and ProQuest Dissertation and Theses Global.

Search Strategy

The search was developed and conducted by a medical librarian, with input from the other coauthors, and included a mix of keywords and subject headings, including algorithm, bias, mitigation and assessment, health care, and race and ethnicity. The original searches were conducted on August 24, 2022, and we found 18,028 citations. The searches were independently peer reviewed by another medical librarian, using a modified Peer Review of Electronic Search Strategies (PRESS) checklist. The author team also assessed several sources of grey literature through a targeted web search of government entities, including the National Institute of Standards and Technology. Full reproducible search strategies for all included databases are detailed in online appendix exhibit A1.[17]

Study Selection

After the search, all identified studies were uploaded into Covidence, a software system for managing systematic reviews. Duplicates were removed by the software ($n = 6{,}381$). A final set of 11,647 citations was left to be screened in the study title and abstract phase.

Eligibility Criteria

We systematically screened papers to include applications, frameworks, reviews and perspectives, and tools that dealt with racial and ethnic bias mitigation in algorithms used either to guide care decisions for individual patients or to inform decisions for achieving population health goals, such as efficiently allocating health care resources; mitigation strategies could be applied at any stage in the algorithm development lifecycle, from pre- to postdeployment. We excluded conference abstracts and dissertations.

After study titles and abstracts were screened for relevance, full texts of publications were reviewed to verify that they met inclusion and exclusion criteria. All aspects of screening were performed by at least two independent reviewers. At each stage, disagreements between reviewers were resolved by a third reviewer. The study selection process is represented in appendix exhibit A2.[17]

Data Extraction And Synthesis

Before extracting studies, we tested the extraction matrix on four studies—one for each study type: applications, tools, reviews and perspectives, and frameworks. We revised the matrix to reflect consensus among reviewers and developed an extraction manual to ensure coding consistency. Because there were relatively few tool articles and we extracted the same data elements from the tools and applications articles, we then combined applications and tools into one category. Two reviewers independently extracted data from the full text of all eligible articles. Conflicts between reviewers were resolved by a third reviewer.

Limitations

Our study had several limitations. First, our review was limited to clinical algorithms used by health care providers and excluded rule-based algorithms. Second, our search returned few to no examples of mitigation methods for text-based algorithms and generative artificial intelligence, and we cannot speak to the best mitigation methods for these types of algorithms. Other limitations included restrictions imposed on our sample (English-language only, 2011–22 study interval). Despite our efforts to implement best practices for a comprehensive and inclusive search strategy, we may have missed some relevant studies during data extraction if specific mitigation strategies were missed or misclassified by independent reviewers.

Study Results

Of the remaining 11,579 articles, 11,233 were excluded during title and abstract screening. A total of 346 papers were sought for retrieval; one was unavailable and the full texts of the remaining 345 papers (plus an additional nine papers identified from other methods via targeted web searches) were assessed for eligibility on the basis of the study criteria. After excluding 245 articles that did not meet study criteria, a final total of 109 articles were included, comprising 106 from the database search plus three identified from other sources (a full list of the studies included in our review is available in appendix exhibit A3).[17]

Of the three article types (applications and tools, frameworks, and reviews and perspectives), the most common types were reviews and perspectives (n = 48 [44 percent]) and health care applications and tools (n = 45 [47 percent]). The least common type was framework (n = 16 [15 percent]). Most articles (n = 101 [93 percent]) identified were published within the last four years (exhibit 1).

Mitigation Strategies By Article Type

To facilitate interpretation of our findings, we classified mitigation strategies into three strategy categories (exhibit 2): technical (such as data collection, algorithm design, preprocessing, processing, postprocessing, and monitoring postdeployment), operational (governance, design principles, and interdisciplinary multistakeholders), and systemwide (training and education, collaborative platforms, standards, incentives, and regulation).

There were notable differences in the kinds of mitigation strategies discussed across article types. Most applications and tool articles focused on technical solutions to bias mitigation, such as better selection of prediction targets and reweighting sample data to resemble the target population. Frameworks and reviews and perspectives, in contrast, more often included discussions of nontechnical solutions, such as the role of governance and need for systemwide standards.

Technical Strategies

Technical strategies were grouped into six categories on the basis of the point in the algorithmic lifecycle at which bias was addressed: data collection (strategies that involved the collection of higher-quality data from representative patient populations), algorithm design (strategies related to the selection of the outcome, predictor variables, and algorithms used for the prediction problem), preprocessing (strategies that occurred before training the algorithm—that is, the process of teaching it to make predictions or decisions based on a data set—including changing the data via weighting[18–22] or sampling[18,23–25] methods to make it more representative of the population in which the algorithm would be used), processing (strategies that alter the training of the algorithm, including adjusting the algorithm's objective function to incorporate some aspect fairness in addition to statistical fit[24–31]; this could include a mathematical formula that attempts to maximize or minimize the algorithm's goal—for instance, to minimize the errors while ensuring that error rates across race groups are similar), postprocessing (strategies that update the results after prediction through recalibration[32] or varying cut points or thresholds at which the model output is used to define a category of risk or recommend that a clinician take action to achieve fairness[20,33–35]), and monitoring postdeployment (strategies related to tracking the performance of the algorithm after it has been trained, such as checking that the algorithm's performance is not degrading over time and measuring the impact on treatment allocation and health outcomes).

In exhibit 3 we summarize some of the most common technical strategies for mitigating racial and ethnic bias. These include stratifying or calibrating algorithms by race, weighting methods, and adjusting the algorithm's objective function. These strategies span clinical applications and algorithms. Almost every article reported some success in mitigating bias using these strategies, although this sometimes came at a cost to other statistical performance measures.[21]

Operational Strategies

Operational strategies were those applied across algorithms deployed within organizations. These included governance of algorithms, incorporation of design principles (including accountability, explainability, interpretability, transparency, and usability) into algorithm development and maintenance, and the engagement of interdisciplinary teams and varied stakeholders.

Operational strategies reflect the need for technical expertise, as well as institutional knowledge regarding potential sources of bias. For example, in health care settings, it is often not possible or feasible to measure the outcome of interest, such as the need for care, so a proxy for that outcome, such as cost of care or number of health care visits, must be used. Given well-documented disparities in access and care, proxy variables such as cost reflect these differences and may exacerbate disparities.[36] Interdisciplinary teams and multistakeholder engagement can bring necessary perspective and debate to decisions about what outcomes to use and the potential risks of bias. For example, the inclusion of race in clinical algorithms requires careful consideration: Some studies found benefit in removing race and ethnicity from clinical algorithms,[8,27,37–39] whereas others concluded that it was preferable to include them.[27,37,40] Different algorithms have different objectives, data inputs, and intended applications, all of which can influence the decision to include or exclude race. Governance boards can help monitor these complicated decisions, and design principles such as transparency ensure that all algorithm users know when predictor variables such as race are included in algorithms, and the rationale for their conclusion.

Systemwide Strategies

Articles presenting systemwide strategies were largely authored by researchers, industry leaders, and government officials, and were often written collaboratively by authors affiliated with different institutions and oriented toward broader health and social policy issues. Systemwide strategies included updates to training and education about risks for algorithmic bias, collaborative platforms to aid organizations with algorithmic auditing, the creation of algorithm standards, increased incentives for bias mitigation, and regulation of algorithms.

Gaps In The Literature

Of the sixteen articles that presented frameworks for addressing bias in clinical algorithms, only five identified health equity as a key component, and only three tested frameworks in real-world clinical settings.[41–45]

Of the forty-five health care applications and tools tested in health care settings, nearly all addressed bias-mitigation strategies used during predeployment, such as sampling and weighting techniques and regularization methods. There was less evidence regarding clinical algorithms that had been deployed (a central component discussed in the frameworks), and only one article[26] presented a mitigation strategy applied in clinical practice. We did not identify any prospective studies. Only eighteen of forty-five health care applications included links to code used to implement the mitigation methods.

Although the articles presented numerous techniques that could be used to mitigate racial and ethnic bias in algorithms, most did not address the specific and complex questions of when and how they should be used. Only seven articles compared the performance of different techniques to mitigate racial and ethnic bias specifically.[19,21,31,34,37,39,46] We also found minimal or no information regarding the involvement of no-physician stakeholders in the design, evaluation, or deployment of or reporting on clinical algorithms.

Discussion

This scoping review included 109 articles describing empirical applications, frameworks, reviews and perspectives, and tools related to mitigating racial and ethnic bias in algorithms used to guide health care decisions. Although other related reviews have been conducted,[6,10] we believe that this review is the largest, most comprehensive summary of bias-mitigation strategies and methods in the academic and grey literatures to date.

The bias-mitigation approaches we reviewed tended to be either highly specific technical guidance or high-level, nontechnical surveys of strategies. This dichotomy is particularly challenging because bias in clinical algorithms, depending on mitigation category, requires solutions based on statistical expertise, social science expertise, clinical expertise, or some combination of the three, underscoring the need for a professionally diverse, appropriately trained workforce. For example, social scientists tend to conceive of bias as cognitive dispositions or inclinations in human thinking and reasoning, to be addressed during algorithmic design or preprocessing. Statisticians and data scientists, however, often consider bias as estimate errors to be programmatically addressed during the preprocessing and processing stages. Further downstream, clinicians view bias as a contributing factor to health inequities, which can be exacerbated by disparities in health care access, allocation, and outcomes. Increasing use of clinical algorithms has resulted in the need for new competencies among health professionals,[47] who can identify sources of bias across the algorithmic lifecycle and apply a health equity lens to evaluate and inform their use.[48]

In our review, we identified mitigation techniques that specifically address racial and ethnic bias. This requires special consideration, because the data used to train algorithms often reflect structural inequities in health care systems arising from racism and its interactions with social determinants of health. The topic of whether to include race and ethnicity as a predictor was an important focus. Although some study authors included race and ethnicity as predictors, others opted to remove them. In most studies, it is important to acknowledge that race and ethnicity are not biological factors but, rather, social constructs historically used to categorize and differentiate groups. Therefore, the decision to include or exclude race or ethnicity in a clinical model must be considered carefully in every context, and its impact on health equity should be thoroughly examined.[49] When feasible, relying on more direct measures, rather than proxies, of the outcomes being studied can prevent biases associated with the use of race and ethnicity as predictors.

Our review highlights the numerous choices required during the process of mitigating bias in clinical algorithms, encompassing the choice of variables to include as predictors, the selection of fairness metrics used to measure bias, and the choice of mitigation method. Even quantifying algorithmic bias (often measured using fairness metrics) can be difficult. Fairness metrics involve inherent trade-offs, and static fairness criteria may lead to delayed long-term harm.

Finally, mitigation methods are numerous, and few studies address the selection of appropriate methodologies, which depends on factors including the type of clinical algorithm, the specific clinical or research question being addressed, data availability, and ethical or legal considerations. Continued evaluation of these choices is needed to help determine in what clinical scenarios one strategy may work better than another.

Researchers, algorithm developers, and clinicians should follow procedures, document their results, and adhere to recognized standards for bias mitigation in clinical algorithms. A trustworthy organizational culture encourages clinicians and developers to prevent bias and reduce inequities by reporting discrimination so that root cause analysis can be performed and identified risks can be removed from the system. This not only helps ensure that algorithms are safe, fair, and transparent but also provides needed evidence on what strategies are effective in practice.

Health equity should be fundamental to designing, evaluating, and deploying clinical algorithms in real-world clinical settings to ensure that their use does not result in discrimination. One promising approach, recently articulated by the Office of the National Coordinator for Health Information Technology, is health equity by design.[50,51] Analogous to the concept of quality by design that undergirds the principles of good clinical and regulatory practice,[52] health equity by design is a multifaceted approach in which equity is a core feature that can be used to deliberately mitigate discriminatory effects of clinical algorithms across pre- and postprocessing stages. In this scoping review, we found few examples of such principles operationalized in practice, and we hope that this will be a focus for organizations moving forward.

Policy Implications

On the basis of our findings, we offer the following concrete recommendations for bias mitigation in clinical algorithms.

Ensure Professional Diversity

Developers who build algorithms and health care organizations that deploy and monitor them should cultivate and sustain professionally diverse, appropriately trained workforces that comprise the different areas of expertise (clinical, social science, technical) needed to identify and mitigate bias.

Require Auditable Clinical Algorithms

Developers and end users of algorithms must provide clear and accurate information about the intended use and risks of clinical algorithms.

Foster Transparent Organizational Culture

Developers, health care organizations, and journals should openly disclose limitations as part of communicating algorithmic outputs, acknowledge biases, and report the results of mitigation strategies.

Implement Health Equity By Design

Developers should incorporate principles of health equity by design throughout the development and deployment of clinical algorithms to mitigate discriminatory effects across pre- and postprocessing stages.

Accelerate Research

Funding agencies and health care organizations should increase research efforts focused on bias mitigation methodologies to expand the empirical evidence base informing choices regarding mitigation strategies.

Establish Governance Structures

Policy makers should encourage covered entities to establish governance structures and evaluation schemes to prevent harms arising from algorithmic bias.

Amplify Patient Voices

Developers must engage diverse patients and local communities in the design and preprocessing of clinical algorithms to inform patient-centered and culturally affirming care.

In the context of HHS's proposed rule, these recommendations provide the path forward for researchers, developers, health professionals, and policy makers to identify and mitigate harmful bias and prevent discrimination amid increasing use of clinical algorithms in health care decision making.

Many of the articles in our review include specific examples of patient harms, however inadvertent, that can result from bias in algorithms designed to support clinical decision making.[26] Given the risks posed by the rapid implementation of algorithms in health care, regulatory measures aimed at ensuring that patients do not face discrimination as a result of clinical algorithms are appropriate. However, despite evidence of the potential for harm from biased algorithms, our review demonstrates that bias mitigation remains a nascent field of research: Most of the articles we reviewed were published in the last few years. Furthermore, we observed wide variation among published mitigation strategies meant to be applied at various stages across the entire algorithm development life cycle. Some are as general as recommendations to assemble diverse team members for algorithm development, whereas others are as specific as a weighting methodology that includes published code for replication. Given the real risks posed by algorithms being rapidly deployed in health care, we believe HHS is right to revise Section 1557 of the Patient Protection and Affordable Care Act to ensure that patients do not face discrimination because of clinical algorithms.

Our review demonstrates that researchers, algorithm developers, and health care professionals have a broad range of available methods to identify and mitigate algorithmic bias. At the same time, there is no single or discrete set of approaches that HHS could require covered entities to use to eliminate algorithmic bias. Therefore, more research is needed to determine which bias mitigation methods are optimal and in what scenarios, depending on factors such as patient population, clinical setting, algorithm design, and types of bias to be addressed. Regulators and policy makers should promote sharing of resources for identifying and mitigating bias and support further research to generate empirical evidence to inform the selection of effective mitigation strategies. Policy makers should also encourage covered entities to draw from the principles and approaches presented in the literature we reviewed to establish governance structures and evaluation regimens to ensure that the algorithms they deploy do not harm patients.

Conclusion

Our scoping review serves as a significant response to the HHS notice of proposed rulemaking. It provides stakeholders, including developers, health professionals, and policy makers, with a comprehensive, up-to-date analysis of strategies, resources, and recommendations for mitigating harmful racial and ethnic bias in algorithms used in clinical decision-making. By conducting an extensive and wide-ranging review, we have gathered valuable insights across various technical, operational, and systemwide strategies. Our review encompasses a wide range of approaches and interventions that address racial and ethnic bias specifically, but can be applied more generally to prevent algorithmic discrimination.

Acknowledgment

Research reported in this publication was supported by the National Center For Advancing Translational Sciences of the National Institutes of Health (Award No. UL1TR002553). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors acknowledge the following institutions for their support: Duke University School of Nursing, Duke University Clinical and Translational Sciences Institute, Northwestern University Feinberg School of Medicine (National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR grant number: 90ARHF0003), University of California, Berkeley, and the University of Chicago Booth School of Business and Center for Applied Artificial Intelligence. The authors acknowledge the following individuals for their contributions to this project: Jonathan McCall, Donnalee Frega, Judith Hays, Seanna Horan, Shelley Rusincovitch, and Margaret Graton. To access the authors' disclosures, click on the Details tab of the article online.

BIOS for 2023–00553 (Cary)

Bio1: Michael Paul Cary, Jr. (michael.cary@duke.edu), Duke University, Durham, North Carolina.

Bio2: Anna Zink, University of Chicago, Chicago, Illinois.

Bio3: Sijia Wei, Northwestern University, Chicago, Illinois.

Bio4: Andrew Olson, Duke University.

Bio5: Mengying Yan, Duke University.

Bio6: Rashaud Senior, Duke University.

Bio7: Sophia Bessias, Duke University.

Bio8: Kais Gadhoumi, Duke University.

Bio9: Genevieve Jean-Pierre, Duke University.

Bio10: Demy Wang, Duke University.

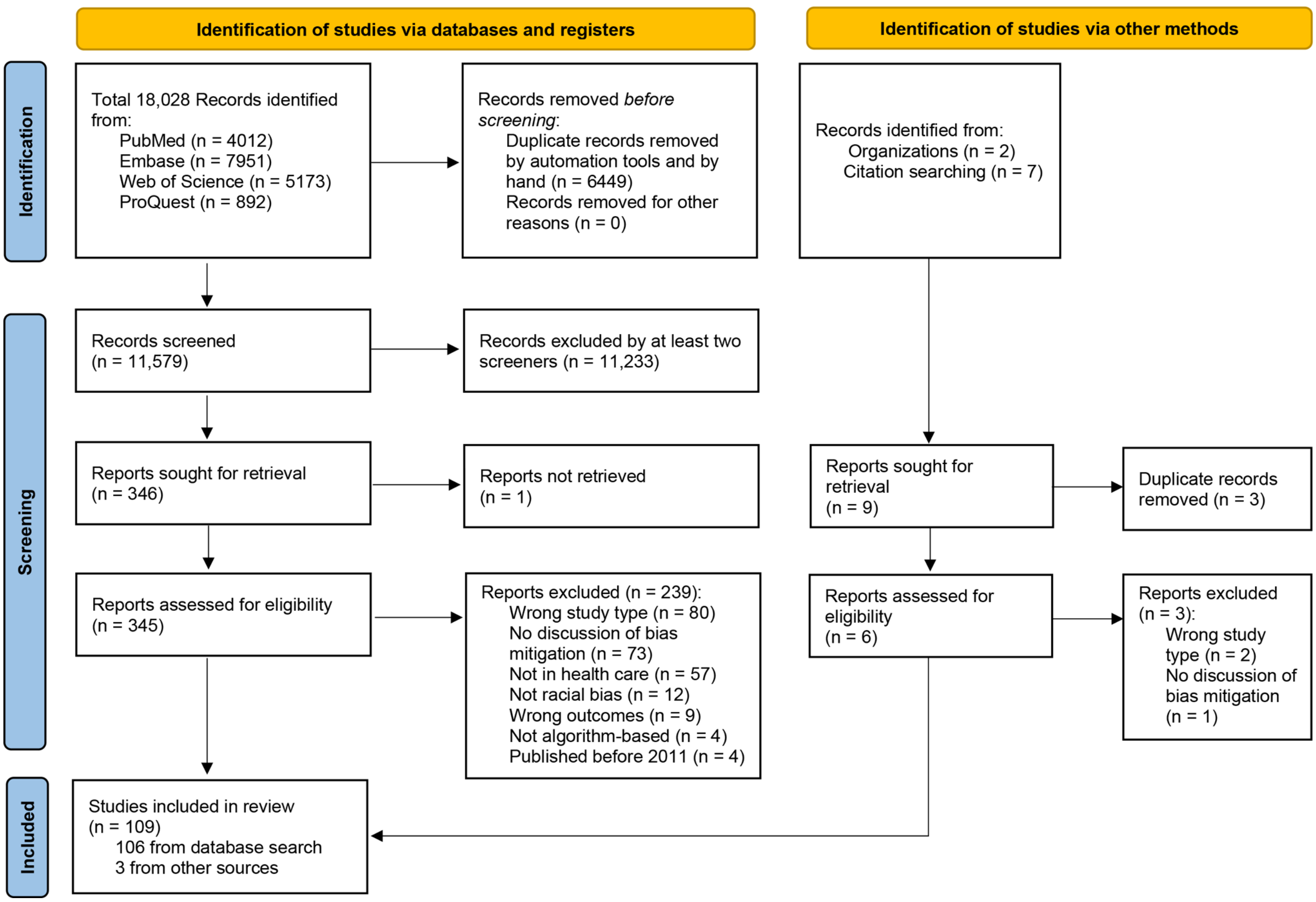
Bio11: Leila S. Ledbetter, Duke University.

Bio12: Nicoleta J. Economou-Zavlanos, Duke University.

Bio13: Ziad Obermeyer, University of California Berkeley, Berkeley, California.

Bio14: Michael J. Pencina, Duke University.

Appendix Exhibit A1.: PRISMA Diagram of the Literature Search and Selection Process



Initial Search Results

Database Search

- 18,028 references imported for screening
  - 6446 duplicates removed
- 11,582 studies screened against title and abstract
  - 11,233 studies excluded
- 1 could not retrieve
- 345 studies assessed for full-text eligibility

Other Sources

- 9 from other sources
  - 3 duplicates removed
- 6 retrieved and full text reviewed
  - 3 studies excluded
  - 2 wrong study type
  - 1 no discussion of bias mitigation

Initial Search Results	
Database Search	**Other Sources**
○ 239 studies excluded	• 3 included into the final review
■ 73 no discussion of bias mitigation	
■ 62 not reviews, guidelines, position statements, or commentaries	
■ 48 not related to healthcare	
■ 15 wrong study design	
■ 12 not specific to racial bias	
■ 9 wrong outcomes	
■ 9 wrong setting	
■ 4 not algorithm-based	
■ 4 published prior to 2011	
• Abstract	
• Not peer reviewed	
○ 106 studies included	

COMBINED TOTAL = 109

Original Exported from Covidence
• 18,037 references imported for screening as 18,037 studies
○ 6449 duplicates removed
• 11,588 studies screened against title and abstract
○ 11233 studies excluded
• 355 studies assessed for full-text eligibility
○ 246 studies excluded
■ 74 no discussion of bias mitigation
■ 66 not reviews, guidelines, position statements, or commentaries
■ 48 not related to healthcare
■ 17 wrong study design
■ 12 not specific to racial bias
■ 9 wrong outcomes
■ 9 wrong setting
■ 4 not algorithm-based
■ 4 published prior to 2011
■ 1 abstract
■ 1 not peer reviewed
■ 1 could not retrieve
○ 0 studies ongoing
○ 0 studies awaiting classification
• 109 studies included

Reasons for Exclusion	
Original Out of Covidence	**Combined**
• 74 no discussion of bias mitigation	• wrong study type (n = [66 + 15 +1 +1 =] 83)
• 66 not reviews, guidelines, position statements, or commentaries	• no discussion of bias mitigation (n = 74)
	• not related to health care (n = 48 + 9 = 57)
• 48 not related to healthcare	• not specific to racial bias (n = 12)
• 15 wrong study design	• wrong outcomes (n = 9)
• 12 not specific to racial bias	• not algorithm-based (n = 4)
• 9 wrong outcomes	• published prior to 2011 (n = 4)
• 9 wrong setting	• not retrievable (n = 1)
• 4 not algorithm-based	
• 4 published prior to 2011	
• 1 abstract	
• 1 not peer reviewed	
• 1 not retrievable	

Appendix Exhibit A2.: Search Strategies

Mitigating Bias and Advancing Health Equity Throughout the Development, Evaluation, and Deployment of Clinical Algorithms in Healthcare: A Scoping Review and Implications for Policy

Librarian Searcher: Leila Ledbetter

Peer-Reviewer: Steph Hendren

Date of completed search: August 24, 2022

Date of updated search: N/A

Total number of articles (before de-duplication): 18,028

Total number of articles (after de-duplication): 11,647

Database / Study Registry (including vendor/platform): Medline (via PubMed)

Set #	Search Strategy	Results
1 *Algorithms*	"Algorithms"[Mesh] OR "Artificial Intelligence"[Mesh:NoExp] OR "Machine Learning"[Mesh] OR "Deep Learning"[Mesh] OR Algorithm[tiab] OR Algorithms[tiab] OR Algorithmic[tiab] OR "Artificial Intelligence"[tiab] OR "Machine Learning"[tiab] OR "neural network"[tiab] OR "neural networks"[tiab] OR "deep learning"[tiab] OR "augmented intelligence"[tiab] OR "deep metric learning"[tiab] OR "Predictive model"[tiab] OR "predictive models"[tiab] OR "prediction model"[tiab] OR "prediction models"[tiab] OR AI[tiab] OR "risk score"[tiab] OR "risk scores"[tiab] OR "risk model"[tiab] OR "risk models"[tiab]	752,687

Set #	Search Strategy	Results
2 Bias	“Bias”[Mesh] OR “Health Inequities”[Mesh] OR “Healthcare Disparities”[Mesh] OR “Health Status Disparities”[Mesh] OR “Race Factors”[Mesh] OR “Systemic Racism”[Mesh] OR “Racism”[Mesh] OR Biases[tiab] OR bias[tiab] OR biased[tiab] OR debias[tiab] OR debiased[tiab] OR “de-bias”[tiab] OR “de-biased”[tiab] OR racism[tiab] or racial[tiab] or racist[tiab] OR prejudice[tiab] OR prejudices[tiab] OR prejudiced[tiab] OR discriminate[tiab] OR discriminating[tiab] OR discriminates[tiab] OR discriminatory[tiab] OR discriminated[tiab] OR discrimination[tiab] OR stereotype[tiab] OR stereotypes[tiab] OR stereotyped[tiab] OR stereotyping[tiab] OR stereotypical[tiab] OR profiling[tiab] OR disparity[tiab] OR disparities[tiab] OR inequities[tiab] OR inequity[tiab] OR inequality[tiab] OR inequalities[tiab] OR equities[tiab] OR equity[tiab] OR equality[tiab] OR equalities[tiab] OR ethics[tiab] OR ethical[tiab] OR ethically[tiab] OR unethical[tiab] OR unethically [tiab] OR accountable[tiab] OR accountability[tiab]	1,079,154
3 healthcare	“Insurance, Health”[Mesh] OR “hospitals”[MeSH Terms] OR “Risk Assessment”[Mesh] OR “Risk Management”[Mesh] OR “Disease Management”[Mesh] OR “Medical Records Systems, Computerized”[Mesh] OR “Electronic Health Records”[Mesh] OR “Decision Support Systems, Clinical”[Mesh] OR “Delivery of Health Care”[Mesh] OR “Health Services”[Mesh] OR “Medicare”[Mesh] OR “Medicaid”[Mesh] OR Health[tiab] OR “health care”[tiab] OR healthcare[tiab] OR hospitals[tiab] OR hospital[tiab] OR “medical center”[tiab] OR “medical centers”[tiab] OR “Health systems”[tiab] OR “Health system”[tiab] OR “health-care”[tiab] OR medical[tiab] OR clinical[tiab] OR risk[tiab] OR “care management”[tiab] OR “Health Services”[tiab] OR “disease management”[tiab] OR “disease prevention”[tiab] OR “health insurance”[tiab] OR “Medicare”[tiab] OR Medicaid[tiab] OR uninsured[tiab] OR insured[tiab] OR ((electronic[tiab] OR computerized[tiab] OR automated[tiab] OR administrative[tiab]) AND (medical[tiab] or health[tiab]) AND (record[tiab] OR records[tiab]))	10,437,508
4 Mitigation/ Assessment/ tools/ frameworks	“Checklist”[Mesh] OR mitigate[tiab] OR mitigates[tiab] OR mitigated[tiab] OR mitigating[tiab] OR mitigation[tiab] OR reduce[tiab] OR reduction[tiab] OR reduced[tiab] OR reduces[tiab] OR reducing[tiab] OR diminish[tiab] OR diminished[tiab] OR diminishing[tiab] OR diminishes[tiab] OR diminishment[tiab] OR alleviate[tiab] OR alleviates[tiab] OR alleviated[tiab] OR alleviating[tiab] OR amend[tiab] OR correct[tiab] OR corrects[tiab] OR corrected[tiab] OR correcting[tiab] OR correction[tiab] OR rectify[tiab] OR rectifies[tiab] OR rectified[tiab] OR rectifying[tiab] OR rectification[tiab] OR reform[tiab] OR reforms[tiab] OR reformed[tiab] OR reforming[tiab] OR reformation[tiab] OR ameliorate[tiab] OR strategy[tiab] OR strategies[tiab] OR prevent[tiab] OR prevents[tiab] OR preventing[tiab] OR prevention[tiab] OR ensure[tiab] OR ensures[tiab] OR ensuring[tiab] OR ensured[tiab] OR detect[tiab] OR detects[tiab] OR detected[tiab] OR detecting[tiab] OR detection[tiab] OR limit[tiab] OR limits[tiab] OR limited[tiab] OR limiting[tiab] OR limitation[tiab] OR limitations[tiab] OR avoid[tiab] OR avoids[tiab] OR avoided[tiab] OR avoiding[tiab] OR remove[tiab] OR removes[tiab] OR removed[tiab] OR removing[tiab] OR removal[tiab] OR decrease[tiab] OR decreases[tiab] OR decreased[tiab] OR decreasing[tiab] OR address[tiab] OR addresses[tiab] OR addressed[tiab] OR addressing[tiab] OR combat[tiab] OR combats[tiab] OR combated[tiab] OR combatted[tiab] OR combatting[tiab] OR eliminate[tiab] OR eliminates[tiab] OR eliminated[tiab] OR eliminating[tiab] OR elimination[tiab] OR minimize[tiab] OR minimizes[tiab] OR minimized[tiab] OR minimizing[tiab] OR minimise[tiab] OR minimises[tiab] OR minimised[tiab] OR minimising[tiab] OR Assess[tiab] OR assesses[tiab] OR assessed[tiab] OR assessing[tiab] OR assessment[tiab] OR evaluate[tiab] OR evaluates[tiab] OR evaluated[tiab] OR evaluating[tiab] OR evaluation[tiab] OR framework[tiab] OR frameworks[tiab] OR validate[tiab] OR validates[tiab] OR validated[tiab] OR validating[tiab] OR validation[tiab] OR measure[tiab] OR measures[tiab] OR measured[tiab] OR measuring[tiab] OR measurement[tiab] OR monitor[tiab] OR monitors[tiab] OR monitored[tiab] OR monitoring[tiab] OR analyze[tiab] OR analyzes[tiab] OR analyzed[tiab] OR analyzing[tiab] OR analysis[tiab] OR analyse[tiab] OR analyses[tiab] OR analysed[tiab] OR analysing[tiab] OR screen[tiab] OR screens[tiab] OR screened[tiab] OR screening[tiab] OR report[tiab] OR reports[tiab] OR reported[tiab] OR reporting[tiab] OR compare[tiab] OR compares[tiab] OR compared[tiab] OR comparing[tiab] OR comparison[tiab] OR diagnose[tiab] OR diagnoses[tiab] OR diagnosed[tiab] OR diagnosing[tiab] OR diagnostic[tiab] OR diagnosis[tiab] OR audit[tiab] OR audits[tiab] OR auditing[tiab] OR audited[tiab] OR comply[tiab] OR complies[tiab] OR complied[tiab] OR complying[tiab] OR compliance[tiab] OR validate[tiab] OR validates[tiab] OR validated[tiab] OR validating[tiab] OR validation[tiab] OR “algorithmic hygiene”[tiab] OR “algorithmic playbook”[tiab] OR “algorithmic bias playbook”[tiab] OR checklist[tiab] OR checklists[tiab] OR tool[tiab] OR tools[tiab] OR toolkit[tiab] OR toolkits[tiab] OR guideline[tiab] OR guidelines[tiab] OR rubric[tiab] OR rubrics[tiab] OR guidance[tiab] OR standard[tiab] OR standards[tiab] OR method[tiab] OR methods[tiab] OR methodology[tiab] OR methodologies[tiab] OR checkpoint[tiab] OR checkpoints[tiab] OR recommendation[tiab] OR recommendations[tiab] OR	22,393,661

Set #	Search Strategy	Results
	benchmark[tiab] OR benchmarks[tiab] OR regulation[tiab] OR regulations[tiab] OR regulatory[tiab] OR generalizable[tiab] OR generalizability[tiab] OR “prejudice remover”[tiab] OR fair[tiab] OR unfair[tiab] OR fairness[tiab] OR unfairness[tiab] OR “classification parity”[tiab] OR trust[tiab] OR trustworthy[tiab] OR trustworthiness[tiab] OR PROBAST[tiab] OR “Prediction Model Risk Of Bias Assessment Tool”[tiab] OR “Minimum Information for Medical AI Reporting guideline”[tiab] OR “Artificial Intelligence Risk Management Framework”[tiab] OR “AI RMF”[tiab] OR “Medical Information Mart for Intensive Care”[tiab] OR MIMIC[tiab] OR “balanced accuracy”[tiab] OR “disparate impact”[tiab] OR “equal opportunity”[tiab] OR “equalized odds”[tiab] OR “statistical parity”[tiab]	
5 Race/ ethnicity	“Ethnicity”[Mesh] OR “Minority Health”[Mesh] OR “Ethnic and Racial Minorities”[Mesh] OR “Blacks”[Mesh] OR “African Americans”[Mesh] OR “Mexican Americans”[Mesh] OR “Hispanic or Latino”[Mesh] OR “Indians, North American”[Mesh] OR “American Indians or Alaska Natives”[Mesh] OR “Indigenous Peoples”[Mesh] OR “Asian Americans”[Mesh] OR “Native Hawaiian or Other Pacific Islander”[Mesh] OR “Transients and Migrants”[Mesh] OR “Emigrants and Immigrants”[Mesh] OR “minority group”[tiab] OR ethnic[tiab] OR ethnicity[tiab] OR “minority groups”[tiab] OR “minority population”[tiab] OR “minority populations”[tiab] OR “people of color”[tiab] OR “person of color”[tiab] OR BIPOC[tiab] OR “minority health”[tiab] OR race[tiab] OR races[tiab] OR racial[tiab] OR racially[tiab] OR “African Americans”[tiab] OR “African ancestry”[tiab] OR Black[tiab] OR Blacks[tiab] OR “African American”[tiab] OR “Hispanic American”[tiab] OR “Hispanic-American”[tiab] OR “Hispanic Americans”[tiab] OR “Hispanic-Americans”[tiab] OR Hispanics[tiab] OR Hispanic[tiab] OR “Mexican American”[tiab] OR “Mexican-American”[tiab] OR “Mexican Americans”[tiab] OR “Mexican-Americans”[tiab] OR “Cuban American”[tiab] OR “Cuban-American”[tiab] OR “Cuban Americans”[tiab] OR “Cuban-Americans”[tiab] OR “Latin American”[tiab] OR “Latin-American”[tiab] OR “Latin Americans”[tiab] OR “Latin-Americans”[tiab] OR Latinos[tiab] OR Latino[tiab] OR Latinas[tiab] OR Latina[tiab] OR Latinx[tiab] OR latine[tiab] OR latines[tiab] OR “Spanish speaking”[tiab] OR “Spanish-speaking”[tiab] OR “Spanish speakers”[tiab] OR “Spanish-speakers”[tiab] OR Mexican[tiab] OR Mexicans[tiab] OR “Puerto Rican”[tiab] OR “Puerto Ricans”[tiab] OR “Asian American”[tiab] OR “Asian Americans”[tiab] OR Asian[tiab] OR “Japanese Americans”[tiab] OR “Japanese American”[tiab] OR “Chinese Americans”[tiab] OR “Chinese American”[tiab] OR “Vietnamese Americans”[tiab] OR “Vietnamese American”[tiab] OR “Asian Indian Americans”[tiab] OR “Asian Indian American”[tiab] OR “Cambodian Americans”[tiab] OR “Cambodian American”[tiab] OR “Hmong Americans”[tiab] OR “Hmong American”[tiab] OR “Korean Americans”[tiab] OR “Korean American”[tiab] OR “Filipino Americans”[tiab] OR “Filipino American”[tiab] OR “Indigenous People”[tiab] OR “Indigenous People”[tiab] OR “Indigenous Population”[tiab] OR “Indigenous Populations”[tiab] OR tribes[tiab] OR “Native Peoples”[tiab] OR “Native People”[tiab] OR “First Nation”[tiab] OR “native person”[tiab] OR Alaskan[tiab] OR Hawaiian[tiab] OR Alaskans[tiab] OR Hawaiians[tiab] OR “Caribbean American”[tiab] OR “Caribbean-American”[tiab] OR “Caribbean Americans”[tiab] OR “Caribbean-Americans”[tiab] OR migrant[tiab] OR migrants[tiab] OR immigrant[tiab] OR immigrants[tiab] OR emigrant[tiab] OR emigrants[tiab] OR demographic[tiab] OR demographics[tiab] OR Polynesians[tiab] OR Polynesian[tiab] OR “pacific islander”[tiab] OR “pacific islanders”[tiab]	1,056,938
6	#1 AND #2 AND #3 AND #4 AND #5	4,024
7 Remove animal studies	#6 NOT (animals[MeSH Terms] NOT humans[MeSH Terms])	4,012

2nd Database: Embase (via Elsevier)

Set #	Search Strategy	Results
1 Algorithms	‘Algorithm’/de OR ‘Algorithm bias’/de OR ‘Artificial Intelligence’/de OR ‘automated reasoning’/de OR ‘Machine Learning’/exp OR ‘Deep Learning’/de OR ‘predictive model’/de OR ‘risk model’/de OR Algorithm:ti,ab OR Algorithms:ti,ab OR Algorithmic:ti,ab OR ‘Artificial Intelligence’:ti,ab OR ‘Machine Learning’:ti,ab OR ‘deep learning’:ti,ab OR ‘augmented intelligence’:ti,ab OR ‘deep metric learning’:ti,ab OR ‘Predictive model’:ti,ab OR ‘predictive models’:ti,ab OR ‘prediction model’:ti,ab	921,128

Set #	Search Strategy	Results
	OR ‘prediction models’:ti,ab OR AI:ti,ab OR ‘risk score’:ti,ab OR ‘risk scores’:ti,ab OR ‘risk model’:ti,ab OR ‘risk models’:ti,ab	
2 Bias	‘Prejudice’/de OR ‘Health Disparity’/de OR ‘Race’/de OR ‘Racism’/exp OR ‘Implicit Bias’/de OR Biases:ti,ab OR bias:ti,ab OR biased:ti,ab OR debias:ti,ab OR debiased:ti,ab OR ‘de-bias’:ti,ab OR ‘de-biased’:ti,ab OR racism:ti,ab OR racial:ti,ab OR racist:ti,ab OR prejudice:ti,ab OR prejudices:ti,ab OR prejudiced:ti,ab OR discriminate:ti,ab OR discriminating:ti,ab OR discriminates:ti,ab OR discriminatory:ti,ab OR discriminated:ti,ab OR discrimination:ti,ab OR stereotype:ti,ab OR stereotypes:ti,ab OR stereotyped:ti,ab OR stereotyping:ti,ab OR stereotypical:ti,ab OR profiling:ti,ab OR disparity:ti,ab OR disparities:ti,ab OR inequities:ti,ab OR inequity:ti,ab OR inequality:ti,ab OR inequalities:ti,ab OR equities:ti,ab OR equity:ti,ab OR equality:ti,ab OR equalities:ti,ab OR ethics:ti,ab OR ethical:ti,ab OR ethically:ti,ab OR unethical:ti,ab OR unethically:ti,ab OR accountable:ti,ab OR accountability:ti,ab	1,349,533
3 healthcare	‘Health Insurance’/de OR hospital/exp OR ‘Risk Assessment’/exp OR ‘Risk Management’/de OR ‘Disease Management’/de OR ‘Electronic Medical Records System’/de OR ‘Health Care delivery’/de OR ‘Health Services’/de OR ‘Health care’/de OR Medicare/de OR Medicaid/de OR ‘electronic health record’/exp OR Health:ti,ab OR ‘health care’:ti,ab OR healthcare:ti,ab OR hospitals:ti,ab OR hospital:ti,ab OR ‘medical center’:ti,ab OR ‘medical centers’:ti,ab OR ‘Health systems’:ti,ab OR ‘Health system’:ti,ab OR health-care:ti,ab OR medical:ti,ab OR clinical:ti,ab OR risk:ti,ab OR ‘care management’:ti,ab OR ‘Health Services’:ti,ab OR ‘disease management’:ti,ab OR ‘disease prevention’:ti,ab OR ‘health insurance’:ti,ab OR Medicare:ti,ab OR Medicaid:ti,ab OR uninsured:ti,ab OR insured:ti,ab OR ‘electronic health records’:ti,ab OR ((electronic OR computerized OR automated OR administrative) NEAR/2 (medical OR health) NEAR/2 (records OR records)):ti,ab	13,131,397
4 Mitigation/ Assessment	‘checklist’/exp OR mitigate:ti,ab OR mitigates:ti,ab OR mitigated:ti,ab OR mitigating:ti,ab OR mitigation:ti,ab OR reduce:ti,ab OR reduction:ti,ab OR reduced:ti,ab OR reduces:ti,ab OR reducing:ti,ab OR diminish:ti,ab OR diminished:ti,ab OR diminishing:ti,ab OR diminishes:ti,ab OR diminishment:ti,ab OR alleviate:ti,ab OR alleviates:ti,ab OR alleviated:ti,ab OR alleviating:ti,ab OR amend:ti,ab OR amends:ti,ab OR amended:ti,ab OR amending:ti,ab OR correct:ti,ab OR corrects:ti,ab OR corrected:ti,ab OR correcting:ti,ab OR correction:ti,ab OR rectify:ti,ab OR rectifies:ti,ab OR rectified:ti,ab OR rectifying:ti,ab OR rectification:ti,ab OR reform:ti,ab OR reforms:ti,ab OR reformed:ti,ab OR reforming:ti,ab OR reformation:ti,ab OR ameliorate:ti,ab OR strategy:ti,ab OR strategies:ti,ab OR prevent:ti,ab OR prevents:ti,ab OR preventing:ti,ab OR prevention:ti,ab OR ensure:ti,ab OR ensures:ti,ab OR ensuring:ti,ab OR ensured:ti,ab OR detect:ti,ab OR detects:ti,ab OR detected:ti,ab OR detecting:ti,ab OR detection:ti,ab OR limit:ti,ab OR limits:ti,ab OR limited:ti,ab OR limiting:ti,ab OR limitation:ti,ab OR limitations:ti,ab OR avoid:ti,ab OR avoids:ti,ab OR avoided:ti,ab OR avoiding:ti,ab OR remove:ti,ab OR removes:ti,ab OR removed:ti,ab OR removing:ti,ab OR removal:ti,ab OR decrease:ti,ab OR decreases:ti,ab OR decreased:ti,ab OR decreasing:ti,ab OR address:ti,ab OR addresses:ti,ab OR addressed:ti,ab OR addressing:ti,ab OR combat:ti,ab OR combats:ti,ab OR combated:ti,ab OR combatted:ti,ab OR combatting:ti,ab OR eliminate:ti,ab OR eliminates:ti,ab OR eliminated:ti,ab OR eliminating:ti,ab OR elimination:ti,ab OR minimize:ti,ab OR minimizes:ti,ab OR minimized:ti,ab OR minimizing:ti,ab OR minimise:ti,ab OR minimises:ti,ab OR minimised:ti,ab OR minimising:ti,ab OR Assess:ti,ab OR assesses:ti,ab OR assessed:ti,ab OR assessing:ti,ab OR assessment:ti,ab OR evaluate:ti,ab OR evaluates:ti,ab OR evaluated:ti,ab OR evaluating:ti,ab OR evaluation:ti,ab OR framework:ti,ab OR frameworks:ti,ab OR validate:ti,ab OR validates:ti,ab OR validated:ti,ab OR validating:ti,ab OR validation:ti,ab OR measure:ti,ab OR measures:ti,ab OR measured:ti,ab OR measuring:ti,ab OR measurement:ti,ab OR monitor:ti,ab OR monitors:ti,ab OR monitored:ti,ab OR monitoring:ti,ab OR analyze:ti,ab OR analyzes:ti,ab OR analyzed:ti,ab OR analyzing:ti,ab OR analysis:ti,ab OR analyse:ti,ab OR analyses:ti,ab OR analysed:ti,ab OR analysing:ti,ab OR screen:ti,ab OR screens:ti,ab OR screened:ti,ab OR screening:ti,ab OR report:ti,ab OR reports:ti,ab OR reported:ti,ab OR reporting:ti,ab OR compare:ti,ab OR compares:ti,ab OR compared:ti,ab OR comparing:ti,ab OR comparison:ti,ab OR diagnose:ti,ab OR diagnoses:ti,ab OR diagnosed:ti,ab OR diagnosing:ti,ab OR diagnostic:ti,ab OR diagnosis:ti,ab OR audit:ti,ab OR audits:ti,ab OR auditing:ti,ab OR audited:ti,ab OR comply:ti,ab OR complies:ti,ab OR complied:ti,ab OR complying:ti,ab OR compliance:ti,ab OR validate:ti,ab OR validates:ti,ab OR validated:ti,ab OR validating:ti,ab OR validation:ti,ab OR ‘algorithmic hygiene’:ti,ab OR ‘algorithmic playbook’:ti,ab OR ‘algorithmic bias playbook’:ti,ab OR checklist:ti,ab OR checklists:ti,ab OR tool:ti,ab OR tools:ti,ab OR toolkit:ti,ab OR toolkits:ti,ab OR guideline:ti,ab OR guidelines:ti,ab OR rubric:ti,ab OR rubrics:ti,ab OR guidance:ti,ab OR standard:ti,ab OR standards:ti,ab OR method:ti,ab OR methods:ti,ab OR methodology:ti,ab OR methodologies:ti,ab OR checkpoint:ti,ab OR checkpoints:ti,ab OR recommendation:ti,ab OR recommendations:ti,ab OR benchmark:ti,ab OR benchmarks:ti,ab OR regulation:ti,ab OR regulations:ti,ab	28,945,812

Set #	Search Strategy	Results
	OR regulatory:ti,ab OR generalizable:ti,ab OR generalizability:ti,ab OR 'prejudice remover':ti,ab OR fair:ti,ab OR unfair:ti,ab OR fairness:ti,ab OR unfairness:ti,ab OR 'classification parity':ti,ab OR trust:ti,ab OR trustworthy:ti,ab OR trustworthiness:ti,ab OR PROBAST:ti,ab OR 'Prediction Model Risk Of Bias Assessment Tool':ti,ab OR 'Minimum Information for Medical AI Reporting guideline':ti,ab OR 'Artificial Intelligence Risk Management Framework':ti,ab OR 'AI RMF':ti,ab OR 'Medical Information Mart for Intensive Care':ti,ab OR MIMIC:ti,ab OR 'balanced accuracy':ti,ab OR 'disparate impact':ti,ab OR 'equal opportunity':ti,ab OR 'equalized odds':ti,ab OR 'statistical parity':ti,ab	
5 Race/ ethnicity	Ethnicity/de OR 'Minority Health'/de OR 'Ethnic group'/de OR 'minority group'/de OR 'Black person'/exp OR 'African Americans'/de OR 'Mexican Americans'/de OR 'Hispanic'/exp OR 'Puerto Rican'/de OR 'Asian Americans'/de OR 'American Indian'/de OR 'Alaska Native'/de OR 'Indigenous People'/de OR 'First Nation'/de OR 'Migrant'/exp OR 'demographics'/exp OR 'Polynesian'/exp OR 'Pacific Islander'/exp OR 'minority group':ti,ab OR ethnic:ti,ab OR ethnicity:ti,ab OR 'minority groups':ti,ab OR 'minority population':ti,ab OR 'minority populations':ti,ab OR 'people of color':ti,ab OR 'person of color':ti,ab OR BIPOC:ti,ab OR 'minority health':ti,ab OR race:ti,ab OR races:ti,ab OR racial:ti,ab OR racially:ti,ab OR 'African Americans':ti,ab OR 'African ancestry':ti,ab OR Black:ti,ab OR Blacks:ti,ab OR 'African American':ti,ab OR 'Hispanic American':ti,ab OR Hispanic-American:ti,ab OR 'Hispanic Americans':ti,ab OR Hispanic-Americans:ti,ab OR Hispanics:ti,ab OR Hispanic:ti,ab OR 'Mexican American':ti,ab OR Mexican-American:ti,ab OR 'Mexican Americans':ti,ab OR Mexican-Americans:ti,ab OR 'Cuban American':ti,ab OR Cuban-American:ti,ab OR 'Cuban Americans':ti,ab OR Cuban-Americans:ti,ab OR 'Latin American':ti,ab OR Latin-American:ti,ab OR 'Latin Americans':ti,ab OR Latin-Americans:ti,ab OR Latinos:ti,ab OR Latino:ti,ab OR Latinas:ti,ab OR Latina:ti,ab OR Latinx:ti,ab OR latine:ti,ab OR latines:ti,ab OR 'Spanish speaking':ti,ab OR Spanish-speaking:ti,ab OR 'Spanish speakers':ti,ab OR Spanish-speakers:ti,ab OR Mexican:ti,ab OR Mexicans:ti,ab OR 'Puerto Rican':ti,ab OR 'Puerto Ricans':ti,ab OR 'Asian American':ti,ab OR 'Asian Americans':ti,ab OR Asian:ti,ab OR 'Japanese Americans':ti,ab OR 'Japanese American':ti,ab OR 'Chinese Americans':ti,ab OR 'Chinese American':ti,ab OR 'Vietnamese Americans':ti,ab OR 'Vietnamese American':ti,ab OR 'Asian Indian Americans':ti,ab OR 'Asian Indian American':ti,ab OR 'Cambodian Americans':ti,ab OR 'Cambodian American':ti,ab OR 'Hmong Americans':ti,ab OR 'Hmong American':ti,ab OR 'Korean Americans':ti,ab OR 'Korean American':ti,ab OR 'Filipino Americans':ti,ab OR 'Filipino American':ti,ab OR 'Indigenous People':ti,ab OR 'Indigenous People':ti,ab OR 'Indigenous Population':ti,ab OR 'Indigenous Populations':ti,ab OR tribes:ti,ab OR 'Native Peoples':ti,ab OR 'Native People':ti,ab OR 'First Nation':ti,ab OR 'native person':ti,ab OR Alaskan:ti,ab OR Alaskans:ti,ab OR 'Caribbean American':ti,ab OR Caribbean-American:ti,ab OR 'Caribbean Americans':ti,ab OR Caribbean-Americans:ti,ab OR migrant:ti,ab OR migrants:ti,ab OR immigrant:ti,ab OR immigrants:ti,ab OR emigrant:ti,ab OR emigrants:ti,ab OR demographic:ti,ab OR demographics:ti,ab OR Polynesians:ti,ab OR Polynesian:ti,ab OR 'pacific islander':ti,ab OR 'pacific islanders':ti,ab OR Hawaiian:ti,ab OR Hawaiians:ti,ab	1,548,451
6	#1 AND #2 AND #3 AND #4 AND #5	8,230
7	#6 AND [humans]/lim	7,951

3rd Database: Web of Science (Clarivate)

Set #	Search Strategy	Results
1 Algorithms	TS=("Artificial Intelligence" OR "Machine Learning" OR "Deep Learning" OR Algorithm OR Algorithms OR Algorithmic OR "augmented intelligence" OR "deep metric learning" OR "Predictive model" OR "predictive models" OR "prediction model" OR "prediction models" OR AI OR "risk score" OR "risk scores" OR "risk model" OR "risk models")	2,913,410
2 Bias	TS=("Health Status Disparities" OR "Race Factors" OR Biases OR bias OR biased OR debias OR debiased OR "debias" OR "de-biased" OR racism OR racial OR racist OR prejudice OR prejudices OR prejudiced OR discriminate OR discriminating OR discriminates OR discriminatory OR discriminated OR discrimination OR stereotype OR stereotypes OR stereotyped OR stereotyping OR stereotypical OR profiling OR disparity OR disparities OR inequities OR inequity OR inequality OR inequalities OR equities OR equity OR equality OR equalities OR ethics OR ethical OR ethically OR unethical OR unethically OR accountable OR accountability)	3,804,114

Set #	Search Strategy	Results
3 Healthcare	TS=("Health Insurance" OR "Risk Assessment" OR "Risk Management" OR "Disease Management" OR "Medical Records Systems" OR "Delivery of Health Care" OR "Health Services" OR Health OR "health care" OR healthcare OR hospitals OR hospital OR "medical center" OR "medical centers" OR "Health systems" OR "Health system" OR health-care OR medical OR risk OR clinical OR "care management" OR "disease prevention" OR Medicare OR Medicaid OR uninsured OR insured OR "Electronic Health Records" OR ((electronic OR computerized OR automated OR administrative) AND (medical or health) AND (record OR records)))	10,034,108
4 Mitigation/ Assessment	TS=(mitigate OR mitigates OR mitigated OR mitigating OR mitigation OR reduce OR reduction OR reduced OR reduces OR reducing OR diminish OR diminished OR diminishing OR diminishes OR diminishment OR alleviate OR alleviates OR alleviated OR alleviating OR amend OR amends OR amended OR amending OR correct OR corrects OR corrected OR correcting OR correction OR rectify OR rectifies OR rectified OR rectifying OR rectification OR reform OR reforms OR reformed OR reforming OR reformation OR ameliorate OR strategy OR strategies OR prevent OR prevents OR preventing OR prevention OR ensure OR ensures OR ensuring OR ensured OR detect OR detects OR detected OR detecting OR detection OR limit OR limits OR limited OR limiting OR limitation OR limitations OR avoid OR avoids OR avoided OR avoiding OR remove OR removes OR removed OR removing OR removal OR decrease OR decreases OR decreased OR decreasing OR address OR addresses OR addressed OR addressing OR combat OR combats OR combated OR combatted OR combatting OR eliminate OR eliminates OR eliminated OR eliminating OR elimination OR minimize OR minimizes OR minimized OR minimizing OR minimise OR minimises OR minimised OR minimising OR Assess OR assesses OR assessed OR assessing OR assessment OR evaluate OR evaluates OR evaluated OR evaluating OR evaluation OR framework OR frameworks OR validate OR validates OR validated OR validating OR validation OR measure OR measures OR measured OR measuring OR measurement OR monitor OR monitors OR monitored OR monitoring OR analyze OR analyzes OR analyzed OR analyzing OR analysis OR analyse OR analyses OR analysed OR analysing OR screen OR screens OR screened OR screening OR report OR reports OR reported OR reporting OR compare OR compares OR compared OR comparing OR comparison OR diagnose OR diagnoses OR diagnosed OR diagnosing OR diagnostic OR diagnosis OR audit OR audits OR auditing OR audited OR comply OR complies OR complied OR complying OR compliance OR validate OR validates OR validated OR validating OR validation OR Checklist OR "algorithmic hygiene" OR "algorithmic playbook" OR "algorithmic bias playbook" OR checklist OR checklists OR tool OR tools OR toolkit OR toolkits OR guideline OR guidelines OR rubric OR rubrics OR guidance OR standard OR standards OR method OR methods OR methodology OR methodologies OR checkpoint OR checkpoints OR recommendation OR recommendations OR benchmark OR benchmarks OR regulation OR regulations OR regulatory OR generalizable OR generalizability OR "prejudice remover" OR fair OR unfair OR fairness OR unfairness OR "classification parity" OR trust OR trustworthy OR trustworthiness OR PROBAST OR "Prediction Model Risk Of Bias Assessment Tool" OR "Minimum Information for Medical AI Reporting guideline" OR "Artificial Intelligence Risk Management Framework" OR "AI RMF" OR "Medical Information Mart for Intensive Care" OR MIMIC OR "balanced accuracy" OR "disparate impact" OR "equal opportunity" OR "equalized odds" OR "statistical parity")	43,307,162
5 Race	TS=("minority group" OR ethnic OR ethnicity OR "minority groups" OR "minority population" OR "minority populations" OR "people of color" OR "person of color" OR BIPOC OR "minority health" OR race OR races OR racial OR racially OR Blacks OR "African Americans" OR "African ancestry" OR Black OR "African American" OR "Mexican Americans" OR "Hispanic or Latino" OR "Hispanic American" OR Hispanic-American OR "Hispanic Americans" OR Hispanic-Americans OR Hispanics OR Hispanic OR "Mexican American" OR Mexican-American OR "Mexican Americans" OR Mexican-Americans OR "Cuban American" OR Cuban-American OR "Cuban Americans" OR Cuban-Americans OR "Latin American" OR Latin-American OR "Latin Americans" OR Latin-Americans OR Latinos OR Latino OR Latinas OR Latina OR Latinx OR latine OR latines OR "Spanish speaking" OR Spanish-speaking OR "Spanish speakers" OR Spanish-speakers OR Mexican OR Mexicans OR "Puerto Rican" OR "Puerto Ricans" OR "Asian Americans" OR "Asian American" OR Asian OR "Japanese Americans" OR "Japanese American" OR "Chinese Americans" OR "Chinese American" OR "Vietnamese Americans" OR "Vietnamese American" OR "Asian Indian Americans" OR "Asian Indian American" OR "Cambodian Americans" OR "Cambodian American" OR "Hmong Americans" OR "Hmong American" OR "Korean Americans" OR "Korean American" OR "Filipino Americans" OR "Filipino American" OR "American Indians" OR "Alaska Natives" OR "Indigenous Peoples" OR "Indigenous People" OR "Indigenous Population" OR "Indigenous Populations" OR tribes OR "Native Peoples" OR "Native People" OR "First Nation" OR "native person" OR Alaskan OR Alaskans OR "Caribbean American" OR Caribbean-American OR "Caribbean Americans" OR Caribbean-Americans OR migrant OR migrants OR immigrant OR immigrants OR emigrant OR emigrants or demographic OR	1,968,002

Set #	Search Strategy	Results
	demographics OR Polynesian OR Polynesians OR “Pacific Islander” OR “Pacific Islanders” OR Hawaiian OR Hawaiians)	
6	#1 AND #2 AND #3 AND #4 AND #5	5,173

4th Database: Proquest Computer Science Database

Set #	Search Strategy	Results
1 *Algorithms*	NOFT(“Artificial Intelligence” OR “Machine Learning” OR “Deep Learning” OR Algorithm OR Algorithms OR Algorithmic OR “Artificial Intelligence” OR “augmented intelligence” OR “deep metric learning” OR “Predictive model” OR “predictive models” OR “prediction model” OR “prediction models” OR AI OR “risk score” OR “risk scores” OR “risk model” OR “risk models”)	410,974
2 Bias	NOFT(“Health Status Disparities” OR “Race Factors” OR Biases OR bias OR biased OR debias OR debiased OR “de-bias” OR “de-biased” OR racism OR racial OR racist OR prejudice OR prejudices OR prejudiced OR discriminate OR discriminating OR discriminates OR discriminatory OR discriminated OR discrimination OR stereotype OR stereotypes OR stereotyped OR stereotyping OR stereotypical OR profiling OR disparity OR disparities OR inequities OR inequity OR inequality OR inequalities OR equities OR equity OR equality OR equalities OR ethics OR ethical OR ethically OR unethical OR unethically OR accountable OR accountability)	192,745
3 Healthcare	NOFT(“Health Insurance” OR “Risk Assessment” OR “Risk Management” OR “Disease Management” OR “Medical Records Systems” OR “Delivery of Health Care” OR “Health Services” OR Health OR “health care” OR healthcare OR hospitals OR hospital OR “medical center” OR “medical centers” OR “Health systems” OR “Health system” OR health-care OR medical OR clinical OR risk OR “care management” OR “disease prevention” OR Medicare OR Medicaid OR uninsured OR insured OR “Electronic Health Records” OR ((electronic OR computerized OR automated OR administrative) AND (medical or health) AND (record OR records)))	774,748
4 Mitigation/ assessment	NOFT(mitigate OR mitigates OR mitigated OR mitigating OR mitigation OR reduce OR reduction OR reduced OR reduces OR reducing OR diminish OR diminished OR diminishing OR diminishes OR diminishment OR alleviate OR alleviates OR alleviated OR alleviating OR amend OR amends OR amended OR amending OR correct OR corrects OR corrected OR correcting OR correction OR rectify OR rectifies OR rectified OR rectifying OR rectification OR reform OR reforms OR reformed OR reforming OR reformation OR ameliorate OR strategy OR strategies OR prevent OR prevents OR preventing OR prevention OR ensure OR ensures OR ensuring OR ensured OR detect OR detects OR detected OR detecting OR detection OR limit OR limits OR limited OR limiting OR limitation OR limitations OR avoid OR avoids OR avoided OR avoiding OR remove OR removes OR removed OR removing OR removal OR decrease OR decreases OR decreased OR decreasing OR address OR addresses OR addressed OR addressing OR combat OR combats OR combated OR combatted OR combatting OR eliminate OR eliminates OR eliminated OR eliminating OR elimination OR minimize OR minimizes OR minimized OR minimizing OR minimise OR minimises OR minimised OR minimising OR Assess OR assesses OR assessed OR assessing OR assessment OR evaluate OR evaluates OR evaluated OR evaluating OR evaluation OR framework OR frameworks OR validate OR validates OR validated OR validating OR validation OR measure OR measures OR measured OR measuring OR measurement OR monitor OR monitors OR monitored OR monitoring OR analyze OR analyzes OR analyzed OR analyzing OR analysis OR analyse OR analyses OR analysed OR analysing OR screen OR screens OR screened OR screening OR report OR reports OR reported OR reporting OR compare OR compares OR compared OR comparing OR comparison OR diagnose OR diagnoses OR diagnosed OR diagnosing OR diagnostic OR diagnosis OR audit OR audits OR auditing OR audited OR comply OR complies OR complied OR complying OR compliance OR validate OR validates OR validated OR validating OR validation OR Checklist OR “algorithmic hygiene” OR “algorithmic playbook” OR “algorithmic bias playbook” OR checklist OR checklists OR tool OR tools OR toolkit OR toolkits OR guideline OR guidelines OR rubric OR rubrics OR guidance OR standard OR standards OR method OR methods OR methodology OR methodologies OR checkpoint OR checkpoints OR recommendation OR recommendations OR benchmark OR benchmarks OR regulation OR regulations OR regulatory OR generalizable OR generalizability OR “prejudice remover” OR fair OR unfair OR fairness OR unfairness OR “classification parity” OR trust OR trustworthy OR trustworthiness OR PROBAST OR “Prediction Model Risk Of Bias Assessment Tool” OR “Minimum Information for Medical AI Reporting guideline” OR “Artificial Intelligence Risk Management Framework” OR “AI RMF” OR “Medical Information Mart for Intensive Care” OR MIMIC OR “balanced	4,758,704

Set #	Search Strategy	Results
	accuracy" OR "disparate impact" OR "equal opportunity" OR "equalized odds" OR "statistical parity")	
5 Race	NOFT("minority group" OR ethnic OR ethnicity OR "minority groups" OR "minority population" OR "minority populations" OR "people of color" OR "person of color" OR BIPOC OR "minority health" OR race OR races OR racial OR racially OR Blacks OR "African Americans" OR "African ancestry" OR Black OR "African American" OR "Mexican Americans" OR "Hispanic or Latino" OR "Hispanic American" OR Hispanic-American OR "Hispanic Americans" OR Hispanic-Americans OR Hispanics OR Hispanic OR "Mexican American" OR Mexican-American OR "Mexican Americans" OR Mexican-Americans OR "Cuban American" OR Cuban-American OR "Cuban Americans" OR Cuban-Americans OR "Latin American" OR Latin-American OR "Latin Americans" OR Latin-Americans OR Latinos OR Latino OR Latinas OR Latina OR Latinx OR latine OR latines OR "Spanish speaking" OR Spanish-speaking OR "Spanish speakers" OR Spanish-speakers OR Mexican OR Mexicans OR "Puerto Rican" OR "Puerto Ricans" OR "Asian Americans" OR "Asian American" OR Asian OR "Japanese Americans" OR "Japanese American" OR "Chinese Americans" OR "Chinese American" OR "Vietnamese Americans" OR "Vietnamese American" OR "Asian Indian Americans" OR "Asian Indian American" OR "Cambodian Americans" OR "Cambodian American" OR "Hmong Americans" OR "Hmong American" OR "Korean Americans" OR "Korean American" OR "Filipino Americans" OR "Filipino American" OR "American Indians" OR "Alaska Natives" OR "Indigenous Peoples" OR "Indigenous People" OR "Indigenous Population" OR "Indigenous Populations" OR tribes OR "Native Peoples" OR "Native People" OR "First Nation" OR "native person" OR Alaskan OR Alaskans OR "Caribbean American" OR Caribbean-American OR "Caribbean Americans" OR Caribbean-Americans OR migrant OR migrants OR immigrant OR immigrants OR emigrant OR emigrants OR demographic OR demographics OR Polynesian OR Polynesians OR "Pacific Islander" OR "Pacific Islanders" OR Hawaiian OR Hawaiians)	127,761
6	1 AND 2 AND 3 AND 4 AND 5	161

5th Database: Proquest Dissertation & Theses Global

Set #	Search Strategy	Results
1 *Algorithms*	NOFT("Artificial Intelligence" OR "Machine Learning" OR "Deep Learning" OR Algorithm OR Algorithms OR Algorithmic OR "augmented intelligence" OR "deep metric learning" OR "Predictive model" OR "predictive models" OR "prediction model" OR "prediction models" OR AI OR "risk score" OR "risk scores" OR "risk model" OR "risk models")	274,653
2 Bias	NOFT("Health Status Disparities" OR "Race Factors" OR Biases OR bias OR biased OR debias OR debiased OR "de-bias" OR "de-biased" OR racism OR racial OR racist OR prejudice OR prejudices OR prejudiced OR discriminate OR discriminating OR discriminates OR discriminatory OR discriminated OR discrimination OR stereotype OR stereotypes OR stereotyped OR stereotyping OR stereotypical OR profiling OR disparity OR disparities OR inequities OR inequity OR inequality OR inequalities OR equities OR equity OR equality OR equalities OR ethics OR ethical OR ethically OR unethical OR unethically OR accountable OR accountability)	363,230
3	NOFT("Health Insurance" OR "Risk Assessment" OR "Risk Management" OR "Disease Management" OR "Medical Records Systems" OR "Delivery of Health Care" OR "Health Services" OR Health OR "health care" OR healthcare OR hospitals OR hospital OR "medical center" OR "medical centers" OR "Health systems" OR "Health system" OR health-care OR medical OR risk OR clinical OR "care management" OR "disease prevention" OR Medicare OR Medicaid OR uninsured OR insured OR "Electronic Health Records" OR ((electronic OR computerized OR automated OR administrative) AND (medical or health) AND (record OR records)))	989,225
4 Mitigation/ assessment	NOFT(mitigate OR mitigates OR mitigated OR mitigating OR mitigation OR reduce OR reduction OR reduced OR reduces OR reducing OR diminish OR diminished OR diminishing OR diminishes OR diminishment OR alleviate OR alleviates OR alleviated OR alleviating OR amend OR amends OR amended OR amending OR correct OR corrects OR corrected OR correcting OR correction OR rectify OR rectifies OR rectified OR rectifying OR rectification OR reform OR reforms OR reformed OR reforming OR reformation OR ameliorate OR strategy OR strategies OR prevent OR prevents OR preventing OR prevention OR ensure OR ensures OR ensuring OR ensured OR detect OR detects OR detected OR detecting OR detection OR limit OR limits OR limited OR limiting OR limitation OR limitations OR avoid OR avoids OR avoided OR avoiding OR remove OR removes OR removed OR removing OR removal OR decrease OR	3,783,315

Set #	Search Strategy	Results
	decreases OR decreased OR decreasing OR address OR addresses OR addressed OR addressing OR combat OR combats OR combated OR combatted OR combatting OR eliminate OR eliminates OR eliminated OR eliminating OR elimination OR minimize OR minimizes OR minimized OR minimizing OR minimise OR minimises OR minimised OR minimising OR Assess OR assesses OR assessed OR assessing OR assessment OR evaluate OR evaluates OR evaluated OR evaluating OR evaluation OR framework OR frameworks OR validate OR validates OR validated OR validating OR validation OR measure OR measures OR measured OR measuring OR measurement OR monitor OR monitors OR monitored OR monitoring OR analyze OR analyzes OR analyzed OR analyzing OR analysis OR analyse OR analyses OR analysed OR analysing OR screen OR screens OR screened OR screening OR report OR reports OR reported OR reporting OR compare OR compares OR compared OR comparing OR comparison OR diagnose OR diagnoses OR diagnosed OR diagnosing OR diagnostic OR diagnosis OR audit OR audits OR auditing OR audited OR comply OR complies OR complied OR complying OR compliance OR validate OR validates OR validated OR validating OR validation OR Checklist OR “algorithmic hygiene” OR “algorithmic playbook” OR “algorithmic bias playbook” OR checklist OR checklists OR tool OR tools OR toolkit OR toolkits OR guideline OR guidelines OR rubric OR rubrics OR guidance OR standard OR standards OR method OR methods OR methodology OR methodologies OR checkpoint OR checkpoints OR recommendation OR recommendations OR benchmark OR benchmarks OR regulation OR regulations OR regulatory OR generalizable OR generalizability OR “prejudice remover” OR fair OR unfair OR fairness OR unfairness OR “classification parity” OR trust OR trustworthy OR trustworthiness OR PROBAST OR “Prediction Model Risk Of Bias Assessment Tool” OR “Minimum Information for Medical AI Reporting guideline” OR “Artificial Intelligence Risk Management Framework” OR “AI RMF” OR “Medical Information Mart for Intensive Care” OR MIMIC OR “balanced accuracy” OR “disparate impact” OR “equal opportunity” OR “equalized odds” OR “statistical parity”)	
5	NOFT(“minority group” OR ethnic OR ethnicity OR “minority groups” OR “minority population” OR “minority populations” OR “people of color” OR “person of color” OR BIPOC OR “minority health” OR race OR races OR racial OR racially OR Blacks OR “African Americans” OR “African ancestry” OR Black OR “African American” OR “Mexican Americans” OR “Hispanic or Latino” OR “Hispanic American” OR Hispanic-American OR “Hispanic Americans” OR Hispanic-Americans OR Hispanics OR Hispanic OR “Mexican American” OR Mexican-American OR “Mexican Americans” OR Mexican-Americans OR “Cuban American” OR Cuban-American OR “Cuban Americans” OR Cuban-Americans OR “Latin American” OR Latin-American OR “Latin Americans” OR Latin-Americans OR Latinos OR Latino OR Latinas OR Latina OR Latinx OR latine OR latines OR “Spanish speaking” OR Spanish-speaking OR “Spanish speakers” OR Spanish-speakers OR Mexican OR Mexicans OR “Puerto Rican” OR “Puerto Ricans” OR “Asian Americans” OR “Asian American” OR Asian OR “Japanese Americans” OR “Japanese American” OR “Chinese Americans” OR “Chinese American” OR “Vietnamese Americans” OR “Vietnamese American” OR “Asian Indian Americans” OR “Asian Indian American” OR “Cambodian Americans” OR “Cambodian American” OR “Hmong Americans” OR “Hmong American” OR “Korean Americans” OR “Korean American” OR “Filipino Americans” OR “Filipino American” OR “American Indians” OR “Alaska Natives” OR “Indigenous Peoples” OR “Indigenous People” OR “Indigenous Population” OR “Indigenous Populations” OR tribes OR “Native Peoples” OR “Native People” OR “First Nation” OR “native person” OR Alaskan OR Alaskans OR “Caribbean American” OR Caribbean-American OR “Caribbean Americans” OR Caribbean-Americans OR migrant OR migrants OR immigrant OR immigrants OR emigrant OR emigrants OR demographic OR demographics OR Polynesian OR Polynesians OR “Pacific Islander” OR “Pacific Islanders” OR Hawaiian OR Hawaiians)	449,815
6	1 AND 2 AND 3 AND 4 AND 5	731

Notes

1. Office for Civil Rights (OCR), Office of the Secretary, HHS. Nondiscrimination in health programs and activities. Final rule. Fed Regist. 2016;81(96):31375–473. [PubMed: 27192742]
2. Department of Health and Human Services. Covered entities and business associates [Internet]. Washington (DC): HHS; 2015 Nov 23 [last updated 2017 Jun 16; cited 2023 Aug 18]. Available from: https://www.hhs.gov/hipaa/for-professionals/covered-entities/index.html
3. Forrest S. Artificial intelligence/machine learning (AI/ML)-enabled medical devices: tailoring a regulatory framework to encourage responsible innovation in AI/ML [Internet]. Silver Spring (MD): Food and Drug Administration; [cited 2023 Aug 18]. Available from: https://www.fda.gov/media/160125/download

4. Schwartz R, Vassilev A, Greene K, Perine L, Burt A, Hall P. Towards a standard for identifying and managing bias in artificial intelligence [Internet]. Gaithersburg (MD): National Institute of Standards Technology; 2022 Mar [cited 2023 Aug 18]. Accessed June 26, 2023. (NIST Special Publication No. 1270). Available from: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1270.pdf
5. Food and Drug Administration. General principles of software validation; final guidance for industry and FDA staff [Internet]. Silver Spring (MD): FDA; 2002 Jan [cited 2023 Aug 18]. Available from: https://www.fda.gov/media/73141/download
6. Agency for Healthcare Research and Quality. Impact of healthcare algorithms on racial and ethnic disparities in health and healthcare [Internet]. Rockville (MD): AHRQ; 2022 Jan 25 [cited 2023 Aug 18]. Available from: https://effectivehealthcare.ahrq.gov/products/racial-disparities-health-healthcare/protocol
7. National Institute of Standards and Technology. AI risk management framework [Internet]. Gaithersburg (MD): NIST; 2023 Jan [cited 2023 Aug 18]. Available from: https://www.nist.gov/itl/ai-risk-management-framework
8. Coalition for Health AI. Blueprint for trustworthy AI implementation guidance and assurance for healthcare [Internet]. McLean (VA): Mitre Corporation; 2023 Apr 4 [cited 2023 Aug 18]. Available from: https://coalitionforhealthai.org/papers/blueprint-for-trustworthy-ai_V1.0.pdf
9. White House, Office of Science and Technology Policy. Blueprint for an AI bill of rights: making automated systems work for the American people [Internet]. Washington (DC): White House; [cited 2023 Aug 18]. Available from: https://www.whitehouse.gov/ostp/ai-bill-of-rights/
10. Huang J, Galal G, Etemadi M, Vaidyanathan M. Evaluation and mitigation of racial bias in clinical machine learning models: scoping review. JMIR Med Inform. 2022;10(5):e36388. [PubMed: 35639450]
11. Kaur D, Uslu S, Rittichier KJ, Durresi A. Trustworthy artificial intelligence: a review. ACM Comput Surv. 2022;55(2):1–38.
12. Daneshjou R, Smith MP, Sun MD, Rotemberg V, Zou J. Lack of transparency and potential bias in artificial intelligence data sets and algorithms: a scoping review. JAMA Dermatol. 2021;157(11):1362–9. [PubMed: 34550305]
13. Bear Don't Walk OJ 4th, Reyes Nieva H, Lee SS-J, Elhadad N. A scoping review of ethics considerations in clinical natural language processing. JAMIA Open. 2022;5(2):ooac039. [PubMed: 35663112]
14. Pearson A, Wiechula R, Court A, Lockwood C. The JBI model of evidence-based healthcare. Int J Evid Based Healthc. 2005;3(8):207–15. [PubMed: 21631749]
15. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med. 2018;169(7):467–73. [PubMed: 30178033]
16. Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ. 2021;372(160):n160. [PubMed: 33781993]
17. To access the appendix, click on the Details tab of the article online.
18. Weissman GE, Teeple S, Eneanya ND, Hubbard RA, Kangovi S. Effects of neighborhood-level data on performance and algorithmic equity of a model that predicts 30-day heart failure readmissions at an urban academic medical center. J Card Fail. 2021;27(9):965–73. [PubMed: 34048918]
19. Puyol-Antón E, Ruijsink B, Mariscal Harana J, Piechnik SK, Neubauer S, Petersen SE, et al. Fairness in cardiac magnetic resonance imaging: assessing sex and racial bias in deep learning-based segmentation. Front Cardiovasc Med. 2022;9:859310. [PubMed: 35463778]
20. Thompson HM, Sharma B, Bhalla S, Boley R, McCluskey C, Dligach D, et al. Bias and fairness assessment of a natural language processing opioid misuse classifier: detection and mitigation of electronic health record data disadvantages across racial subgroups. J Am Med Inform Assoc. 2021;28(11):2393–403. [PubMed: 34383925]

21. Foryciarz A, Pfohl SR, Patel B, Shah N. Evaluating algorithmic fairness in the presence of clinical guidelines: the case of atherosclerotic cardiovascular disease risk estimation. BMJ Health Care Inform. 2022;29(1):e100460.
22. Coston A, Rambachan A, Chouldechova A. Characterizing fairness over the set of good models under selective labels. PMLR. 2021;139:2144–55.
23. Akbilgic O, Langham MR Jr, Davis RL. Race, preoperative risk factors, and death after surgery. Pediatrics. 2018;141(2):e20172221. [PubMed: 29321256]
24. Borgese M, Joyce C, Anderson EE, Churpek MM, Afshar M. Bias assessment and correction in machine learning algorithms: a use-case in a natural language processing algorithm to identify hospitalized patients with unhealthy alcohol use. AMIA Annu Symp Proc. 2021;2021:247–54. [PubMed: 35308909]
25. Radovanović S, Petrović A, Delibašić B, Suknović M. Making hospital readmission classifier fair—what is the cost? Paper presented at: Central European Conference on Information and Intelligent Systems; 2019 Oct 2–4; Varazdin, Croatia.
26. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019;366(6464):447–53. [PubMed: 31649194]
27. Buckley A, Sestito S, Ogundipe T, Roig J, Rosenberg HM, Cohen N, et al. Racial and Ethnic disparities among women undergoing a trial of labor after cesarean delivery: performance of the VBAC Calculator with and without patients' race/ethnicity. Reprod Sci. 2022;29(7):2030–8. [PubMed: 35534768]
28. Segar MW, Hall JL, Jhund PS, Powell-Wiley TM, Morris AA, Kao D, et al. Machine learning-based models incorporating social determinants of health vs traditional models for predicting in-hospital mortality in patients with heart failure. JAMA Cardiol. 2022;7(8):844–54. [PubMed: 35793094]
29. Yan M, Pencina MJ, Boulware LE, Goldstein BA. Observability and its impact on differential bias for clinical prediction models. J Am Med Inform Assoc. 2022;29(5):937–43. [PubMed: 35211742]
30. Allen A, Mataraso S, Siefkas A, Burdick H, Braden G, Dellinger RP, et al. A Racially unbiased, machine learning approach to prediction of mortality: algorithm development study. JMIR Public Health Surveill. 2020;6(4):e22400. [PubMed: 33090117]
31. Mosteiro P, Kuiper J, Masthoff J, Scheepers F, Spruit M. Bias discovery in machine learning models for mental health. Information (Basel). 2022;13(5):237.
32. Pfohl SR, Foryciarz A, Shah NH. An empirical characterization of fair machine learning for clinical risk prediction. J Biomed Inform. 2021;113:103621. [PubMed: 33220494]
33. Afrose S, Song W, Nemeroff CB, Lu C, Yao DD. Subpopulation-specific machine learning prognosis for underrepresented patients with double prioritized bias correction. Commun Med (Lond). 2022;2:111. [PubMed: 36059892]
34. Burlina P, Joshi N, Paul W, Pacheco KD, Bressler NM. Addressing artificial intelligence bias in retinal diagnostics. Transl Vis Sci Technol. 2021;10(2):13.
35. Adeli E, Zhao Q, Pfefferbaum A, Sullivan EV, Fei-Fei L, Niebles JC, et al. Representation learning with statistical independence to mitigate bias. IEEE Winter Conf Appl Comput Vis. 2021;2021:2512–22. [PubMed: 34522832]
36. Mullainathan S, Obermeyer Z. On the inequity of predicting A while hoping for B. AEA Pap Proc. 2021;111:37–42.
37. Samorani M, Blount LG. Machine learning and medical appointment scheduling: creating and perpetuating inequalities in access to health care. Am J Public Health. 2020;110(4):440–1. [PubMed: 32159974]
38. Gama RM, Clery A, Griffiths K, Heraghty N, Peters AM, Palmer K, et al. Estimated glomerular filtration rate equations in people of self-reported black ethnicity in the United Kingdom: inappropriate adjustment for ethnicity may lead to reduced access to care. PLoS One. 2021;16(8):e0255869. [PubMed: 34383841]
39. Park Y, Hu J, Singh M, Sylla I, Dankwa-Mullan I, Koski E, et al. Comparison of methods to reduce bias from clinical prediction models of postpartum depression. JAMA Netw Open. 2021;4(4):e213909. [PubMed: 33856478]

40. Pierson E, Cutler DM, Leskovec J, Mullainathan S, Obermeyer Z. An algorithmic approach to reducing unexplained pain disparities in underserved populations. Nat Med. 2021;27(1):136–40. [PubMed: 33442014]
41. Dankwa-Mullan I, Scheufele EL, Matheny ME, Quintana Y, Chapman WW, Jackson G, et al. A proposed framework on integrating health equity and racial justice into the artificial intelligence development lifecycle. J Health Care Poor Underserved. 2021;32(2):300–17.
42. Richardson S, Lawrence K, Schoenthaler AM, Mann D. A framework for digital health equity. NPJ Digit Med. 2022;5(1):119. [PubMed: 35982146]
43. Sikstrom L, Maslej MM, Hui K, Findlay Z, Buchman DZ, Hill SL. Conceptualising fairness: three pillars for medical algorithms and health equity. BMJ Health Care Inform. 2022;29(1):e100459.
44. Uche-Anya E, Anyane-Yeboa A, Berzin TM, Ghassemi M, May FP. Artificial intelligence in gastroenterology and hepatology: how to advance clinical practice while ensuring health equity. Gut. 2022;71(9):1909–15. [PubMed: 35688612]
45. Nelson ALH, Zanti S. A framework for centering racial equity throughout the administrative data life cycle. Int J Popul Data Sci. 2020;5(1):1367. [PubMed: 34007882]
46. Reeves M, Bhat HS, Goldman-Mellor S. Resampling to address inequities in predictive modeling of suicide deaths. BMJ Health Care Inform. 2022;29(1):e100456.
47. Lomis K, Jeffries P, Palatta A, Sage M, Sheikh J, Sheperis C, et al. Artificial intelligence for health professions educators. NAM Perspect. 2021 Sep 8. [Epub ahead of print].
48. Raza S (Cornell University). Connecting fairness in machine learning with public health equity. Arxiv [preprint on the Internet]. 2023 Apr 8 [cited 2023 Aug 18]. Available from: https://arxiv.org/abs/2304.04761
49. Vyas DA, Eisenstein LG, Jones DS. Hidden in plain sight—reconsidering the use of race correction in clinical algorithms. N Engl J Med. 2020;383(9):874–82. [PubMed: 32853499]
50. Argentieri R, Mason TA, Hefcart J, Henry J. Embracing health equity by design. Health IT Buzz [serial on the Internet]. 2022 Feb 22 [cited 2023 Aug 18]. https://www.healthit.gov/buzz-blog/health-it/embracing-health-equity-by-design
51. Wong WF, LaVeist TA, Sharfstein JM. Achieving health equity by design. JAMA. 2015;313(14):1417–8. [PubMed: 25751310]
52. Clinical Trials Transformation Initiative. Quality by design [Internet]. Durham (NC): CTTI; [cited 2023 Aug 18]. Available from: https://ctti-clinicaltrials.org/our-work/quality/quality-by-design/

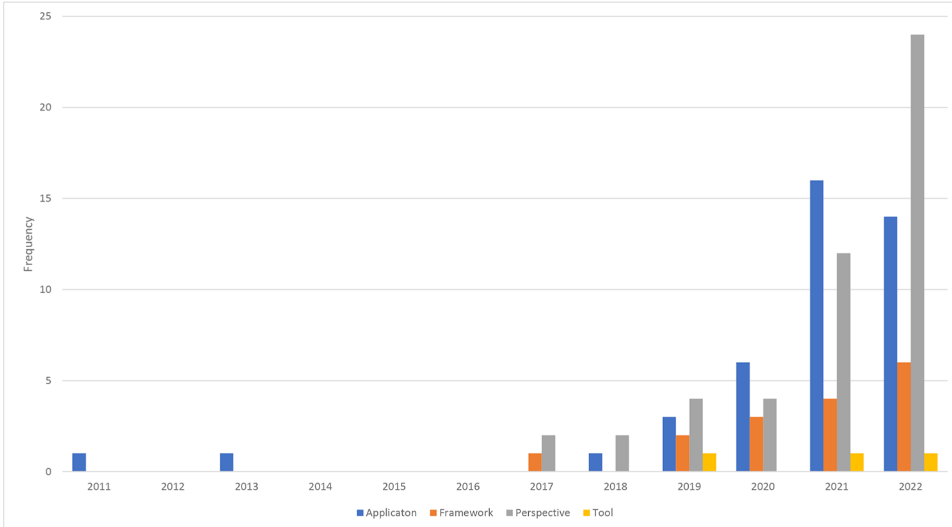


Exhibit 1:

Number of articles by article type (empirical health care application, framework, review or perspective, tool), from a review of studies on bias in clinical algorithms, 2011–22

Source: SOURCE Authors' analysis of studies that identify and mitigate bias in clinical algorithms.

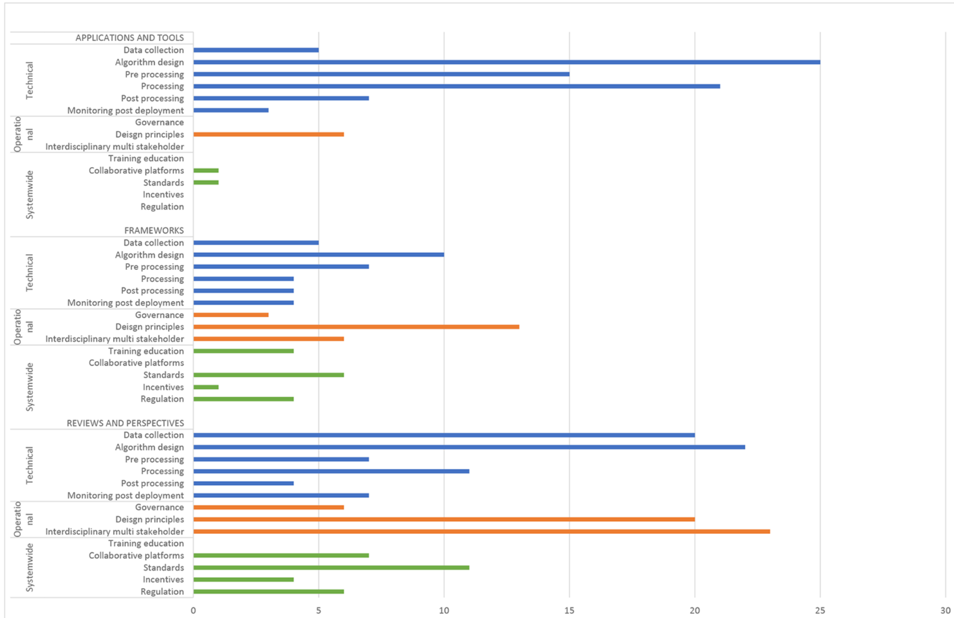


Exhibit 2:

Number of articles on bias mitigation strategies by category (technical, operational, and systemwide) by article type, from a review of studies on bias in clinical algorithms, 2011–22

Source: SOURCE Authors' analysis of studies that identify and mitigate bias in clinical algorithms.

exhibit 3:

Bias mitigation strategies reported in empirical health care applications by category, from a review of studies on bias in clinical algorithms, 2011–22

Category	Strategies Reported	Studies
Algorithm design	New outcome variable	Obermeyer 2019, Landy 2021, Pierson 2021
	Remove race and ethnicity and social determinants of health from the model	Samorani 2020, Gama 2021, Park 2021, Buckley 2022, Huang 2022
	Add race and ethnicity and social determinants of health to the model	Hammond 2020, Weissman 2021, Segar 2022
	Determining when to add or remove sensitive variables	Yan 2022
	Use different algorithm	Pierson 2021, Segar 2022
	Stratify models by race	Shores 2013, Akbilgic 2018, Do 2020, Borgese 2021, Thompson 2021, Afrose 2022, Puyol-Anton 2022, Segar 2022, Foryciarz 2022
	Weighting methods	Coston 2019, Radovanovic 2019, Allen 2020, Park 2021, Mosteiro 2022
	Sampling methods	Afrose 2022, Puyol-Anton 2022, Park 2022, Reeves 2022
	Data augmentation	Burlina 2021
Pre-processing	Disparate impact remover to debias variables	Park 2022
	Adjust the algorithm's objective function[a]	Samorani 2020, Adeli 2021, Park 2021, Pfohl 2021, Foryciarz 2022, Mosteiro 2022, Puyol-Anton 2022, Perez 2022
	Bias correction	Afrose 2022
Processing	Adversarial and transfer learning	Radovanović 2019, Gao 2020, Toseef 2022
	Varying cutoff points/thresholds	Radovanović 2019, Gianattasio 2020, Thompson 2021, Rodolfa 2021
Postprocessing	Recalibration	Barda 2021

Sources/Notes: SOURCES Source details are located in the appendix (see note 17 in text). Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations, (see note 26 in text). Landy R, Young CD, Skarzynski M, Skarzynski M, Cheung LC, Berg CD, et al. Using Prediction Models to Reduce Persistent Racial and Ethnic Disparities in the Draft 2020 USPSTF Lung Cancer Screening Guidelines. JNCI J Natl Cancer Inst. 2021;113 (11):1590–4. Pierson E, Cutler DM, Leskovec J, Mullainathan S, Obermeyer Z. An algorithmic approach to reducing unexplained pain disparities in underserved populations, (see note 40 in text). Samorani M, Blount LG. Machine Learning and Medical Appointment Scheduling: Creating and Perpetuating Inequalities in Access to Health Care, (see note 37 in text). Buckley A, Sestito S, Ogundipe T, Roig J, Rosenberg HM, Cohen N, et al. Racial and Ethnic Disparities Among Women Undergoing a Trial of Labor After Cesarean Delivery: Performance of the VBAC Calculator with and without Patients' Race/Ethnicity, (see note 27 in text). Gama RM, Clery A, Griffiths K, Heraghty N, Peters AM, Palmer K, et al. Estimated glomerular filtration rate equations in people of self-reported black ethnicity in the United Kingdom: Inappropriate adjustment for ethnicity may lead to reduced access to care, (see note 38 in text). Huang J, Galal G, Etemadi M, Vaidyanathan M. Evaluation and Mitigation of Racial Bias in Clinical Machine Learning Models: Scoping Review, (see note 10 in text). Park Y, Hu J, Singh M, Sylla I, Dankwa-Mullan I, Koski E, et al. Comparison of Methods to Reduce Bias From Clinical Prediction Models of Postpartum Depression, (see note 39 in text). Weissman GE, Teeple S, Eneanya ND, Hubbard RA, Kangovi S. Effects of Neighborhood-level Data on Performance and Algorithmic Equity of a Model That Predicts 30-day Heart Failure Readmissions at an Urban Academic Medical Center, (see note 18 in text). Hammond G, Johnston K, Huang K, Joynt Maddox KE. Social Determinants of Health Improve Predictive Accuracy of Clinical Risk Models for Cardiovascular Hospitalization, Annual Cost, and Death. Circ Cardiovasc Qual Outcomes. 2020;13(6):e006752. Segar MW, Hall JL, Jhund PS, Powell-Wiley TM, Morris AA, Kao D, et al. Machine Learning-Based Models Incorporating Social Determinants of Health vs Traditional Models for Predicting In-Hospital Mortality in Patients With Heart Failure, (see note 28 in text). Yan M, Pencina MJ, Boulware LE, Goldstein BA. Observability and its impact on differential bias for clinical prediction models, (see note 29 in text). Shores NJ, Dodge JL, Feng S, Terrault NA. Donor Risk Index for African American liver transplant recipients with hepatitis C virus. Hepatology. 2013;58(4):1263–9. Akbilgic O, Langham MR, Davis RL. Race, Preoperative Risk Factors, and Death After Surgery, (see note 23 in text). Do H, Nandi S, Putzel P, Smyth P, Zhong J. Joint Fairness Model with Applications to Risk Predictions for Under-represented Populations [Internet]. Ithaca (NY): Cornell University, ArXiv; 2021 May 10 [last updated 2022 Feb 23; cited 2023 Aug 23]. Available from: https://arxiv.org/abs/2105.04648. Afrose S, Song W, Nemeroff CB, Lu C, Yao DD. Subpopulation-specific machine learning prognosis for underrepresented patients with double prioritized bias correction, (see note 33 in text). Borgese M, Joyce C, Anderson EE, Churpek MM, Afshar M. Bias Assessment and Correction in Machine

Learning Algorithms: A Use-Case in a Natural Language Processing Algorithm to Identify Hospitalized Patients with Unhealthy Alcohol Use, (see note 24 in text). Puyol-Antón E, Ruijsink B, Mariscal Harana J, Piechnik SK, Neubauer S, Petersen SE, et al. Fairness in Cardiac Magnetic Resonance Imaging: Assessing Sex and Racial Bias in Deep Learning-Based Segmentation, (see note 19 in text). Thompson HM, Sharma B, Bhalla S, Boley R, McCluskey C, Dligach D, et al. Bias and fairness assessment of a natural language processing opioid misuse classifier: detection and mitigation of electronic health record data disadvantages across racial subgroups, (see note 20 in text). Foryciarz A, Pfohl SR, Patel B, Shah N. Evaluating algorithmic fairness in the presence of clinical guidelines: the case of atherosclerotic cardiovascular disease risk estimation, (see note 21 in text). Coston A, Rambachan A, Chouldechova A. Characterizing Fairness Over the Set of Good Models Under Selective Labels, (see note 22 in text). Radovanović S, Petrović A, Delibašić B, Suknović M. Making hospital readmission classifier fair – What is the cost? Proceedings of the Central European Conference on Information and Intelligent Systems, (see note 25 in text). Allen A, Mataraso S, Siefkas A, Burdick H, Braden G, Dellinger RP, et al. A Racially Unbiased, Machine Learning Approach to Prediction of Mortality: Algorithm Development Study, (see note 30 in text). Mosteiro P, Kuiper J, Masthoff J, Scheepers F, Spruit M. Bias Discovery in Machine Learning Models for Mental Health, (see note 31 in text). Park J, Arunachalam R, Silenzio V, Singh VK. Fairness in Mobile Phone-Based Mental Health Assessment Algorithms: Exploratory Study. JMIR Form Res. 2022;6(6):e34366. Reeves M, Bhat HS, Goldman-Mellor S. Resampling to address inequities in predictive modeling of suicide deaths, (see note 46 in text). Burlina P, Joshi N, Paul W, Pacheco KD, Bressler NM. Addressing Artificial Intelligence Bias in Retinal Diagnostics, (see note 34 in text). Adeli E, Zhao Q, Pfefferbaum A, Sullivan EV, Fei-Fei L, Niebles JC, et al. Representation Learning with Statistical Independence to Mitigate Bias, (see note 35 in text). Pfohl SR, Foryciarz A, Shah NH. An empirical characterization of fair machine learning for clinical risk prediction, (see note 32 in text). Perez Alday EA, Rad AB, Reyna MA, Sadr N, Gu A, Li Q, et al. Age, sex and race bias in automated arrhythmia detectors. J Electrocardiol. 2022;74:5–9. Gao Y, Cui Y. Deep transfer learning for reducing health care disparities arising from biomedical data inequality. Nat Commun. 2020;1 (1):5131. Toseef M, Li X, Wong KC. Reducing healthcare disparities using multiple multiethnic data distributions with fine-tuning of transfer learning. Brief Bioinform. 2022;23(3):bbac078. Gianattasio KZ, Ciarleglio A, Power MC. Development of Algorithmic Dementia Ascertainment for Racial/Ethnic Disparities Research in the US Health and Retirement Study. Epidemiol Camb Mass. 2020;31(1):126–33. Rodolfa KT, Lamba H, Ghani R. Empirical observation of negligible fairness–accuracy trade-offs in machine learning for public policy. Nat Mach Intell. 2021;3(10):896–904. Barda N, Yona G, Rothblum GN, Greenland P, Leibowitz M, Balicer E, et al. Addressing bias in prediction models by improving subpopulation calibration. J Am Med Inform Assoc. 2021;28(3):549–58. NOTES These techniques were sourced from a variety of clinical applications with different clinical algorithms. The bias mitigation strategies reported excluded data collection suggestions such as using more recent data or collecting more (diverse) data and monitoring postdeployment.

[a]There are numerous methods to adjust the algorithm's objective function, including regularization, constrained optimization, and others.