Reducing Racial Biases within Healthcare Applications of Artificial Intelligence (AI) With Transparency

Mishayla Harve[1], Sakthi Priya Ramamoorthy[1], Viresh Pati[1], Garen Bainbridge[1], Abigayle Kankolenski[1], Bratee Podder[1], Matthew Sampt[1]

[1]Georgia Institute of Technology, Atlanta, GA, USA
https://doi.org/10.38126/JSPG250110
Corresponding author: mharve@gatech.edu
Keywords: artificial intelligence; AI; healthcare; healthcare equity; racial bias; transparency; data collection; accountability

Executive Summary: Artificial intelligence (AI) is increasingly being used in healthcare for applications such as drug discovery, diagnostics, disease management, and delivery of services. However, integrating AI and healthcare raises concerns about reinforcing existing societal prejudices: AI systems are known to exhibit racial biases by making inaccurate and unreliable decisions based on race when it is irrelevant to the task. Furthermore, government directives currently lack consistent standards for regulating AI and offer insufficient guidance on preventing the perpetuation of harmful racial biases, especially in healthcare. To improve patients' quality of life interacting with AI systems, it is essential to ensure transparency regarding these systems. Additionally, it is vital to ensure that innovation dedicated to improving healthcare enhances the integrity of the patient's experience rather than compounds existing systemic disparities. The authors propose three recommendations to address racial biases in healthcare applications of AI and emphasize the need for legislation placing AI regulation in healthcare at the forefront of healthcare policy agendas.

I. Introduction

Artificial intelligence (AI) has exploded in countless domains, spurring public interest and financial investment. Because of its learning abilities, AI can be applied to any field, such as the creative industry and education (US Government of Accountability Office 2023). AI has also shown promise in healthcare as an assistive tool for clinicians (Akinrinmade et al. 2023). Researchers at the University of Pavia in Italy successfully used transformer-based AI models to convert radiology reports into structured, clinical reports (Bergomi et al. 2024). AI systems have also been used to detect diseases more accurately, personalize treatment plans, and improve patient scheduling.

Notably, the rapid adoption of these emerging technologies introduces the risk of propagating biases already embedded within our society. This is particularly pronounced in healthcare, where AI systems have exhibited a concerning degree of racial bias (Samorani et al. 2021; Obermeyer et al., 2019; Agarwal et al. 2023) These biases can reinforce racial and socioeconomic disparities in healthcare, and if poorly addressed, marginalized communities could encounter heightened obstacles in obtaining equitable healthcare services.

The challenges with racial bias in healthcare AI are compounded by the fact that decisions aided by AI, unlike human judgments, can be amplified and scaled to impact vast patient populations simultaneously (Nazer et al. 2023). The probabilistic nature of AI-based decision-making lends to a false sense of legitimacy, as accountability for its outcomes is difficult to assign (Hao 2019). Coupled with the lack of transparency associated with many AI systems and the involvement of multiple parties in its development, it becomes difficult to assign clear responsibility for harmful outcomes (Rodrigues 2020).

Due to the significant ethical concerns associated with AI applications in healthcare, policymakers, healthcare leaders, and AI developers must prioritize regulation. Existing international, federal, and state-level AI policies only vaguely establish guidance on AI implementation without perpetuating dangerous racial bias within healthcare. Furthermore, it is imperative to ensure transparency regarding how and when these systems are used, given the indirect way AI interacts with patients in most medical contexts. There is a crucial distinction between a scenario where a doctor transparently uses AI as a tool for consultancy services and one where patients unknowingly interact with AI systems without explicit consent. This lack of awareness critically undermines patient autonomy.

II. Major forms of AI algorithmic biases cause racial inequities in healthcare

AI inherits society's flaws and biases, especially when it functions effectively. In healthcare, most modern AI systems learn patterns from vast amounts of real-life data. This inevitability introduces and integrates human flaws and biases into algorithms (Fazelpour & Danks, 2021). In many models, consolidated data features are passed as inputs to make predictions or produce outputs (Maind and Walker 2014). If these results are deemed accurate, as measured by chosen or learned metrics, the AI model reinforces the underlying patterns that will make similar predictions more often (Barto and Sutton 2007). Conversely, if the model's outputs are inaccurate, the model minimizes the value of features associated with the undesired results (Barto and Sutton 2007). Therefore when AI models reinforce desirable features and decisions, they might also perpetuate the algorithmic biases embedded in the data they were trained on. Hence, if bias exists in society, it must also be true that data and AI models are vulnerable to reflecting it.

i. Training and "garbage in, garbage out"

"Garbage In, Garbage Out" is an explicit form of statistical bias, in which an AI model receives a problematic or unbalanced dataset that excludes or heavily leans toward particular characteristics (Mueller 2020). This causes the AI model to create biased decisions that further reinforce societal prejudices despite being unrelated to the decision at hand (Mueller 2020). If the training data used by an AI system is flawed, then the outcomes of the AI will also be flawed.

For example, three algorithms used to detect dermatological malignancies—ModelDerm, DeepDerm, and HAM10000—struggled to diagnose darker-skinned patients at earlier stages of the malignancy (Liu et al. 2022). In this example, AI models struggled to provide helpful diagnoses to a large population of people of color—not because it intentionally prioritized one ethnicity, but because the system was trained on a non-diverse dataset (Liu et al. 2022). Even without harmful motivations, the AI models were unhelpful in improving patient care systemically and exacerbated accessibility issues in an already disadvantaged community.

ii. Deployment and implicit bias

Implicit bias may be more complex but presents equally significant consequences if left unregulated. Consider a patient scheduling system where race is one of many features. A logical goal of any scheduling paradigm is to avoid allocating prime time slots to patients who frequently miss appointments. Complex socioeconomic factors (e.g. medical deserts) contribute to higher no-show rates in particular communities. An AI model could use race as a proxy to make decisions that deny optimal time slots to patients of a racial group. Without malicious intent, AI in medicine can propagate patient discrimination by not accounting for societal factors, such as social determinants of health (CDC 2024).

In a previous study, a machine learning-driven scheduling system caused Black patients to wait around 30% longer than non-Black patients in a large specialty outpatient clinic (Shanklin 2022). Many studies have found that race is correlated with no-show probability, meaning existing datasets may inadvertently lead these algorithms to assign longer waiting times to certain racial groups (Shimotsu et al., 2015). Consequently, Black patients, who might already face barriers that contribute to no-show rates, could be unfairly penalized with less desirable appointment times, further reinforcing and exacerbating the bias in the AI system. Although the explicit use of race has been regulated in contexts like lending, housing, and hiring, it remains largely unregulated in healthcare (Hersch & Shinall 2015). It is important to note that other forms of algorithmic bias arise when developing AI, such as problem specification, data curation, and choosing validation criteria (Lin & Chen 2022).

III. A brief assessment of current directives targeting AI bias

While these existing proposed interventions offer a foundation for addressing AI bias, their effectiveness in fully addressing the issues presented in AI-driven healthcare still needs to be improved. On the international level, The Five Safes Program framework, developed by the UK Health Authority, facilitates the ethical use of patient data through the UK Data Service SecureLab (United Kingdom Data Service), ensuring access to sensitive information without disclosing identifiable details (United Kingdom Data Service). This framework helps mitigate racial biases by promoting diverse patient datasets but does not fully address cybersecurity threats or data privacy concerns prevalent in the US (McClain et al. 2023). In the United States, Executive Order 14410 mandates AI developers share safety test results and establish standards for AI systems to ensure their safety and trustworthiness (US Executive Order 14410 2023). While this order could complement frameworks like the Five Safes Program, it isn't as durable as established laws, which could limit its effectiveness in regulating AI. On a state level, Georgia House Bill 887 aims to prohibit the use of AI in certain healthcare decisions and would task the Georgia Composite Medical Board with issuing related rules (Thomas 2024). However, the bill hasn't yet become law and lacks detailed guidelines for enforcing ethical AI usage in healthcare. In summary, effective legislation addressing AI bias, especially in healthcare, is few and far between.

The shortcomings of the current directives must be addressed. The areas of improvement for the mentioned directives are evident: they are limited in scope, lack comprehensive enforcement mechanisms, and fail to address the rapidly evolving nature of AI technology. Still, several organizations and advocacy groups are pushing for effective political action in targeting racial bias. In February 2024, the RAISE (Responsible AI for Social and Ethical Healthcare) conference, organized by the Department of Biomedical Informatics at Harvard Medical School, addressed the current shortcomings in healthcare systems (Goldberg et al. 2024). Some of their calls to action include AI vendors revealing training data sources to patients and regulators, reporting missing or lesser demographic inclusion in datasets to the public, and being transparent about their performance for different demographic characteristics. Other advocacy groups such as the Coalition for Health AI (CHAI) are developing preliminary guidelines to drive high-quality healthcare by adopting credible, fair, and transparent health AI systems (Coalition for Health AI 2023). Dr. Nishit Patel, the associate professor of Medicine and AI integration lead at the University of California, San Francisco, calling for AI policies that include heavy input from medical professionals in the development of new AI models, require more transparent explanations behind AI logic and decision-making processes, and support the iterative improvement of AI tools without overly burdensome regulatory barriers (Rossy 2023). These advocacy movements highlight the need for stronger flexible regulations to effectively address AI bias and ensure fair outcomes in healthcare.

IV. Policy recommendations

The policy recommendations outlined below are built on transparency because it is crucial for addressing racial biases and promoting fairness in AI systems, particularly in healthcare, without stifling innovation. They aim to codify existing AI-focused policies to limit discrimination as AI becomes more widespread. By setting a standard for healthcare AI, these policies can improve patient experiences, increase access to care, and establish better regulations benefiting all demographics, even beyond the medical sector. As AI evolves, these policies prioritize patient protection. The ranked suggestions outline who must be transparent to consumers interacting with healthcare AI are below:

i. Mandating healthcare disclosure and transparency on their use of AI to patients

Healthcare policymakers should mandate transparency and patient choice of AI systems used for healthcare services before the procedure is enacted. At the beginning of each patient visit, healthcare providers using AI systems must disclose usage, risk, methods, and rationale to patients, alongside resources for understanding AI functionality. Patients unwilling to participate in AI-enabled services must receive alternative medical procedures of equivalent quality. For example, under this policy, a physician would need to inform the patient about the potential risks and benefits of using AI-enabled diagnostic software using it on the patient, which could involve having the patient sign a consent form. If the patient chooses not to proceed with the AI-enabled procedure, the physician must diagnose the patient without using the AI system. This process is similar to how anesthesia is delivered in medical settings. While anesthesia is highly beneficial for keeping patients comfortable during surgery, it can also lead to health complications. Physicians must disclose these risks to patients and obtain their consent before proceeding with its use (Waisel & Truog 1997).

Additionally, while AI developers and companies undergo intense federal oversight before being made available for medical applications, healthcare providers, and major stakeholders in the particular healthcare setting (like large patient groups), will ultimately have the discretion to decide whether to implement the AI model. If the dataset or AI model's development process conflicts with their patient's values, they are not obligated to use it, even if the larger hospital institution has approved the model. This approach allows healthcare providers to make more informed decisions that prioritize the needs and values of their patients (Lynch 2024). Such measures aim to uphold patient autonomy, ensure equitable healthcare delivery, and foster accountability in AI implementation within healthcare settings.

This policy addresses consumer protection directly, and the consequences of AI bias indirectly. Any policy for consumer protection implies stronger company regulations. Large corporations have historically refused transparency because accountability drives production costs up. Although any form of AI regulation is arguably novel, similar laws have been passed that require corporations to inform consumers before product purchase, such as the Federal Cigarette Labeling and Advertising Act, which forces tobacco companies to disclose the consequences of smoking on their products. Informed consent between the consumer and the company largely benefits the public (Federal Trade Commission 1965).

For AI, transparency may be a long-term solution to combatting racial disparities present in AI. However, as AI continues to change and become even more complex over time, requesting transparency may be a larger ask from AI corporations, increasing resistance to the policy. Enforcing the policy will not require intense public funding. Most of the cost falls onto AI companies, whose products will be held to higher standards, should they be used in a hospital setting. Even if this transparency-focused policy causes some healthcare systems to stop using technology that could have optimized their hospitals, these systems are not dependent on AI. Hospitals have successfully managed their medical and non-medical procedures long before AI was introduced into healthcare settings.

The main objective of this policy is to ensure patient protection, with the secondary goal of addressing harmful racial biases. Due to the novelty of the policy, the efficacy of the policy is unpredictable. While the ultimate effectiveness of this policy alternative remains unclear, it proactively tackles prejudices in an increasingly prominent healthcare technology.

Although there are few case studies on the impact of transparency in AI usage on patient outcomes and bias reduction—likely because AI use in medicine is still in its early stages, as noted in the introduction—transparency has consistently been identified as a top priority, particularly when it concerns consumers. Curt Langlotz, associate director of the Stanford Institute of Human-Centered Artificial Intelligence, notes that AI models often don't generalize well to new users (Lynch 2024). As a result, potential users of healthcare AI are finding it difficult to determine whether a particular AI system will integrate effectively into their practice (Lynch 2024). This underscores the need for greater transparency regarding the data on which these products were trained. While research on the potential effects of this policy is rare, this policy aims to address the necessity of transparency in a healthcare setting.

ii. Establishing a racial equity committee on the National Institute of Standards and Technology (NIST)'s Diversity, Equity, and Inclusion (DEI) board

Unlike conventional diversity and inclusion initiatives, its main objective is to ensure patient protection, with the secondary goal of addressing potentially harmful racial biases. This proposal calls for establishing a Racial Equity Committee on the National Institute of Standards and Technology (NIST) Diversity Equity Inclusion (DEI) board, whose primary focus will be evaluating the AI systems developed by companies associated with NIST. NIST is a non-regulatory, federal agency that oversees domestic measures for promoting science and technology innovation, including cybersecurity protections, AI risk, and AI safety (NIST 2024). Therefore, the proposed Racial Equity Committee within NIST would propose crucial standards for AI systems that mitigate potential biases and discriminatory outcomes. More importantly, the Racial Equity Committee's activities, findings, and recommendations would be made publicly accessible through annual reports and publications. This committee could also be an important voice when NIST further refines vital proposals, such as the NIST AI Risk Management Framework, recently established in July 2024 (NIST 2024).

The level of transparency that would become the standard of the Racial Equity Committee would be comparable to existing regulatory frameworks like the Financial Industry Regulatory Authority (FINRA). FINRA, which oversees brokerage firms and exchange markets in the US, releases an annual regulatory oversight report (Financial Industry Regulatory Authority 2024). Outlining the specific regulations firms must follow, including those related to transparency and fair practices, could serve as a vital model for creating AI-specific regulations in healthcare.

The establishment of a Racial Equity Committee brings numerous advantages. Technically, it is highly feasible, as NIST can tap into its existing expertise and potentially draw from individuals already involved in the Diversity, Equity, and Inclusion (DEI) board. This initiative would lay the groundwork for future legislation across various industries, fostering a culture of fairness and equity in AI development and deployment. The board would be crucial in mitigating racial biases in AI systems while promoting trust and confidence in these technologies. Additionally, there are no direct federal budget implications, making it a financially viable policy within NIST's existing budget, and representing a strategic investment in ethical AI practices. Establishing this board will provide the initial steps towards implementing race-conscious standards towards AI, in other industries such as industrial design and product development.

However, the policy also faces several challenges. Social acceptability is a significant concern, as some stakeholders might perceive this as government overreach into the private sector, potentially leading to resistance from private businesses. Additionally, the current political landscape is divided on race-related research and equity advancements may make a racial equity board controversial amongst the public, lowering the political feasibility of this policy. Furthermore, integrating new experts and aligning their work with existing processes may require considerable initial effort and resources.

While these concerns may be true, a similar initiative has successfully addressed these challenges. The Health Information Technology for Economic and Clinical Health (HITECH) Act of 2009 found a way to balance community priorities with important innovation. The HITECH Act was passed to enforce the security and privacy of confidential patient information through electronic health records (EHRs) (Burde 2011). HITECH garnered bipartisan support by leveraging needs from both types of stakeholders: citizens concerned with privacy and technology companies.

At the same time, additional reservations about this policy remain unresolved. For example, integrating new experts and aligning their work with existing processes may require considerable initial effort and resources. Furthermore, while the board aims to set vital standards for race-conscious AI, measuring its efficacy and ensuring its long-term sustainability and scalability across different industries and sectors may prove complex and challenging.

iii. Safeguards for AI systems and mandatory maintenance

The proposal brings to light important safeguards that must be implemented to prevent organizations, developers, and those who intend to use AI for healthcare applications from taking advantage of policy loopholes. One of the main techniques used to circumvent implemented standards is "jailbreaking." Another issue surrounding jailbreaking is the lack of transparency among AI developers on safeguards and processes they had implemented to minimize the risk of health disinformation, along with a lack of response to reported safeguard vulnerabilities. An important safeguard for AI systems is to ensure they adhere to the practices outlined in agreed-upon contracts throughout their research, development, and deployment stages.

On the positive side, standardizing assessments through regulatory agencies such as NIST and FDA ensures all datasets are treated equally, promoting fairness across the industry. This standardization and the establishment of robust safeguards help prevent the exploitation of regulatory loopholes, thereby ensuring higher integrity in AI applications. By ensuring transparency and maintenance, the policy contributes to the long-term success and reliability of AI models in healthcare. Additionally, enhancing transparency aligns with consumer expectations of equity, increasing trust in AI products.

Several organizations have been established with this purpose in mind. For instance, a regulatory framework could be modeled after Singapore's approach to AI governance. Singapore's AI Verify is a testing framework and software toolkit that allows organizations to validate the performance of their AI systems against established principles, such as transparency, through standardized tests. Through these tests, AI-focused companies can communicate the development of their innovations to shareholders and consumers (Singapore Personal Data Protection Commission). While these tests do not guarantee that AI systems are entirely "free from risks or biases", they are easily accessible and provide valuable insights.

There are notable objections to this approach. Implementing this policy requires significant bipartisan effort, including passing comprehensive regulations through appropriate agencies, which is a complex and time-consuming process. Conducting annual reviews and ensuring compliance can be resource and cost-intensive, diverting time and resources from other regulatory priorities. Organizations may also oppose the policy due to additional costs associated with compliance, including payments for transparency-certified reports and maintaining a team to ensure ongoing adherence to standards. Moreover, ensuring compliance and defining the operational scope for organizations and regulatory bodies present significant logistical challenges.

Without financial incentives, most AI developers and organizations will likely oppose it. Larger organizations may resist transparency due to increased costs and potential competitive disadvantages it could bring. Additionally, high regulatory and maintenance costs deter smaller organizations and freelance developers from contributing to open-source AI innovations. Policymakers could enhance the proposed policy through incentives and support mechanisms. These interventions could include tax breaks, subsidies, and tiered compliance requirements based on company size, using scaling methods similar to tax brackets.

Despite the considerable effort required to set up the necessary infrastructure, it establishes a strong foundation for the future of AI in medicine. The time needed for data reassessment and other maintenance tasks aligns with industry standards and consumer expectations of equity when interacting with AI products.

V. Conclusion

While AI innovation in healthcare has proven its potential to improve the patient experience, it can also produce harmful consequences, particularly by reinforcing racial biases. Biased systems pose a significant threat in the healthcare sector, underscoring the divide in healthcare accessibility and quality and combatting socioeconomic disadvantages.

We introduce three policy recommendations aligned with increasing transparency in AI healthcare applications, signifying a proactive approach toward addressing biases and promoting equity when improving the quality of care.

References

Agarwal, R., M. Bjarnadottir, L. Rhue, M. Dugas, K. Crowley, J. Clark, and G. Gao. 2022. "Addressing Algorithmic Bias and the Perpetuation of Health Inequities: An AI Bias Aware Framework." Health Policy and Technology 12 (1): 100702. https://doi.org/10.1016/j.hlpt.2022.100702.

Akinrinmade, Abidemi O., Temitayo M. Adebile, Chioma Ezuma-Ebong, Kafayat Bolaji, Afomachukwu Ajufo, Aisha O. Adigun, Majed Mohammad, et al. 2023. "Artificial Intelligence in Healthcare: Perception and Reality." *Cureus* 15 (9). https://doi.org/10.7759/cureus.45594.

Barto, Andrew, and Richard Sutton. 2007. "REINFORCEMENT LEARNING in ARTIFICIAL INTELLIGENCE." *Elsevier Science*. Elsevier Science. https://www.sciencedirect.com/science/article/pii/S0166411597801057?casa_token=o8AxBDIMhskAAAAA:BQVFih_6SiIlq_WzdXgaA1jEYKc3QDEih35I6DmFFDkYVmQUDFXGBeBxO6RJR5JBff4AGZJ3hA.

Bergomi, Laura, Tommaso M Buonocore, Paolo Antonazzo, Lorenzo Alberghi, Riccardo Bellazzi, Lorenzo Preda, Chandra Bortolotto, and Enea Parimbelli. 2024. "Reshaping Free-Text Radiology Notes into Structured Reports with Generative Question Answering Transformers." *Artificial Intelligence in Medicine* 154 (August): 102924–24. https://doi.org/10.1016/j.artmed.2024.102924

Burde, Howard. 2011. "THE HITECH ACT: An Overview." *AMA Journal of Ethics* 13 (3): 172–75. https://doi.org/10.1001/virtualmentor.2011.13.3.hlaw1-1103..

CDC (Centers for Disease Control and Prevention). 2024. "Social Determinants of Health." Public Health Professionals Gateway. May 14, 2024. https://www.cdc.gov/public-health-gateway/php/about/social-determinants-of-health.html.

Coalition for Health AI. 2023. "Blueprint for Trustworthy AI: Implementation Guidance and Assurance for Healthcare (Version 1.0)." Coalition for Health AI. https://coalitionforhealthai.org/papers/blueprint-for-trustworthy-ai_V1.0.pdf.

Collen, McClain, Michelle Faverio, Monica Anderson, and Eugenie Park. 2023. "Views of Data Privacy Risks, Personal Data and Digital Privacy Laws." *How Americans View Data Privacy* (blog). October 18, 2023. https://www.pewresearch.org/internet/2023/10/18/views-of-data-privacy-risks-personal-data-and-digital-privacy-laws/.

Daneshjou, Roxana, Kailas Vodrahalli, Roberto A. Novoa, Melissa Jenkins, Weixin Liang, Veronica Rotemberg, Justin Ko, et al. 2022. "Disparities in Dermatology AI Performance on a Diverse, Curated Clinical Image Set." *Science Advances* 8 (32): eabq6147. https://doi.org/10.1126/sciadv.abq6147.

Fazelpour, Sina, and David Danks. 2021. "Algorithmic Bias: Senses, Sources, Solutions." *Philosophy Compass* 16 (8). https://doi.org/10.1111/phc3.12760.

Federal Trade Commission (FTC). 1965. *Federal Cigarette Labeling and Advertising Act. 15 USC. §§ 1331-1340; 21 USC. § 387c*. https://www.ftc.gov/legal-library/browse/statutes/federal-cigarette-labeling-advertising-act.

Financial Industry Regulatory Authority. 2024. "2024 FINRA Annual Regulatory Oversight Report ." Financial Industry Regulatory Authority. January 9, 2024. https://www.finra.org/rules-guidance/guidance/reports/2024-finra-annual-regulatory-oversight-report.

Goldberg, Carey Beth, Laura Adams, David Blumenthal, Patricia Flatley Brennan, Noah Brown, Atul J. Butte, Morgan Cheatham, et al. 2024. "To Do No Harm — and the Most Good — with AI in Health Care." *NEJM AI* 1 (3). https://doi.org/10.1056/AIp2400036.

Hao, Karen. 2019. "This Is How AI Bias Really Happens—and Why It's so Hard to Fix." MIT Technology Review. February 4, 2019.

Hersch J, Shinall JB. Fifty years later: The legacy of the Civil Rights Act of 1964. Journal of Policy Analysis and Management. 2015;34(2):424–456. doi: 10.1002/pam.21824.

Lin, Ting-An, and Po-Hsuan Cameron Chen. 2022. "Artificial Intelligence in a Structurally Unjust Society." *Feminist Philosophy Quarterly* 8 (3/4). https://ojs.lib.uwo.ca/index.php/fpq/article/view/14191.

Lynch, Shana. 2024. "How Can We Better Regulate Health AI?" Stanford HAI. Stanford University. July 15, 2024. https://hai.stanford.edu/news/how-can-we-better-regulate-health-ai?mkt_tok=NTcwLVJDSC03NTgAAAGUx3IeIDAtQc_59Tsdh5U2-t6W5TUeEnj7dFVNga2IHiI3aMXrCueUIIdCwj0IUpCy-IO13jMwTMv4Ei76lXjpKcxLbycRSXKM22szpgI.

Maind, Sonali, and Priyanka Wankar. 2014. "Research Paper on Basic of Artificial Neural Network." *International Journal on Recent and Innovation Trends in Computing and Communication* 2 (1): 96–100.

Müller, Vincent. 2020. "Ethics of Artificial Intelligence and Robotics." In *Stanford Encyclopedia of Philosophy (Fall 2023 Edition)*. https://plato.stanford.edu/archives/fall2023/entries/ethics-ai.

National Institute of Standards and Technology (NIST). 2024. "Artificial Intelligence: The Vitals." National Institute of Standards and Technology (NIST). https://www.nist.gov/system/files/documents/2023/11/02/AI%20Fact%20Sheet%200615%20FINAL.pdf.

National Institute of Standards and Technology (NIST). 2024. "Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile," July. https://doi.org/10.6028/nist.ai.600-1.

Nazer, Lama H., Razan Zatarah, Shai Waldrip, Janny Xue Chen Ke, Mira Moukheiber, Ashish K. Khanna, Rachel S. Hicklen, et al. 2023. "Bias in Artificial Intelligence Algorithms and Recommendations for Mitigation." *PLOS Digital Health* 2 (6): e0000278. https://doi.org/10.1371/journal.pdig.0000278.

Obermeyer, Ziad, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. 2019. "Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations." Science 366 (6464): 447–53. https://doi.org/10.1126/science.aax2342.

Rodrigues, Rowena. 2020. "Legal and Human Rights Issues of AI: Gaps, Challenges and Vulnerabilities." *Journal of Responsible Technology* 4 (December):100005. https://doi.org/10.1016/j.jrt.2020.100005.

Rossy, Ryan. 2023. "Transforming Health Care with AI: A Conversation with Dr. Nishit Patel." USF Health. University of South Flordia. 2023. https://www.usf.edu/health/news/2023/ai-in-healthcare-dr-patel.aspx?utm_source=usfhealth_home&utm_medium=image-link&utm_content=small_image&utm_campaign=health%20home.

Samorani, Michele, Shannon L. Harris, Linda Goler Blount, Haibing Lu, and Michael A. Santoro. 2021. "Overbooked and Overlooked: Machine Learning and Racial Bias in Medical Appointment Scheduling." Manufacturing & Service Operations Management, August. https://doi.org/10.1287/msom.2021.0999.

Singapore Personal Data Protection Commission. n.d. "PDPC | Singapore's Approach to AI Governance." Www.pdpc.gov.sg. https://www.pdpc.gov.sg/Help-and-Resources/2020/01/Model-AI-Governance-Framework.

Shanklin, Robert, Michele Samorani, Shannon Harris, and Michael A. Santoro. 2022. "Ethical Redress of Racial Inequities in AI: Lessons from Decoupling Machine Learning from Optimization in Medical Appointment Scheduling." *Philosophy & Technology* 35 (4): 96. https://doi.org/10.1007/s13347-022-00590-8.

Shimotsu, Scott, Anne Roehrl, Maribet McCarty, Katherine Vickery, Laura Guzman-Corrales, Mark Linzer, and Nancy Garrett. 2015. "Increased Likelihood of Missed Appointments ('No Shows') for Racial/Ethnic Minorities in a Safety Net Health System." Journal of Primary Care & Community Health 7 (1): 38–40. https://doi.org/10.1177/2150131915599980.

Thomas, Mandisha. n.d. *Insurance; Use of Artificial Intelligence in Making Certain Decisions Regarding Coverage; Prohibit*. https://www.legis.ga.gov/legislation/65973.

United Kingdom Data Service. n.d.-a. "What Is the Five Safes Framework?" UK Data Service. https://ukdataservice.ac.uk/help/secure-lab/what-is-the-five-safes-framework/.

United Kingdom Data Service. n.d.-a. "What Is the UK Data Service SecureLab?" UK Data Service. https://ukdataservice.ac.uk/help/secure-lab/what-is-securelab/.

US Government Accountability Office. 2023. "Artificial Intelligence's Use and Rapid Growth Highlight Its Possibilities and Perils | US GAO." Www.gao.gov. September 6, 2023. https://www.gao.gov/blog/artificial-intelligences-use-and-rapid-growth-highlight-its-possibilities-and-perils.

US President. Executive Order. "Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, Executive order 14410 of November 1, 2023." *Federal Register* 88 no. 210 (November 1, 2023): 75191. https://www.govinfo.gov/content/pkg/FR-2023-11-01/pdf/2023-24283.pdf.

Waisel, D. B., & Truog, R. D. 1997. "Informed Consent". *Anesthesiology*, 87(4), 968–978. https://doi.org/10.1097/00000542-199710000-0003.

Mishayla Harve is an undergraduate student at the Georgia Institute of Technology in the School of Public Policy and the Department of Neuroscience. She is motivated to advocate for healthcare accessibility and community wellness.

Sakthi Priya Ramamoorthy is an undergraduate pre-medicine student at the Georgia Institute of Technology in the Department of Neuroscience. She is interested in addressing the gaps in policy so that healthcare can be more accessible to all people.

Viresh Pati is an undergraduate student at the Georgia Institute of Technology in the College of Computing. He is interested in artificial intelligence and its applications in finance, policy, and healthcare.

Garen Bainbridge is an undergraduate student at the Georgia Institute of Technology in the Department of Neuroscience. He is interested in researching the socioeconomic factors contributing to health disparities, especially human neuromuscular illnesses.

Abigayle Kankolenski is an undergraduate student at the Georgia Institute of Technology in the College of Computing and the School of Public Policy. With personal interests in AI, reading, policy, and politics, she is striving for a knowledgeable career in science and technology policy.

Bratee Podder is an undergraduate student at the Georgia Institute of Technology in the College of Computing. She is interested in ensuring that artificial intelligence is sufficiently regulated to contribute to the well-being of people.

Matthew Sampt is an undergraduate student at the Georgia Institute of Technology in the College of Computing. He has a special interest in the intersection between artificial intelligence and healthcare.

Acknowledgments
All authors are associated with the Health and Biotechnology Committee, as part of Georgia Institute of Technology's Organization of Science and Technology Policy Connections (S&T PC). S&T PC is dedicated to exploring policy related to technology and science. We would like to thank the guidance of the faculty and leadership of the organization– without them, we wouldn't be here today.