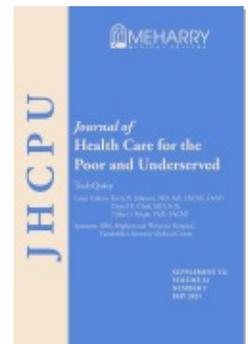




PROJECT MUSE®

A Proposed Framework on Integrating Health Equity and Racial Justice into the Artificial Intelligence Development Lifecycle

Irene Dankwa-Mullan, Elisabeth Lee Scheufele, Michael E. Matheny, Yuri Quintana, Wendy W. Chapman, Gretchen Jackson, Brett R. South



Journal of Health Care for the Poor and Underserved, Volume 32, Number 2, May 2021 Supplement, pp. 300-317 (Article)

Published by Johns Hopkins University Press

➔ For additional information about this article

<https://muse.jhu.edu/article/789672>

A Proposed Framework on Integrating Health Equity and Racial Justice into the Artificial Intelligence Development Lifecycle

*Irene Dankwa-Mullan, MD, MPH

*Elisabeth Lee Scheufele, MD, MS

Michael E. Matheny, MD, MS, MPH

Yuri Quintana, PhD

Wendy W. Chapman, PhD

Gretchen Jackson MD, PhD

Brett R. South, MS, PhD

Abstract: The COVID-19 pandemic has created multiple opportunities to deploy artificial intelligence (AI)-driven tools and applied interventions to understand, mitigate, and manage the pandemic and its consequences. The disproportionate impact of COVID-19 on racial/ethnic minority and socially disadvantaged populations underscores the need to anticipate and address social inequalities and health disparities in AI development and application. Before the pandemic, there was growing optimism about AI's role in addressing inequities and enhancing personalized care. Unfortunately, ethical and social issues that are encountered in scaling, developing, and applying advanced technologies in health care settings have intensified during the rapidly evolving public health crisis. Critical voices concerned with the disruptive potentials and risk for engineered inequities have called for reexamining ethical guidelines in the development and application of AI. This paper proposes a framework to incorporate ethical AI principles into the development process in ways that intentionally promote racial health equity and social justice. Without centering on equity, justice, and ethical AI, these tools may exacerbate structural inequities that can lead to disparate health outcomes.

Key words: Artificial Intelligence, AI ethics; health disparities, COVID-19, AI lifecycle, health equity principles.

Artificial intelligence (AI) applications have been widely deployed to understand, mitigate, and address pandemics, including the ongoing COVID-19 crisis.^{1,2} Ex-

IRENE DANKWA-MULLAN, ELISABETH LEE SCHEUFELE, GRETCHEN JACKSON, and BRETT R. SOUTH are affiliated with IBM Watson Health. **MICHAEL E. MATHENY** is affiliated with Vanderbilt University Medical Center. **YURI QUINTANA** is affiliated with Beth Israel Deaconess Medical Center and Harvard Medical School. **WENDY W. CHAPMAN** is affiliated with The University of Melbourne. Please address all correspondence to Irene Dankwa-Mullan, Center for AI, Research, and Evaluation (CARE), IBM Watson Health, 75 Binney Street, Cambridge, MA, 02142; phone: 1-720-396-0127; email: idankwa@us.ibm.com. *Denotes co-first authors

amples include case tracking, projecting virus transmission under different mitigation scenarios, forecasting mortality trends, and predicting disease outbreaks or hotspots.² The increases in computing capacity and AI-generative platforms, which can rapidly identify novel peptides, genes, and drug candidates, have accelerated the scientific discovery of COVID-19 vaccine candidates and medical therapies.^{2,3} With the ongoing global vaccine roll-out, AI-driven insights and applied interventions continue to play a significant role in adaptive and predictive technology. Some applications include tracking COVID-19 mutations and variants to inform vaccine design and development;^{4,5} predictive impact modeling for describing which populations and regions to vaccinate to rapidly flatten the curve and end the pandemic;⁶ monitoring the supply chain management and vaccine delivery;⁷ as well as post-vaccine surveillance to monitor adverse events and track effectiveness. The pandemic has provided opportunities for leveraging the rapidly evolving data and AI technologies to address this public health crisis. However, concerns about ethics, equity, and justice regarding the development and application of AI technologies in health care settings have intensified during the pandemic.^{1,2,8} The pandemic has been devastating, especially in Black and Hispanic communities that experience a mortality rate three times higher than White communities.⁹ National level data in the United States collected by the American Public Media (APM) Research Lab demonstrate that age-adjusted mortality rates for Black Americans, Indigenous peoples, Latinxs, and Pacific Islanders are 2.1, 2.2, 2.4, and 2.7 times higher than for Whites, respectively.⁹ While there have been great advances in personalized medicine and AI-based biomedical discovery based on genomic profiles, there is also a lack of diverse clinical research data used to generate those treatment strategies, which can result in worse outcomes for underserved members of the community.¹⁰⁻¹³ The rush for biomedical discovery with poorly representative COVID-19 databases may result in further inequities.¹⁴ With heightened visibility around structural racism, the discriminatory stereotypes created and reinforced with particular technologies, and biases reflected in algorithms are an increasing concern.¹⁴

This commentary provides a framework and recommendations to integrate health equity, racial justice, and ethical AI principles into technology development to address health inequities.

Prioritizing Health Equity and Racial Justice in the AI Development Lifecycle

Stakeholders in the design and development of AI technologies have a critical role in ensuring that mission-driven values to promote health equity are prioritized in implementing AI technologies. These technologies can influence payers, health providers, patient behaviors, and their experiences with the health care system in various ways. The application of machine learning to big data can identify patterns for improving health care delivery and decision-support tools can enable evidence-based care.¹⁵ In addition, AI has become a foundational element in many wearable technologies that support health maintenance or disease management.¹⁶

However, there are significant ethical and social concerns involved when designing, developing, and implementing AI tools and applications both domestically and

globally.^{17–21} Bias can be introduced into AI applications and affect numerous facets of an organized pandemic response (e.g., resource allocation and priority-setting, public health surveillance, contact tracing, patient privacy, frontline caregiving, health care worker privacy). Health equity and racial justice principles in applying AI, especially in the COVID-19 era, can provide a conceptual scaffold to ensure that efforts to track the virus, improve outcome predictions, and implement effective interventions will benefit all groups in a population for the current and future pandemics.

For the proposed framework, we define health equity as *the value and principle underlying a commitment to reduce and ultimately eliminate health disparities*.²² Addressing health equity, as asserted by Braveman, Marmot, and other scholars, is a *social justice issue and an ethical imperative, consonant with human rights principles to give special priority to act on significant public health problems that differentially affects those with fewer resources and/or may have more obstacles to achieving optimal health*.^{23,24} Broadly speaking, health disparities have been defined as systematic, unfair, plausibly avoidable differences in health (including its determinants and outcomes) negatively affecting socially vulnerable groups. These social groups are at risk of not achieving their full health potential because of historical discrimination, institutionalized racism, or marginalization (i.e., exclusion from social, political, or economic opportunities, including technologies), among other forces. When developing AI-based solutions in health care, anticipating and addressing potential health disparity concerns is imperative. These concerns must be consciously and appropriately accommodated, or health disparities among racial/ethnic minority and other socially vulnerable populations will continue to widen. Equity and justice principles in the continuum of AI design, development, and use are paramount and foundational. Similar to health equity, racial justice is a moral and value principle that promotes fair treatment of people of all races and ethnicities, resulting in equitable opportunities and outcomes.²⁵ Racial justice includes a deliberate effort to support and achieve racial equity through proactive and preventive measures. We will achieve racial equity when a person's racial or ethnic identity no longer predicts their social or economic opportunities and health outcomes. Simply denouncing or eliminating discrimination or stereotyping and bias is not sufficient to achieve racial justice. Instead, organizations and systems must re-imagine and co-create a different culture and society by implementing interventions that affect multiple sectors, processes, and practices.

Though AI ethics is accepted as critically important in harnessing AI's potential, there are disparate views and varying perspectives on critical ethical issues that inform the AI principles established within governments, the scientific research community, and industry.^{17,26–28} Several groups have attempted to summarize such ethical issues to inform policy statements.¹⁷ The Turing Institute defines AI ethics as a set of principles, values, and approaches that use widely accepted standards to guide moral conduct in the lifecycle of AI systems.^{29,30} The IBM Institute for Business Value defines AI ethics as a multidisciplinary field of study to understand how to optimize its beneficial impact while reducing risks and adverse outcomes for all stakeholders in a way that prioritizes human agency and well-being, as well as environmental flourishing.³¹ Artificial intelligence ethics research largely focuses on designing and building AI systems with an awareness of the values and principles to be followed during development—such

as data responsibility and privacy, fairness, inclusion, moral agency, value alignment, accountability, transparency, trust, and technology misuse.^{32–42} These frameworks and statements can be aligned with health equity and racial justice principles. As a part of the efforts to embrace racial and social justice, the IBM Academy of Technology and other Justice and Diversity Councils have launched initiatives to replace terminology that promotes racial and cultural bias, to promote design justice for racial equity, and to integrate equity and inclusive principles across the solutions.⁴³

This paper compiles the range of ethical issues that inform guidelines and propose examples of how health equity and racial justice might be aligned with AI ethics (see Box 1). The paper also builds on the AI development lifecycle and provides a framework with recommendations for operationalizing ethical AI with health equity and racial justice principles.

Unintended Consequences of Limited Health Equity or Racial Justice Deliberation in AI Development

Although ethical statements are being issued by governments, academics, policymakers and regulators in response to the growing visibility of advanced technologies, the number of AI and algorithmic systems with limited equity and justice considerations continues to increase. There are several ways in which AI systems, including the data and evidence on which they are trained, can cause harm, each with ethical, social, and equity implications. The accuracy and quality of the databases and the sometimes inconclusive or misguided evidence on which the algorithms are developed and implemented, shape decisions that have detrimental and adverse outcomes. A lack of explainability of data sources and transparency as well as design bias and limited evidence in the algorithms for AI, suggest how these issues are intertwined. The result is an exacerbation of structural inequities and adverse outcomes when disadvantaged populations are not included in trial data.^{29,44}

Another unfortunate consequence in product development is the mismatch of the intended use and subsequent actual use. This could happen when there is lack of accountability and moral agency for the entire process from design and development to implementation. For instance, consider an AI tool that may have been developed to identify a population to target with an intervention. Instead, the tool's use may result in discrimination against patients based on factors emphasized in the AI tool, thus influencing future treatment and reimbursement decisions and producing adverse downstream patient outcomes.⁴⁵ Documenting how the dataset was created, curated, validated, implemented, and shared will be important to the development of clinical care guidelines and clinical trials.⁴⁶ The AI Now Institute at New York University created the algorithmic impact assessment to provide awareness and improve processes to identify the potential harms of machine learning algorithms.⁴⁷

In another example, AI-supported clinical decision-support systems may be applied beyond the appropriate scope of use in under-resourced provider or patient settings with unintended consequences.⁴⁸ Human oversight and workflow integration are critical to safety, especially in settings where clinical experts are using clinical decision support systems (CDS) and other technologies and can help avoid harm to vulnerable

Box 1.

ALIGNING HEALTH EQUITY AND RACIAL JUSTICE PRINCIPLES WITH AI ETHICS ISSUES ENCOUNTERED IN AI DEVELOPMENT AND IMPLEMENTATION.

AI Ethics Issue	Health Equity Principles	Racial Justice Principles
Accountability	Promote accountability to the goals and actions of ensuring equity and acknowledge the values for those groups to which the AI tool will be responsible.	Promote accountability to the goals and values of racial and ethnic population groups, ensuring a transparent agenda and purpose to improve health outcomes.
Algorithms	Promote Algorithmic Impact Assessments (AIA) to understand the health and social implications of AI technologies. Incorporate health equity and related disciplines' perspectives to understand social context and develop benchmarks and debiasing approaches for AI models and statistical fairness performance.	Support AIAs in identifying coded racial inequities, preventing algorithmic discrimination and bias. Address hierarchical logic that could produce default discrimination. Train models on diverse data and to be transferable to other relevant settings.
Data—Data Responsibility, Training, and Processing	Promote data responsibility that includes ensuring data diversity and inclusion of social determinants of health data in training and processing. Promote assessment of training datasets that are best suitable for diverse populations and establish standards for documenting how the dataset was created, curated, validated, implemented, and shared.	Ensure data governance, data sovereignty and ownership protocols, the inclusion of race and ethnicity data, and cultural heritage. Support Global Indigenous Data Alliance (GIDA) CARE principles of Collective benefit, Authority to control, Responsibility, and Ethics.

(continued on p. 305)

Box 1. (continued)

AI Ethics Issue	Health Equity Principles	Racial Justice Principles
Design	Ensure equity-centered design thinking that promotes humility-building and self-awareness of biases, addressing power dynamics, and co-creating with community and stakeholders	Ensure design justice, promote inclusion of racial and ethnic groups, prioritize user needs, values, and impact in the design and development of AI
Discrimination	Ensure AI-tools and interventions that promote fair and equal treatment for all members of various groups (race, gender, social class, sexual orientation, physical ability, and other categories). Many forms of discrimination perpetuate inequities, such as invisibility, exclusion, or complacency employed to avoid detection, critique, or questioning.	Ensure AI-tools and interventions that promote equal treatment and opportunities for all racial and ethnic minority groups.
Empathy	Promote visions and values of empathy for socially disadvantaged and marginalized population groups in health care based on fairness, justice, and democracy.	Promote racial empathy and understand cultural perspectives to help inform ethical inclusive AI.
Explainability	Promote explainable AI systems that include how the methods address health equity and improved health outcomes.	Promote explainable AI systems that achieve a pragmatic understanding of how the system addresses racial equity and improved health outcomes.
Fairness, Non-Discrimination	Expand on fairness definitions to include health equity and social justice.	Expand on fairness definitions to include racial justice issues.
Human Oversight	Ensure that human-centered AI allows for oversight and the human dimension in all aspects of the AI development and implementation.	Ensure accountability and human oversight across the AI development process to ensure racial equity.

(continued on p. 306)

Box 1. (continued)

AI Ethics Issue	Health Equity Principles	Racial Justice Principles
Human Autonomy	Support and respect human autonomy and at the same time understand how individual autonomy occurs within social, economic, and structural contexts.	Respect individual autonomy for racial and ethnic populations, and at the same time understand how the consequences may impact use of AI within their social contexts.
Inclusion—Perspectives	Promote diversity in technology and AI and ensure inclusion of under-represented groups or traditionally excluded individuals or groups into processes, activities, and decisions in ways that value perspectives equally. The expertise of patient groups, caregivers, nurses, and community health workers is important.	Promote inclusion of racial and ethnic populations into the processes, activities, and decision-making and policies.
Inclusion—Social Cohesion, Solidarity	Support social cohesion as an essential driver of ethical, inclusive technology.	Support structural changes to address systemic racism and lack of diversity through a collaborative approach across sectors.
Inclusive Technology	Promote efforts to address engineered inequities, and promote the use of inclusive language.	Address racial justice with technologies. Identify and replace terminology that promotes racial and cultural bias.
Moral Agency	Promote an understanding of the social, historical, and economic context and impact decision-making and health outcomes. Understand the translation of the evidence to policy and impact on health equity.	Address racial and social complexities in the scientific and public health field to protect racial and ethnic populations from risks and harms with AI technologies.

(continued on p. 307)

Box 1. (continued)

AI Ethics Issue	Health Equity Principles	Racial Justice Principles
Privacy Protection	Ensure data protection and systemic governance laws regarding individual rights and right to access one's data. Ethically extend Health Insurance Portability and Accountability Act enforcement.	Ensure data and privacy protection laws to prevent harm against racial and ethnic populations.
Robustness, Repeatability	Promote interdisciplinary research that integrates health disparities science where relevant and ensures robustness in AI methodology.	Ensure robust, reproducible AI using diverse data that includes race and ethnicity.
Safety, Cybersecurity	Confront inherent security vulnerabilities in AI systems and impact on health equity.	Ensure safeguards or accountability structures and policies to ensure safety and privacy for racial and ethnic groups. Account for and ensure against potential risks and harms.
Transparency	Prioritize and focus on openness and transparency, accountability, and oversight in the development of technology.	Ensure transparency in the data, models, and use of AI technologies.
Trust	Foster long-term relationships with users, build trust, and promote active inclusion and engagement of communities of color in the AI development process.	Acknowledge reasons for mistrust from racial and ethnic communities and use that to inform inclusive engagement.
Value Alignment	Promote health equity as a social justice issue.	Promote racial equity as a moral and ethical imperative.

populations. Users of AI must maintain accountability when adverse effects arise, especially as some AI applications are maturing to full automation, such as the Apple Watch EKG app that received FDA clearance.⁴⁹ Artificial intelligence should generally be considered *augmented intelligence* to ensure that providers and patients are the final shared decision-makers.

Addressing algorithmic bias and ensuring data diversity have not been consistent in AI technologies' design and development. The AI development lifecycle should employ a strategic approach that considers health equity and ethical principles in managing the data, model-building, training, and deployment from conception to implementation. Gaps in the current data science and machine learning methods include addressing health equity and racial justice as fundamental requirements. Lifecycle processes that overlook health disparities may promulgate and perpetuate bias. Data often incompletely represent a target population.⁵⁰ The data and knowledge sources used to inform AI technologies require rigorous evaluation to ensure clinical performance, analytical performance, and scientific validity, promoting fairness and equitable outcomes. The black-box nature of AI technologies can act as a barrier to adoption. Unintended exacerbation of biases will be perpetuated if the output is not easily understandable or applicable to the user.⁵¹ For example, a tool that predicted a seven-day mortality risk or disease progression in a high-risk subpopulation might become outdated as new science, data, evidence, or methods evolve. Thus, it is essential to put humans in the loop for accountability in decisions that affect patient care.^{52,53}

A social concern is the impact of AI on patient-provider relationships. The human touch, empathy, understanding, and judgment are critical components of healing and patient care. Since positive health care encounters are built on relationships with patients, caregivers, and families, automated decisions or recommendations from an AI tool or algorithm can introduce new and possibly complicating elements into these interactions. Additionally, algorithms trained on and dependent on measurable data may not always capture relevant environmental information, social data, or patient cultural beliefs, preferences, and values. Social determinants of health (SDoH) such as educational level, economic insecurity, and other social factors contribute up to 40% towards determining health outcomes.⁵⁴⁻⁵⁷ Another issue is the effect that AI may have on jobs and the potential task-shifting that comes with automation.¹⁰⁻¹² On a broader scale, the foundational evidence for AI tools must include all relevant populations' data to inform appropriate health equity interventions or decision-making.⁴⁶

The Lifecycle of AI Development in Health Care

Widespread implementation and application of AI in health care have lagged behind expectations due to several factors,⁵⁸ including a lack of robust, integrated data, inadequate trust to foster adoption, notable missteps in consideration of biases, disparities in expected targeted outcomes,^{59,60} and challenges in integrating AI into complex workflows. In 2020, the National Academy of Medicine published a special publication on AI in Healthcare.⁶¹ One of the focus areas was a synthesis of best practices for developing, implementing, and maintaining AI systems used in delivering health care, summarized into a lifecycle framework (Figure 1, below). The AI development lifecycle is a continuous

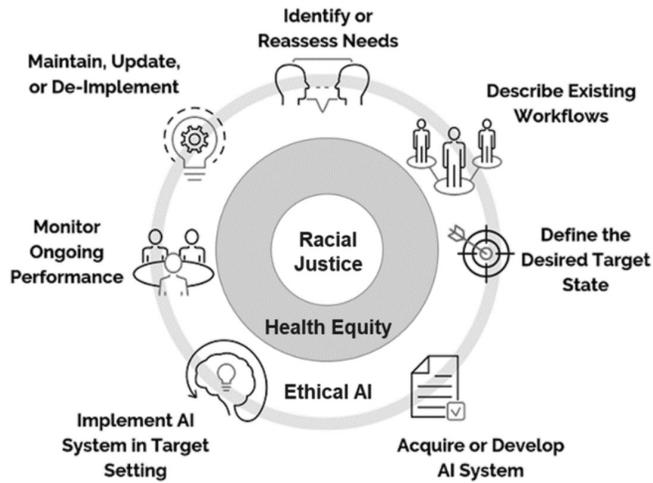


Figure 1. Ethical AI, Health Equity, and Racial Justice integrated across the Lifecycle of AI development.

Note:

Lifecycle phases (outer circle) adopted From the National Academy of Medicine, 2019, *AI in Health Care: The Hope, the Hype, the Promise, the Peril*. Reprinted with permission from the National Academy of Sciences, Courtesy of the National Academies Press, Washington, D.C.

process that begins by assessing needs, describing existing workflows, identifying and defining target states, acquiring infrastructure to develop the AI system, implementing the system, monitoring and evaluating performance, and maintaining, updating, or replacing the system when gaps or new needs arise. The lifecycle of an AI technology can provide a framework to identify opportunities to ensure that health disparity and social justice concerns are integrated into the genesis and application of AI solutions in public health and health care. Integrating health equity and racial justice principles into AI development requires building a responsible culture in innovation and establishing ethical building blocks for the reliable delivery of equitable AI technology.

Practical Applications of Health Equity and Racial Justice in AI Lifecycle Frameworks

We propose a framework in developing AI by incorporating health equity and racial justice principles into the different components of the AI lifecycle in health care. The proposed framework, shown in Figure 2, provides suggestions for every step of the lifecycle to consider equity and inclusivity and guard against biases.

The lifecycle of an AI technology can provide a framework to identify opportunities to ensure health disparity concerns are integrated into the genesis and application of AI solutions in public health and health care. The AI development lifecycle is a continuous process that begins by assessing needs, describing existing workflows, identifying and defining target states, acquiring infrastructure to develop the AI system, implementing the system, monitoring and evaluating performance, and maintaining, updating, or replacing the system when gaps or new needs arise.

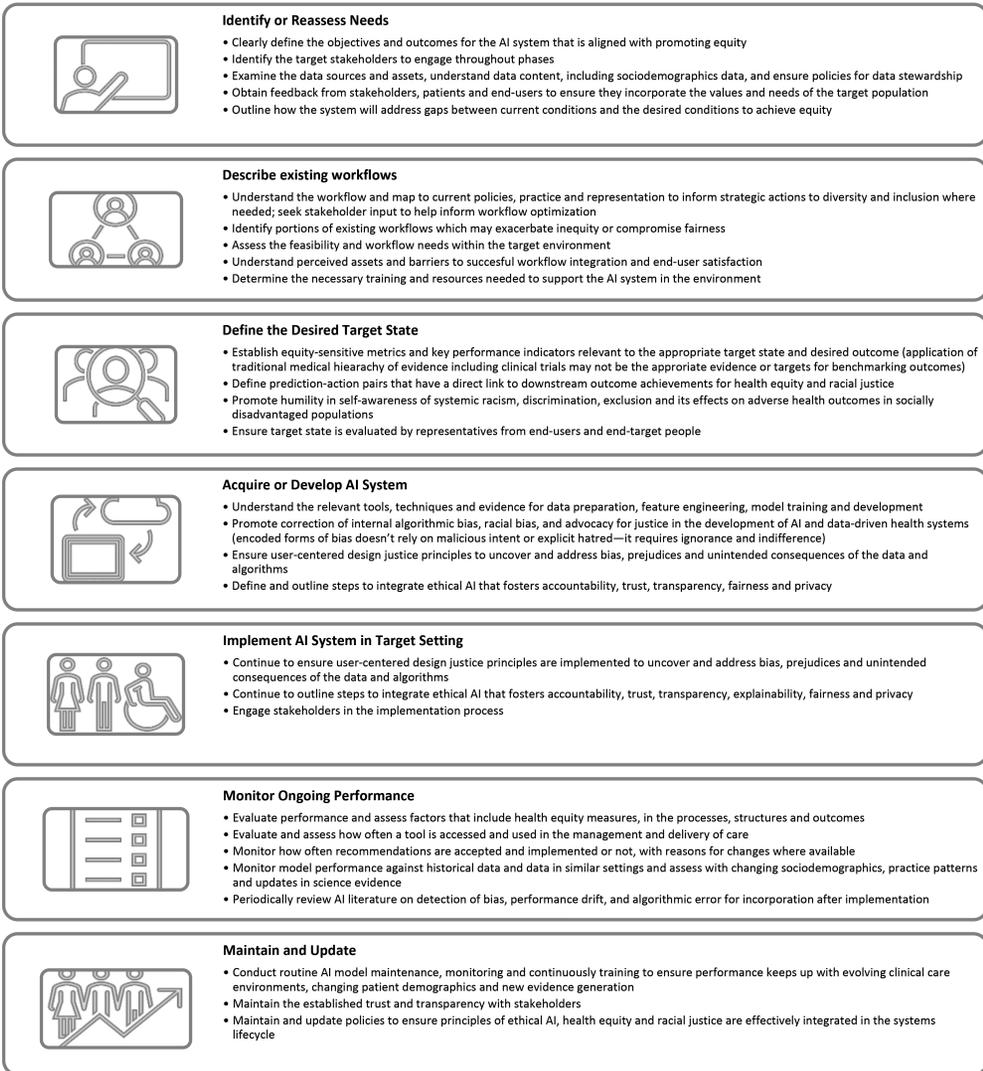


Figure 2. Framework for Integrating Health Equity and Racial Justice into AI Development.

Note:

Lifecycle of AI is from the National Academy of Medicine. 2019. AI in Health Care: The Hope, the Hype, the Promise, the Peril. Adapted with permission from the National Academy of Sciences, Courtesy of the National Academies Press, Washington, D.C.

In the context of ensuring that equity and fairness are central to the lifecycle, and aligned with what has been dubbed the *Quintuple Aim*,⁶¹ the first step of this framework includes identifying or reassessing needs that involve stakeholder, patient, and end-user engagement to ensure incorporating values of the target population. In this step, activities include defining objectives for an AI system that is aligned with promoting equity, including identifying data assets, data content, and policies for data stewardship.

The second step in this framework focuses on describing existing workflows and

their effects on existing needs in policies, practice, feasibility, and workflow, as well as assessment of barriers and understanding the necessary training and resources to support the AI system.

The third step in this framework deals with the need to define desired target states. This step includes activities to establish the equity-sensitive metrics and key performance metrics related to the target outcomes. This step seeks to promote humility and self-awareness of systemic racism, discrimination, exclusion, and its effects on adverse health outcomes in socially disadvantaged populations.

The fourth step in this framework focuses on the task of acquiring and developing the AI system itself. Central to this step is understanding the relevant tools, techniques, and methods for data preparation, feature engineering, model training, and development. This step aims to promote and correct internal algorithmic bias in a way that advocates for justice in the development of AI and data-driven health systems by ensuring user-centered design justice principles are employed to uncover and address racial bias prejudices and unintended consequences of data and algorithms. Thus, defining and outlining steps are needed to integrate ethical AI fostering accountability, trust, transparency, fairness, and privacy and ensuring user-centered design justice principles to uncover and address bias, prejudices, and unintended consequences of the data and algorithms.

The fifth step in the framework focuses on implementing the AI system in the target setting and engaging with stakeholders, patients, and end-users in the implementation process in a way that fosters accountability, trust, transparency, explainability, fairness, and privacy.

The sixth step in the framework involves monitoring ongoing system performance to assess factors that include health equity measures in the processes, structures, and outcomes. These metrics include assessing how often the tool is accessed and used in the management and delivery of care, monitoring how often recommendations are accepted and implemented or not, and reasons for changes. Central to this is the requirement to monitor system performance against historical data and data generated in similar settings to assess changes in socio-demographics, practice patterns, and updates to scientific evidence and real-world data.

The seventh and final step in the framework involves activities focusing on maintaining and updating the system by conducting routine AI model maintenance and continuous training to ensure system performance reflects evolving clinical care environments, changing patient demographics, and new evidence being generated. Maintaining established trust and transparency with stakeholders and continuously maintaining and updating policies to ensure ethical AI principles, health equity, and racial justice are integrated in the system lifecycle.

Conclusions

During a public health crisis, AI's application holds great promise for augmenting decision-making, allocating scarce resources, and aiding in decision-making and policy formulation. Challenges persist in applying AI systems that can cause involuntary and unintended harm with profound ethical and social consequences. Merging the health

equity and racial justice principles with AI Lifecycle provides a framework, approach, and a set of ethical values, principles, and techniques to guide moral conduct in the development of AI systems. Despite increasingly accurate AI tools, limited evidence exists on their applicability in real-world settings. One reason is the gap between proof-of-concept testing and clinical validation. For example, there is a clear process of scientific evaluation in drug development by which regulatory approval is achieved. Although many AI tools in health care are not regulated, a similar framework has been proposed for a systematic and comprehensive AI evaluation in health care to allow safe and effective adoption.⁶² The adoption of this framework and strategy, guided by justice principles, will support algorithm and tool developers, health systems, and researchers in creating user-driven innovations that fit within clinical workflows, facilitate interoperable information exchange, and evaluate AI in real-world health settings and proactively mitigate risks of exacerbating existing health disparities.

Acknowledgments

Funding Statements: This research study was supported by IBM Watson Health. **Competing Interests:** The authors are employed by IBM Watson Health, Cambridge, Mass., USA; Vanderbilt University Medical Center, Nashville Tenn., USA; Beth Israel Deaconess Medical Center, Boston, Mass., USA; Harvard Medical School, Boston, Mass., USA; The University of Melbourne, Victoria, Australia.

References

1. Latif S, Usman M, Manzoor S, et al. Leveraging data science to combat COVID-19: a comprehensive review. *IEEE Transactions on Artificial Intelligence*. 2020 Aug;1(1): 85–103.
<https://doi.org/10.36227/techrxiv.12212516.v2>
2. Vaishya R, Javaid M, Khan IH, et al. Artificial Intelligence (AI) applications for COVID-19 pandemic. *Diabetes Metab Syndr*. 2020 Jul–Aug;14(4):337–9. Epub 2020 Apr 14.
<https://doi.org/10.1016/j.dsx.2020.04.012>
PMid:32305024
3. Keshavarzi Arshadi A, Webb J, Salem M, et al: Artificial intelligence for COVID-19 drug discovery and vaccine development. *Front Artif Intell*. 2020;3.
<https://doi.org/10.3389/frai.2020.00065>
4. Malone B, Simovski B, Moline C, et al. Artificial intelligence predicts the immunogenic landscape of SARS-CoV-2 leading to universal blueprints for vaccine designs. *Sci Rep*. 2020 Dec 23;10(1):22375.
<https://doi.org/10.1038/s41598-020-78758-5>
PMid:33361777
5. Yang Z, Bogdan P, Nazarian S. An in silico deep learning approach to multi-epitope vaccine design: a SARS-CoV-2 case study. *Sci Rep*. 2021 Feb 5;11(1):3238.
<https://doi.org/10.1038/s41598-021-81749-9>
PMid:33547334
6. Shur M. Pandemic equation for describing and predicting COVID19 evolution. *J Healthc Inform Res*. 2021 Jan 7;1–13.

- <https://doi.org/10.1007/s41666-020-00084-2>
PMid:33437912
7. Hariharan R, Sundberg J, Gallino G, et al. An interpretable predictive model of vaccine utilization for Tanzania. *Front Artif Intell.* 2020;3.
<https://doi.org/10.3389/frai.2020.559617>
 8. Hu Y, Jacob J, Parker GJM, et al. The challenges of deploying artificial intelligence models in a rapidly evolving pandemic. *Nat Mach Intell.* 2020 May;2:298–300.
<https://doi.org/10.1038/s42256-020-0185-2>
 9. APM Research Lab. COVID-19 deaths were analyzed by race and ethnicity. St Paul, MN: APM Research Lab, 2020. Available at <https://www.apmresearchlab.org/covid/deaths-by-race-december2020>.
 10. Bentley AR, Callier S, Rotimi CN. Diversity and inclusion in genomic research: why the uneven progress? *J Community Genet.* 2017 Oct;8(4):255–66. Epub 2017 Jul 18.
<https://doi.org/10.1007/s12687-017-0316-6>
PMid:28770442
 11. Bentley AR, Callier SL, Rotimi CN. Evaluating the promise of inclusion of African ancestry populations in genomics. *NPJ Genom Med.* 2020 Feb 25;5:5.
<https://doi.org/10.1038/s41525-019-0111-x>
PMid:32140257
 12. Hindorff LA, Bonham VL, Brody LC, et al. Prioritizing diversity in human genomics research. *Nat Rev Genet.* 2018 Mar;19(3):175–85. Epub 2017 Nov 20.
<https://doi.org/10.1038/nrg.2017.89>
PMid:29151588
 13. Wapner J. Cancer scientists have ignored African DNA in the search for cures. New York, NY: Newsweek, 2018. Available at <https://www.newsweek.com/2018/07/27/cancer-cure-genome-cancer-treatment-africa-genetic-charles-rotimi-dna-human-1024630.html>.
 14. Roosli E, Rice B, Hernandez-Boussard T. Bias at warp speed: how ai may contribute to the disparities gap in the time of COVID-19. *J Am Med Inform Assoc.* 2021 Jan 15;28(1):190–2.
<https://doi.org/10.1093/jamia/ocaa210>
PMid:32805004
 15. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med.* 2019 Jan;25(1):44–56. Epub 2019 Jan 7.
<https://doi.org/10.1038/s41591-018-0300-7>
PMid:30617339
 16. Wu M, Luo J. Wearable technology applications in healthcare: a literature review. *Online J Nurs Inform.* 2019;23(3).
 17. Hagedorff T. The ethics of AI ethics: an evaluation of guidelines. *Minds Mach (Dordr).* 2020;30:99–120.
<https://doi.org/10.1007/s11023-020-09517-8>
 18. Jobin A, Ienca M, Yayena E. The global landscape of AI ethics guidelines. *Nat Mach Intell.* 2019 Sep;1:389–99.
<https://doi.org/10.1038/s42256-019-0088-2>
 19. The Lancet. Artificial intelligence in global health: a brave new world. *Lancet.* 2019 Apr 13;393(10180):1478.
[https://doi.org/10.1016/S0140-6736\(19\)30814-1](https://doi.org/10.1016/S0140-6736(19)30814-1)
 20. Gonzalez Alarcon N, Pombo C. ¿Cómo puede la inteligencia artificial ayudar en una pandemia? (How can artificial intelligence help in a pandemic?) Washington,

- DC: Interamerican Development Bank, 2020. Available at <https://fairlac.iadb.org/en/ia-covid>.
21. United States Agency for International Development. Artificial intelligence in global health. Washington, DC: United States Agency for International Development, 2019. Available at https://www.usaid.gov/sites/default/files/documents/1864/AI-in-Global-Health_webFinal_508.pdf.
 22. Braveman PA, Kumanyika S, Fielding J, et al. Health disparities and health equity: the issue is justice. *Am J Public Health*. 2011 Dec;101 Suppl 1(Suppl 1):S149–55. Epub 2011 May 6.
<https://doi.org/10.2105/AJPH.2010.300062>
PMid:21551385
 23. Marmot M, Allen JJ. Social determinants of health equity. *Am J Public Health*. 2014 Sep;104 Suppl 4(Suppl 4):S517–9.
<https://doi.org/10.2105/AJPH.2014.302200>
PMid:25100411
 24. Penman-Aguilar A, Talih M, Huang D, et al. Measurement of health disparities, health inequities, and social determinants of health to support the advancement of health equity. *J Public Health Manag Pract*. 2016 Jan–Feb;22 Suppl 1(Suppl 1):S33–42.
<https://doi.org/10.1097/PHH.0000000000000373>
PMid:26599027
 25. Jones CP. Toward the science and practice of anti-racism: launching a national campaign against racism. *Ethn Dis*. 2018 Aug 9;28(Suppl 1):231–4.
<https://doi.org/10.18865/ed.28.S1.231>
PMid:30116091
 26. Crawford K, Dobbe R, Dryer T, et al. AI Now 2019 Report. New York, NY: AI Now Institute, 2019. Available at https://ainowinstitute.org/AI_Now_2019_Report.html.
 27. Fjeld J, Achten N, Hilligoss H, et al. Principled artificial intelligence: mapping consensus in ethical and rights-based approaches to principles for AI. Cambridge, MA: Berkman Klein Center, 2020.
<https://doi.org/10.2139/ssrn.3518482>
 28. Floridi L, Cowls J. A unified framework of five principles for AI in society. *Harv Data Sci Rev*. 2019;1(1).
<https://doi.org/10.1162/99608f92.8cd550d1>
 29. Char DS, Shah NH, Magnus D. Implementing machine learning in health care—addressing ethical challenges. *N Engl J Med*. 2018 Mar 15;378(11):981–3.
<https://doi.org/10.1056/NEJMp1714229>
PMid:29539284
 30. Leslie D. Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector. London, England, United Kingdom: The Alan Turing Institute, 2019.
<https://doi.org/10.2139/ssrn.3403301>
 31. Goehring B, Rossi F, Zaharchuk D. Advancing AI ethics beyond compliance: from principles to practice. Armonk, NY: IBM Corporation, 2020. Available at <https://www.ibm.com/downloads/cas/J2LAYLOZ>.
 32. Ribeiro M, Wu T, Guestrin C, et al. Beyond Accuracy: Behavioral Testing of NLP Models with CheckList. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Online, July 2020:4902–12. Stroudsburg, PA: ACL Anthology, 2020.
<https://doi.org/10.18653/v1/2020.acl-main.442>

33. Madaio M, Stark L, Vaughan J, et al. Co-designing checklists to understand organizational challenges and opportunities around fairness in AI. In: 2020 CHI Conference on Human Factors in Computing Systems (CHI '20), Honolulu (HI), April 2020:1–14. New York, NY: Association for Computing Machinery, 2020. <https://doi.org/10.1145/3313831.3376445>
34. Tsamados A, Aggarwal N, Cowls J, et al. The Ethics of Algorithms: Key Problems and Solutions. Rochester, NY: SSRN, 2020. <https://doi.org/10.2139/ssrn.3662302>
35. Leins K, Lau J, Baldwin T. Give me convenience and give her death: who should decide what uses of nlp are appropriate, and on what basis? In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Online, July 2020: 2908–2913. Stroudsburg, PA: ACL Anthology. 2020. <https://doi.org/10.18653/v1/2020.acl-main.261>
36. Kaushik D, Hovy E, Lipton Z: Learning the difference that makes a difference with counterfactually-augmented data. arXiv:1909.12434v2 [cs.CL]. Ithaca, NY: Cornell University, 2020.
37. Bhatt U, Andrus M, Weller A, et al. Machine learning explainability for external stakeholders. arXiv:2007.05408 [cs.CY]. Ithaca, NY: Cornell University, 2020.
38. Kuhlman C, Jackson L, Chunara R. No computation without representation: Avoiding data and algorithm biases through diversity. arXiv:2002.11836 [cs.CY]. Ithaca, NY: Cornell University, 2020. <https://doi.org/10.1145/3394486.3411074>
39. Sap M, Gabriel S, Qin L, et al. Social Bias Frames: reasoning about social and power implications of language. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Online, July 2020: 5477–90. Stroudsburg, PA: ACL Anthology. 2020. <https://doi.org/10.18653/v1/2020.acl-main.486>
40. Lloreda C. Speech recognition tech is yet another example of bias. New York City, NY: Scientific American, 2020. Available at <https://www.scientificamerican.com/article/speech-recognition-tech-is-yet-another-example-of-bias/>.
41. Stray J, Adler S, Hadfield-Menell D. What are you optimizing for? Aligning Recommender systems with human values. In: Participatory Approaches to Machine Learning: ICML 2020 Workshop, Online, July 17, 2020. Available at <https://participatoryml.github.io/papers/2020/42.pdf>.
42. Fiesler C, Garrett N, Beard N. What do we teach when we teach tech ethics? A syllabi analysis. In: The 51st ACM Technical Symposium on Computer Science Education (SIGCSE '20). Portland (OR), February 2020:289–95. New York, NY: Association for Computing Machinery (ACM), 2020. <https://doi.org/10.1145/3328778.3366825>
43. Jones D, Humphrey T. Words matter: driving thoughtful change toward inclusive language in technology. Armonk, NY: IBM Corporation, 2020. Available at <https://www.ibm.com/blogs/think/2020/08/words-matter-driving-thoughtful-change-toward-inclusive-language-in-technology/>.
44. Esmaeilzadeh P: Use of AI-based tools for healthcare purposes: a survey study from consumers' perspectives. BMC Med Inform Decis Mak. 2020 Jul 22;20(1):170. <https://doi.org/10.1186/s12911-020-01191-1> PMID:32698869
45. Obermeyer Z, Powers B, Vogeli C, et al. Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019 Oct 25;366(6464):447–53.

- <https://doi.org/10.1126/science.aax2342>
PMid:31649194
46. Topol EJ. Welcoming new guidelines for AI clinical research. *Nat Med.* 2020 Sep;26(9):1318–20.
<https://doi.org/10.1038/s41591-020-1042-x>
PMid:32908274
 47. Reisman D, Schultz J, Crawford K, et al. Algorithmic impact assessments: a practical framework for public agency accountability. New York, NY: AI Now Institute, 2018. Available at <https://ainowinstitute.org/aiareport2018.pdf>.
 48. Rigby M. Ethical dimensions of using artificial intelligence in health care. *AMA J Ethics.* 2019;21(2):E121–4.
<https://doi.org/10.1001/amajethics.2019.121>
 49. Raja JM, Elsagr C, Roman S, et al. Apple watch, wearables, and heart rhythm: where do we stand? *Ann Transl Med.* 2019 Sep;7(17):417.
<https://doi.org/10.21037/atm.2019.06.79>
PMid:31660316
 50. Gianfrancesco MA, Tamang S, Yazdany J, et al. Potential biases in machine learning algorithms using electronic health record data. *JAMA Intern Med.* 2018 Nov 1;178(11):1544–7.
<https://doi.org/10.1001/jamainternmed.2018.3763>
PMid:30128552
 51. Batahee Y. The artificial intelligence black box and the failure of intent and causation. *Harvard Journal of Law and Technology.* 2018 Spr;31(2):890–938.
 52. McCradden MD, Joshi S, Mazwi M, et al. Ethical limitations of algorithmic fairness solutions in health care machine learning. *Lancet Digit Health.* 2020 May;2(5):e221–3.
[https://doi.org/10.1016/S2589-7500\(20\)30065-0](https://doi.org/10.1016/S2589-7500(20)30065-0)
 53. Nordling L. A fairer way forward for AI in health care. *Nature.* 2019 Sep;573(7775):S103–5.
<https://doi.org/10.1038/d41586-019-02872-2>
PMid:31554993
 54. Social determinants of health: know what affects health. Atlanta, GA: Centers for Disease Control and Prevention, 2021. Available at <https://www.cdc.gov/social-determinants/index.htm>.
 55. University of Wisconsin Population Health Institute. County health rankings & roadmaps. Madison, WI: University of Wisconsin Population Health Institute, 2021.
 56. Magnan S. Social Determinants of Health 101 for Health Care: Five Plus Five. *NAM Perspectives.* Washington, DC: National Academy of Medicine, 2017.
<https://doi.org/10.31478/201710c>
 57. Robert Wood Johnson Foundation. Using social determinants of health data to improve health care and health: a learning report. Princeton, NJ: Robert Wood Johnson Foundation, 2016. Available at <https://www.rwjf.org/en/library/research/2016/04/using-social-determinants-of-health-data-to-improve-health-care-.html>.
 58. Pierson E, Cutler DM, Leskovec J, et al. An algorithmic approach to reducing unexplained pain disparities in underserved populations. *Nat Med.* 2021 Jan;27(1):136–40. Epub 2021 Jan 13.
<https://doi.org/10.1038/s41591-020-01192-7>
PMid:33442014
 59. Kusner MJ, Loftus JR. The long road to fairer algorithms. *Nature.* 2020 Feb;578(7793):34–6.

- <https://doi.org/10.1038/d41586-020-00274-3>
PMid:32020122
60. Cheng F, Kovacs IA, Barabasi AL. Network-based prediction of drug combinations. *Nat Commun.* 2019 Mar 13;10(1):1197.
<https://doi.org/10.1038/s41467-019-09186-x>
PMid:30867426
61. Matheny M, Israni ST, Auerbach A, et al. Artificial intelligence in health care: the hope, the hype, the promise, the peril. Washington, DC: National Academy of Medicine, 2020. Available at <https://nam.edu/artificial-intelligence-special-publication/>.
62. Park Y, Jackson GP, Foreman MA, et al. Evaluating artificial intelligence in medicine: phases of clinical research. *JAMIA Open.* 2020 Sep 8;3(3):326–31.
<https://doi.org/10.1093/jamiaopen/ooaa033>
PMid:33215066