AcademyHealth

September 13, 2024

Brian Anderson, MD
Coalition for Health Artificial Intelligence (CHAI)

RE: Coalition for Health Artificial Intelligence (CHAI) Assurances Standards Guide and Reporting Checklist Request for Information

Dear Coalition for Health Artificial Intelligence Advisory Board:

As the professional home for health services and systems researchers, AcademyHealth is pleased to offer input to guide the development of the Coalition for Health Artificial Intelligence (CHAI) Assurance Standards Guide and Reporting Checklist.

AcademyHealth is the professional home for health services researchers and health policy and the individuals and organizations that use this evidence to improve health and health care. Through a strategic plan that explicitly prioritizes our core values—such as evidence for action, diversity and inclusion, continuous learning, collaboration and community, and trust and integrity—we demonstrate our commitment to improving health outcomes for all, particularly the most vulnerable and disenfranchised populations. Our members are leaders in studying health equity, health disparities, and the social determinants of health. We draw on this expertise to inform our recommendations for CHAI.

As AI's role in health evolves, it has the potential to either drive better equity in health services or widen existing disparities. AI in healthcare can be described as the use of algorithmic systems for a variety of tasks, including decision support, diagnosis, treatment planning, medical imaging analysis, patient monitoring, clinical notetaking, precision medicine, and various administrative processes. CHAI has outlined a six-step lifecycle that details the essential stages and processes involved in developing, deploying, and maintaining AI systems.

We recognize CHAI's acknowledgement that "Fairness, Equity, and Bias" should be included among core principles that must be addressed throughout this AI lifecycle. Moreover, CHAI's constituents will, by May 1, 2025, be required by federal law to take reasonable steps to ensure that their AI does "*not discriminate on the basis of race, color, national origin, sex, age, or disability in its health programs or activities through the use of patient care decision support tools [as defined]*"[45 CFR sec. 92.210]. States are also legislating standards to ensure that AI does not result in worse discrimination than we already have as a society. For CHAI's standards to practically, actionably guide users in addressing bias in alignment with new laws and regulations—they must do far more than shown in the current draft.

Reinforcing this, we further suggest that assessment activities can and should inform steps to mitigate bias and promote fairness and equity. Through health services research (HSR) and other research activities, the field can begin measuring what works, for whom, and in what contexts AI design and deployment impacts health. By guiding people to build this evidence (and learn from it, across the lifecycle) CHAI can make this principle actionable.

Bias doesn't live in a vacuum

1666 K Street NW
Suite 1100
Washington, DC 20006
202 292 6700 PHONE
www.academyhealth.org

The Assurance Standards Guide effectively defines the many forms of bias and the nuanced systematic tendencies or patterns in data, algorithms, or decision-making processes that lead to unequal treatment or outcomes for certain groups. It also recognizes that understanding and mitigating bias in AI systems is essential to achieving fair and equitable AI in health. However, it fails to explicitly address how algorithms, shaped by biased metrics and systemic oppression are not only deeply entangled in and reflective of the world's injustices. Unless intentional steps are taken to inform ethical and equity-oriented design, algorithmic behaviors and outputs actively perpetuate and intensify these injustices.

On page 18, the Guide defines "calibration" as "meaning that outcomes are independent of protected characteristics (or class) – such as race, gender, or their proxies." However, mitigating bias goes beyond defining terms or explaining how they impact individuals; it requires a critical examination of the systems, institutions, and policies that enable and perpetuate these disadvantaged outcomes. For example, a kidney health calculator prevented Black patients from receiving necessary transplants, lung function tests with race corrections led to missed diagnoses of severe pulmonary conditions, race-based calculators used in UTI guidelines for young children resulted in undiagnosed infections in Black girls, causing long-term kidney damage in some cases. To be clear, we do not advocate for eliminating race from AI systems; rather, these cases reveal the urgent need for a critical examination of systems, institutions, and policies—an effort that could have driven the necessary changes to prevent disparate outcomes.

As the Coalition to End Racism in Clinical Algorithms (CERCA) and New York City's Chief Medical Officer Dr. Michelle Morse highlights, "Whose care is centered in these algorithms ... and who is considered the gold standard? Those are the critical questions to ask." Addressing the field's implicit biases and influential oppressive structures is essential to ensure that ethical AI tools are designed to acknowledge, address, and mitigate unequal systems and beliefs, rather than perpetuating them. However, throughout CHAI's 187-page document, there is no mention of the words "racism," "sexism," "ableism," or the intersecting and compounding effects of these systems. These are the structures that perpetuate racial, gender, and ability-based disparities. If we cannot name these social forces, we cannot address or dismantle them. Guidelines must consider not only protected characteristics and their "proxies"—the technical aspects related to data and variables that may inadvertently reflect these characteristics—but also the "corresponding oppressions," which are the broader social and structural issues tied to these characteristics.

Measuring and Mitigating Bias Throughout the AI Life Cycle: Intersectional, Ecological and Interdisciplinary

To address the root causes of bias and create equitable AI systems, CHAI must go beyond the AI lifecycle and examine the broader systems of power, exploitation, and privilege at play. If doing so is beyond the scope of CHAI's aims for these resources, the AI lifecycle should explain or outline guidance on how AI developers can actionably account for these broader dynamics – including by looking to recognized experts on these issues and learning from impacted communities. This requires an intersectional, interdisciplinary, and ecological approach.

Using an intersectional framework of recommendations, guidelines and AI systems can highlight how various forms of oppression (racism, sexism, ableism, classism) overlap and compound one another. This ultimately shapes both the AI lifecycle and its resultant outputs, leading to inequitable access to care, misdiagnosis, or inappropriate treatment plans.

If accessibility needs are not considered alongside factors such as insurance coverage, rurality, and structural and language barriers throughout the design, data collection, model training, and deployment of AI tools, systematically and historically marginalized groups—especially those living outside the norm, such as BIPOC patients with disabilities in rural settings—will continue to face overlooked barriers, and being prescribed technology enabled interventions that fails to serve them equitably.

For example, consider Dr. Jill Inderstrodt, an AcademyHealth member and a research fellow with the NIH's Artificial Intelligence/Machine Learning Consortium to Advance Health Equity and Researcher Diversity (AIM-AHEAD). Her work includes developing an AI-based model to predict and manage late-term preeclampsia, a serious condition affecting pregnant individuals. By incorporating diverse data and avoiding reliance solely on early pregnancy biomarkers, her research promotes inclusivity in AI algorithms. This approach not only ensures that predictive models account for a broader range of variables and experiences but also allows for the inclusion of groups that may have been missing from existing datasets, helping to address potential gaps in representation and ensuring more equitable healthcare outcomes.

Taking an interdisciplinary approach can expose complex layers of inequity and help those designing and deploying AI to avoid inadvertently perpetuating or exacerbating disparities. It is crucial to involve experts from diverse fields—sociology, public health, ethics, data science, and law—in the development, evaluation, and refinement of AI systems and their assurances. Interdisciplinary insights will enable more holistic proxies and interventions that address a broad range of strategies for mitigating biases. It is also crucial that these interdisciplinary experts includes those with "context expertise" related to their specific conditions, communities, characteristics, or lived experiences with racism, sexism, ableism, and classism.

Besides the problem of bias, another critical aspect missing from the guide is the growing concern over AI's environmental impact on health. Given the significant and increasing effects of AI systems on carbon emissions and water consumption, it is essential for CHAI guidelines to incorporate measures addressing environmental sustainability—the ability to maintain or enhance systems, processes, and practices over the long term while protecting resources and minimizing harm to the environment and society. To ensure that AI systems in health services are both fair and sustainable, we propose the following three recommendations for integrating environmental considerations into the Guide:

1. **Energy Efficiency Standards:** Environmental bias stemming from factors such as geographic location, socio-economic status, and access to resources can skew data, decisions, and outcomes, perpetuating inequality. To address this, the field must establish standards for the energy efficiency of AI technologies and infrastructure. This includes protecting natural resources, conserving energy, reducing waste, and using renewable resources. Encouraging the adoption of energy-efficient algorithms, hardware, and data center operations will help minimize the carbon footprint of AI systems and ensure environmental sustainability for future generations.

2. **Sustainable Data Practices:** Promote practices that reduce data storage needs and optimize data processing to minimize environmental impact. Advocate for data minimization strategies and the use of sustainable data storage solutions. Additionally, implement guidelines for assessing the environmental impact throughout the entire lifecycle of AI technologies, from production and deployment to disposal. This includes promoting responsible e-waste management and recycling practices.

3. **Transparency and Reporting:** Require transparency in reporting the environmental impact of AI systems. Encourage organizations to disclose their energy consumption and carbon emissions related to AI operations, and to set and publicly report on sustainability goals.

An ecological approach will enable CHAI guidelines and evaluation measures to address the subtle ways in which institutions, policymakers, and other actors introduce and sustain uneven power structures, thereby perpetuating existing inequities. In the context of accessing care, organizations —through their structures, rules, and practices—can perpetuate inequalities. Critical social theories, such as those focusing on the gendering and racializing organizations, offer valuable insights into how biases extend beyond individual attitudes and become embedded in institutional processes. These theories explain how organizations assign meanings to socially constructed categories like race and gender, influencing the identities of people, groups, and practices.

Beyond Bias as a Buzzword

CHAI should develop standards for conducting audits that assess structural biases throughout the AI lifecycle. These audits would help identify how existing structures and practices contribute to inequalities in care access, treatment, and outcomes. Examining structural biases involves analyzing how policies, routines, and decision-making processes, though seemingly objective, may disproportionately disadvantage certain groups, particularly those historically rendered structurally vulnerable. To address structural bias, we recommend five actions for CHAI:

1. **Evaluate bias at the organizational level**: The Assurance Guide used the word ‘subgroup(s)’ 47 times—why not simply use "groups"? This terminology suggests a hierarchy, reminiscent of the derogatory connotations of "minority," when, in fact, people of color are the majority globally. To mitigate bias from within in, we must use liberatory language and avoid ‘othering’ non-dominant groups. Future drafts should remove language that implies a hierarchy. This process often leads to the creation of racialized burdens—additional, often opaque requirements and barriers imposed on marginalized groups due to biased organizational practices. Racialized burdens can include more complex paperwork, longer wait times, and more stringent eligibility criteria, which disproportionately affect people of color and exacerbate existing disparities in access to care. For example, AI tools trained predominately on white patient data may misdiagnose skin cancer in darker-skinned individuals or models that assess mental health conditions may not account for cultural differences in expressing symptoms leading to misdiagnoses for Black or Latino populations. Racialized burdens neatly carry out the “how” in the production of racial inequality while concealing, or providing an alibi for, the “why.” This perspective is crucial for recognizing that organizations are not neutral entities but often operate in ways that sustain existing power hierarchies.

2. **Developing concrete anti-bias regulations**: To move beyond merely identifying the problem, we must establish clear accountability and allocate dedicated resources. Moreover, addressing these concerns of bias is not just a matter of desire but a legal issue. Recommendations should align with current legal and regulatory decisions. In the Guide, page 24 states that it is essential to "identify when and how users and impacted populations can provide feedback related to fairness, bias, and equity in the design/workflow of the AI solution”. This includes determining whether feedback mechanisms are in place for stakeholders and end users to raise issues regarding

potential bias and fairness of operational processes. It is crucial to specify responsible parties, ensure adequate resource allocation, and address how patients are affected to implement this strategy effectively.

3. **Curating a diverse personnel base**: Diversity should encompass both physical and social categories as well as a range of thought and training. To truly understand different positionalities and axes of oppression, it is essential to engage diverse voices. In both AI and medicine, which are predominantly led by white, male, and middle-class individuals, CHAI guidelines should explicitly recommend that organizations actively promote diversity in hiring and team-building practices to reflect a broad spectrum of perspectives and experiences. Increasing diversity in health technology is essential to safeguard against the pitfalls of "anti-classification," where protected characteristics are not explicitly used in decision-making, as mentioned on page 18. We agree that anti-classification is crucial to avoid tokenism—where marginalized individuals are superficially included without meaningful roles or opportunities. Tokenism in health research or policy often involves symbolic inclusion without genuine engagement, which can lead to inadequate or misaligned health interventions and exacerbate health disparities. When biased AI tokenizes individuals, it may overlook their unique health challenges, resulting in support that fails to address mental health disparities, inclusive healthcare access, or the impacts of discrimination. This neglect not only perpetuates existing inequalities but also undermines the effectiveness of health policies and services meant to support systematically vulnerable groups.

4. **Evaluating bias globally:** We must also consider the broader implications of AI, including the impact on the workers involved in its development and deployment. AI systems might inadvertently encourage unnecessary testing or treatment based on algorithmic recommendations, leading to increased healthcare costs and potential harm to patients. This can be detrimental to already overburdened health care systems in developing countries. The World Health Organization (WHO) has warned the introduction of health-care technologies based on AI could be "dangerous" for people in lower-income countries. This issue is relevant to health service research, as the ethical and labor practices in AI development can affect the quality and equity of health technologies. Companies make choices that influence worker conditions, and our guidelines should address these injustices. Moving forward, it is essential to incorporate considerations of labor exploitation and ethical practices into our evaluation of AI systems to ensure they align with principles of fairness and equity in health services.

5. **Acknowledging the environmental consequence:** Another critical aspect missing from the guide is the growing concern over AI's environmental impact on health. Given the significant and increasing effects of AI systems on carbon emissions and water consumption, it is essential for CHAI guidelines to incorporate measures addressing environmental sustainability—the ability to maintain or enhance systems, processes, and practices over the long term while protecting resources and minimizing harm to the environment and society. To address this, we must establish energy efficiency standards for AI technologies and infrastructure, focusing on protecting natural resources, conserving energy, reducing waste, and utilizing renewable resources. Promoting sustainable data practices—such as reducing data storage needs, optimizing data processing, and implementing data minimization strategies. Additionally, guidelines should be developed to assess the environmental impact of AI throughout its lifecycle, from production to disposal, including responsible e-waste management and recycling. Transparency is also crucial; organizations should be required to report their energy

consumption, carbon emissions, and sustainability goals to ensure accountability and drive environmental responsibility in AI development.

Ethical Concerns: Ensuring Safety, Transparency, and Security in Data

Safeguards are crucial for developing ethical and equitable AI guidelines. Guidelines should emphasize the importance of accountability mechanisms in the ways that health data is collected, used, and potentially misused, through audits, public reporting and explainable data handling, privacy and security practices. In October 2023, the White House released Executive Order 14110, which called for the safe, secure, and trustworthy development and use of artificial intelligence and requires federal agencies to report on their use of AI. This regulatory framework highlights the need for transparency, accountability, and fairness in AI systems that should be informing HSR practices. By integrating these approaches, CHAI can work towards reducing data injustice and centering communities that have been historically and systematically made vulnerable. This will help ensure that the benefits of health AI are distributed more equitably across all populations.

CHAI's efforts around safety, transparency, and security must also focus on data privacy and data justice. In HSR, transparency and accountability are essential for building trust among our stakeholders—patients, healthcare providers, researchers, and policymakers. Ensuring transparency in AI development and deployment processes is not just about making AI systems understandable, but also about critically examining the broader socio-political and economic contexts in which these systems operate. Further, it is important to consider the role, recognition, and opinions of consumers, whose personal data fuels the engine of AI. Data capitalism also raises critical ethical issues around consent, data ownership, and the impact of data-driven decisions on both individuals and society. Despite legal frameworks like HIPAA, the commodification of data often involves extensive data collection and profiling, which can lead to significant power imbalances where a few entities hold disproportionate control over vast amounts of sensitive health information.

Data injustice can manifest in various ways, including limited access to data privacy protections, lack of representation in data collection and analysis, and unequal benefits from data-driven technologies. Addressing data injustice involves advocating for more equitable data rights and protections and ensuring that marginalized and economically disadvantaged communities have a fair share in the benefits and control of their own data.

1. **Inclusive Data Standards:** Develop guidelines to ensure data collection processes are equitable and include diverse populations. This involves establishing protocols for inclusive representation in datasets, particularly for marginalized and economically disadvantaged groups. Such standards will help address the underrepresentation of people of color in clinical trials and other medical health data, ultimately contributing to more comprehensive and equitable research outcomes.

2. **Data Privacy Protections:** Advocate for strong privacy protections that use as little identifiable data as possible and avoid collecting or using highly sensitive data or data from vulnerable populations, as defined by law in appropriate jurisdictions, to ensure compliance with principles of data minimization as privacy protective process. Likewise, security must not only be "robust", it needs to uplevel from what we see today in healthcare, given the volume of data being used. Furthermore, this data must be maximally protected from vulnerabilities in data security that can lead to breaches or ransomware attacks, invasive surveillance and profiling, and the use of data beyond its

intended purposes without explicit patient consent impacting individuals' privacy and autonomy.

The implementation of these recommendations reflects a growing recognition of the need for ethical standards in AI and data practices, especially in health services. For CHAI, this means not only developing and evaluating AI systems that are fair and equitable but also considering the broader socio-political and economic contexts that influence these systems. CHAI should take a proactive role in shaping these guidelines to ensure they address potential power imbalances and inequities perpetuated by AI and data capitalism. By understanding and addressing these ethical, social, and political dimensions, CHAI can help guide the development of AI systems and data practices that promote health equity, protect patient rights, and foster public trust.

Thank you for the opportunity to discuss the perspectives and concerns of the health services research community. For further comment, clarification, or inquiry, please email Josh Caplan at Josh.Caplan@AcademyHealth.org.