Transforming Healthcare Together:
Empowering Health Professionals to Address Bias in the Rapidly Evolving AI-Driven Landscape

IAMSE FALL 2023 WEBCAST AUDIO SEMINAR SERIES

Michael P. Cary, Jr., PhD, RN, FAAN
Sophia Bessias, MSA, MPH

Duke AI HEALTH

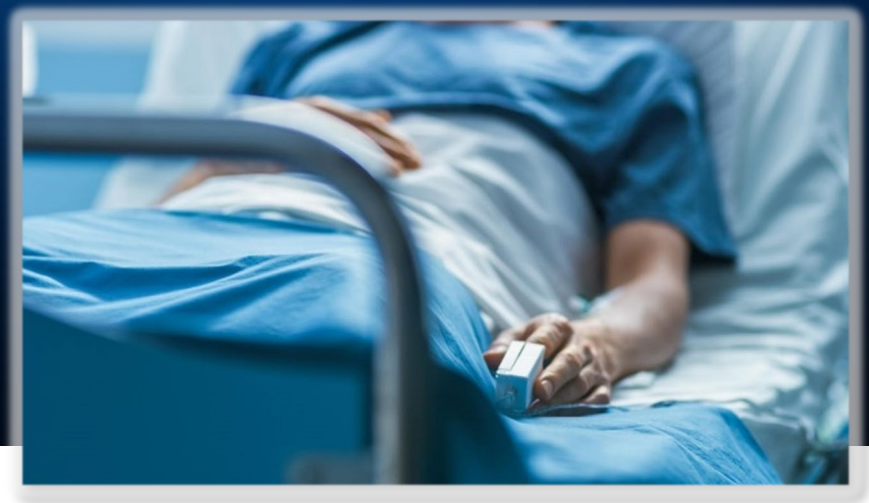
Objectives

1. Establishing Context and Recognizing Challenges
2. Operationalizing Bias Mitigation through AI Governance
3. Navigating the Terrain of Large Language Models (LLMs)
4. Equipping Educators for AI-Driven Healthcare Technologies

Duke AI HEALTH

Promise and Perils of AI in Healthcare

- Improve patient care
- Improve clinician experience *(e.g., reduce burnout)*
- Cost savings
- Operational efficiency



- Non-adoption or over reliance
- No impact on outcomes
- Technical malfunction
- Violation of government regulations
- Non-actionable or **biased recommendations**

https://nam.edu/artificial-intelligence-special-publication/

Duke AI HEALTH

Why is it Important to Identify Racial/Ethnic Bias in Health Algorithms?

Algorithms are used to identify patients with complex health needs in order to provide more comprehensive care management. However, these algorithms can exhibit significant racial bias.

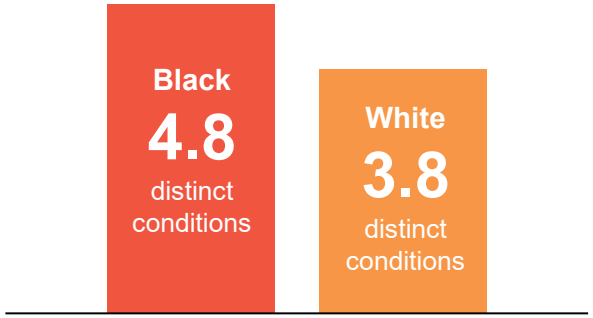
A 2019 study of one such algorithm found:

Black patients who are considerably sicker than White patients are given the same risk score

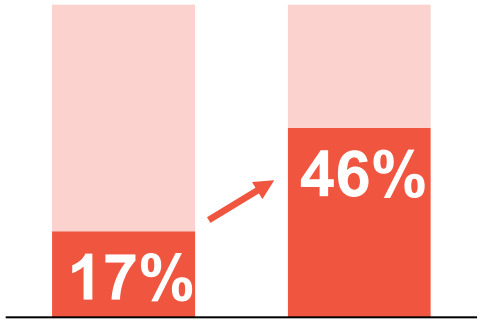
Why is this?

This algorithm assigned risk scores based on past health care spending. Black patients have lower spending than White patients for a given level of health.

At the risk level that would result in automatic identification for the care management program, Black patients had **26.3%** more chronic illnesses than White patients.

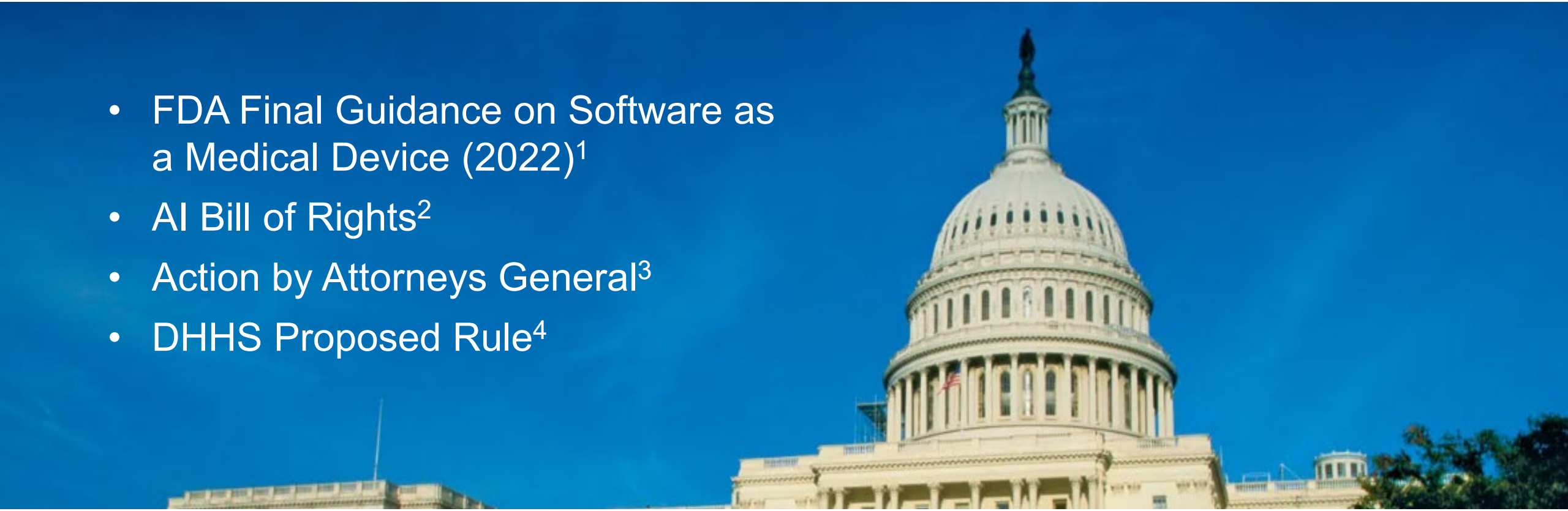


If this bias was eliminated, the percentage of Black patients automatically enrolled in the program would rise from **17.7%** to **46.5%**



Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019 Oct 25;366(6464):447-453

Evolving Landscape for Government Regulation



- FDA Final Guidance on Software as a Medical Device (2022)[1]
- AI Bill of Rights[2]
- Action by Attorneys General[3]
- DHHS Proposed Rule[4]

1. US Food and Drug Administration (FDA) (2022). Final guidance: policy for device software functions and mobile medical applications. https://www.fda.gov/media/80958/download
2. White House Office of Science and Technology Policy (OSTP) (2022). A blueprint for an AI bill of rights. https://www.whitehouse.gov/ostp/ai-bill-of-rights/
3. State of California Department of Justice Office of the Attorney General (2022). Attorney General Bonta launches inquiry into racial and ethnic bias in healthcare algorithms, 2022 Aug 31. https://oag.ca.gov/news/press-releases/attorney-general-bonta-launches-inquiry-racial-and-ethnic-bias-healthcare
4. Department of Health and Human Services (2022). Proposed rule on Nondiscrimination in Health Programs and Activities. 2022 Aug 4. https://www.regulations.gov/document/HHS-OS-2022-0012-0001

Duke AI HEALTH

Proposed Rule on Nondiscrimination in Health Programs and Activities

U.S. Department of Health and Human Services

Enhancing the health and well-being of all Americans

Department of Health and Human Services proposed rule on Nondiscrimination in Health Programs and Activities.

Aug 4, 2022

Use of Clinical Algorithms in Decision-Making (§ 92.210):

‘Covered entities should take steps to ensure that the use of clinical algorithms does not result in discrimination on the basis of race, color, national origin, sex, age, or disability in their health programs and activities.’

Accessed on September 25, 2022, https://www.regulations.gov/document/HHS-OS-2022-0012-0001

Duke AI HEALTH

Scoping Review

We conducted a scoping review of literature to identify strategies that identify and mitigate bias in clinical algorithms.

Literature Search

- August 24, 2022 – searched PubMed, Embase, Web of Science, and ProQuest for publications on mitigating racial bias in clinical algorithms published after 2011.

Study Selection

- Following Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines
  - Included applications, assessment tools, reviews/perspectives, and frameworks that dealt with bias mitigation in EHR-based algorithms
  - Excluded conference abstracts and dissertations
- All aspects of screening were performed by at least two independent reviewers

Data Extraction

- Two reviewers independently extracted data from the full text of all eligible articles.
- Conflicts between reviewers were resolved by a third reviewer.

Publication Notice: "Mitigating Racial And Ethnic Bias And Advancing Health Equity In Clinical Algorithms: A Scoping Review" will be published in the October issue of *Health Affairs* on Monday, October 2 at 4:00PM EST" www.healthaffairs.org

Duke AI HEALTH

Challenges

1. Current research practices have NO “equity lens”
2. Limited education and training
3. Little to no governance in the design or use of data science and AI tools

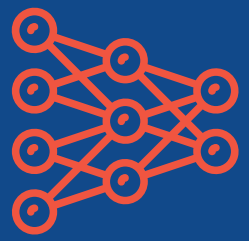
Duke AI HEALTH

Objectives

1. Establishing Context and Recognizing Challenges
2. Operationalizing AI Governance Principles
3. Navigating the Terrain of Large Language Models (LLMs)
4. Equipping Educators for AI-Driven Healthcare Technologies

Duke AI HEALTH

Governance of Algorithms at Health



In light of the risks inherent to deploying algorithms in healthcare, changing regulation, and a system-wide focus on equity, the chancellor, the board of trustees, and the Dean of the School of Medicine charged Duke Health leadership to form a governance framework.

- The Algorithm-Based Clinical Decision Support (ABCDS) Oversight Committee was formed in January 2021.

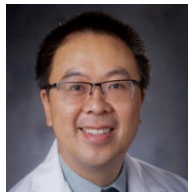
- ABCDS Oversight is a “people-process-technology” that provides governance, evaluation, and monitoring of all algorithms proposed for use in clinical care and/or operations at Duke Health.

Duke AI HEALTH

ABCDS Oversight Committee

Co-Chairs:


M Pencina


E Poon

Director:

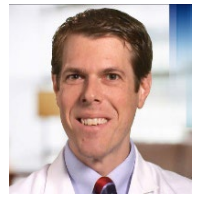

N Economou

ABCDS Oversight Committee

Additional Committee Members:


S Balu


M Cary


M Lipkin


K Lytle

ABCDS Regulatory Subcommittee

Co-Chairs:


A Parrish

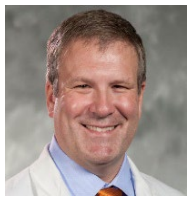

S Elengold


S Ellison

ABCDS Evaluation Subcommittee

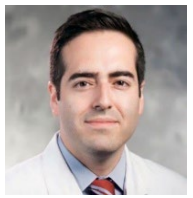
Co-Chairs:


B Goldstein


E Jelovsek

ABCDS Implementation and Monitoring Subcommittee

Co-Chairs:


A Bedoya


C O'Brien

Ops Team:


S Bessias


N Walden

Duke AI HEALTH

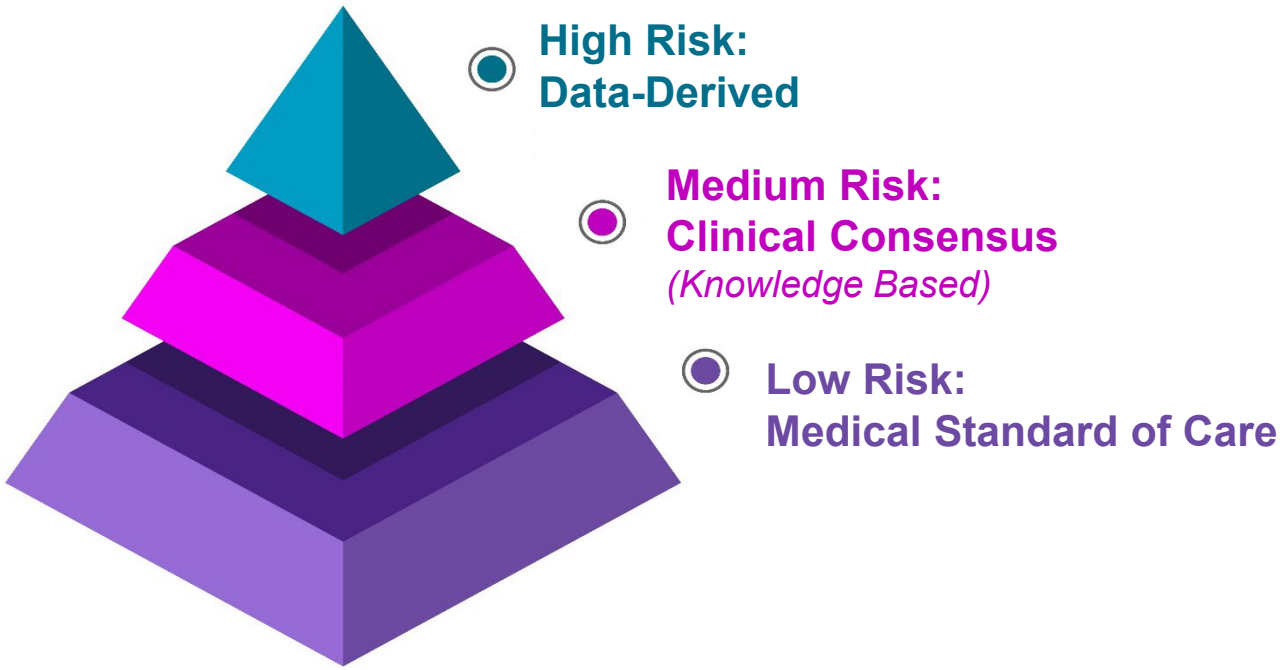
Scope of ABCDS Oversight

Mission

To guide algorithmic tools through their lifecycle by providing governance, evaluation, and monitoring

Core Functions

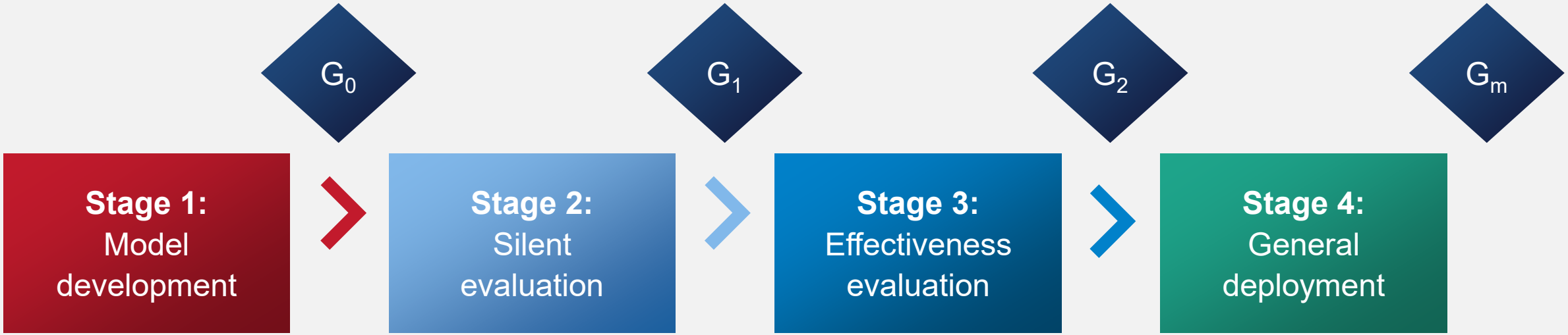
- **Registration:** All electronic algorithms that could impact patient care at Duke Health
- **Evaluation:** High- and medium-risk algorithms



Duke AI HEALTH

ABCDS Evaluation Framework

The algorithm must meet specific requirements to proceed to each new stage of the development lifecycle and is evaluated at “Gates” (Gx) placed between stages.



Bedoya AD, et al., A framework for the oversight and local deployment of safe and high-quality prediction models. *J Am Med Inform Assoc*. 2022 Aug 16;29(9):1631-1636.

Duke AI HEALTH

ABCDS Oversight Gates (G_x)

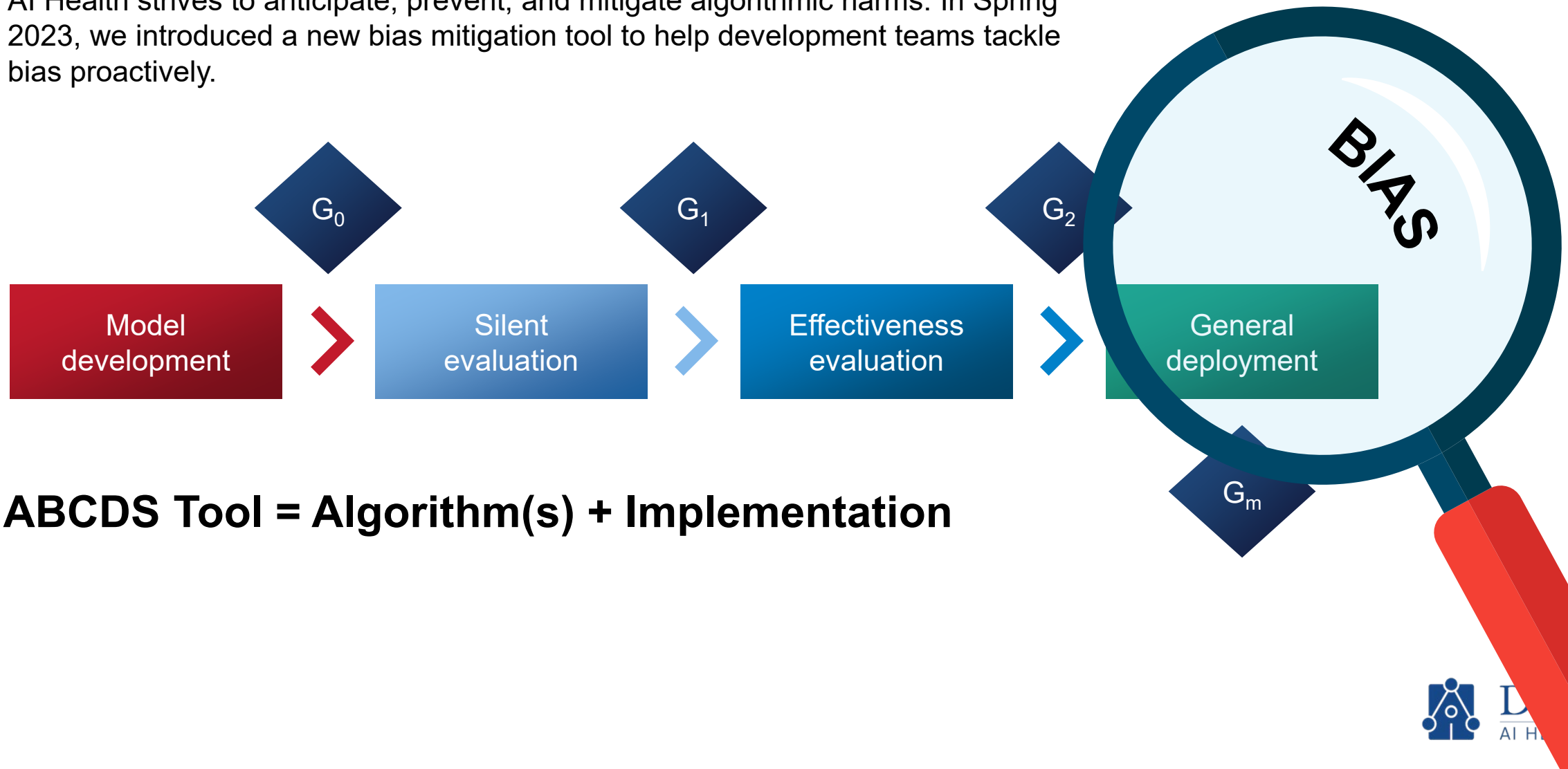
At each G_x gate, Quality & ethical principles are translated into specific evaluation criteria and requirements

1. Transparency & Accountability
2. Clinical Value & Safety
3. **Fairness & Equity**
4. Usability, Reliability & Adoption
5. Regulatory Compliance

Duke AI HEALTH

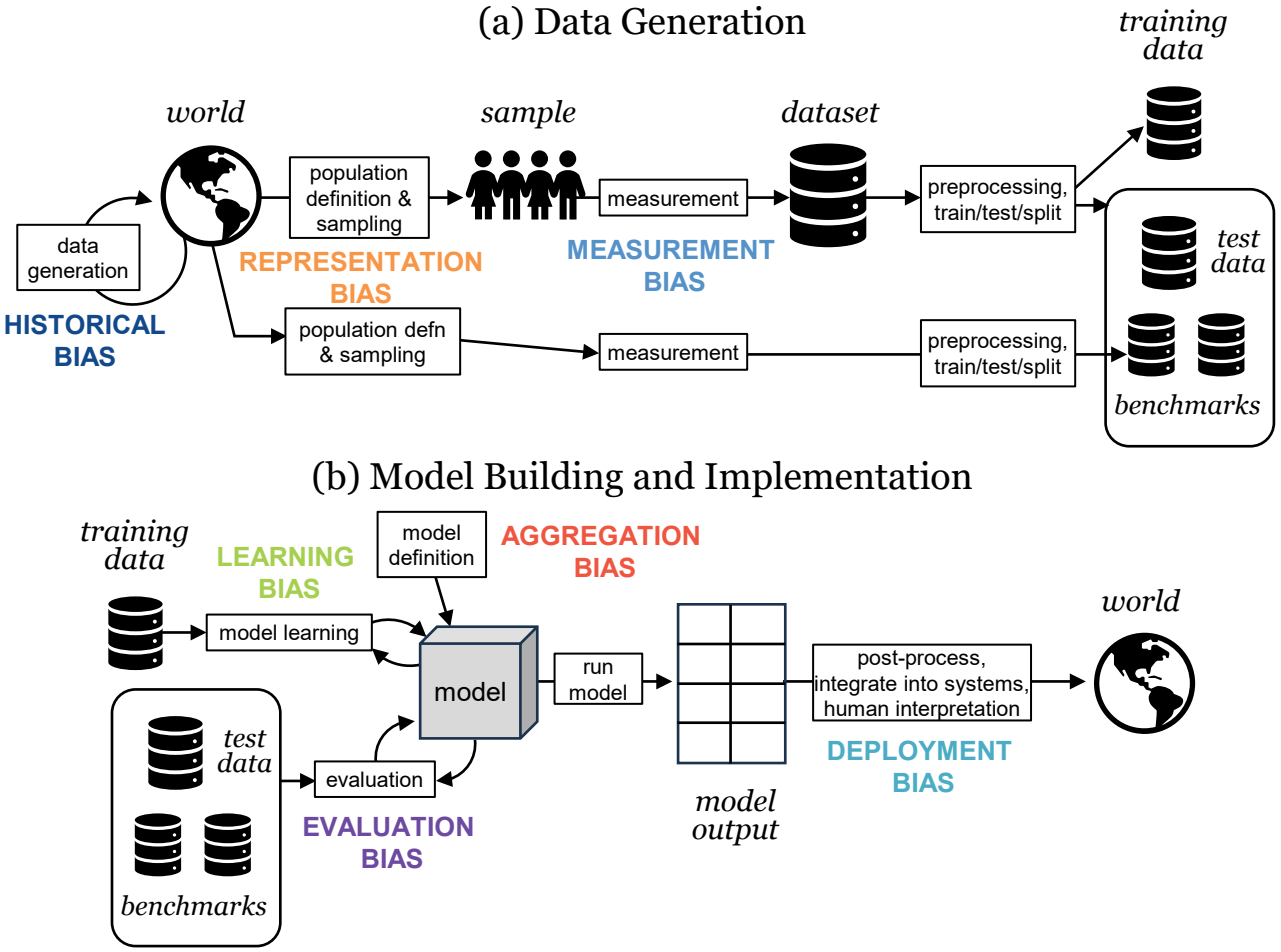
AI Health Focuses on Bias Mitigation

AI Health strives to anticipate, prevent, and mitigate algorithmic harms. In Spring 2023, we introduced a new bias mitigation tool to help development teams tackle bias proactively.



ABCDS Tool = Algorithm(s) + Implementation

Understanding Sources of Bias



- Societal Bias
- Label Bias
- Aggregation Bias
- Learning Bias
- Representation Bias
- Evaluation Bias
- Human Use Bias

Suresh, H., & Guttag, J. (2021). A framework for understanding sources of harm throughout the machine learning life cycle. In Equity and access in algorithms, mechanisms, and optimization (pp. 1-9).

Duke AI HEALTH

From Recognition to Mitigation: Societal Bias

Bias Type	Example	Assessment	Mitigation Strategy
Societal Bias Bias due to training data shaped by present and historical inequities and their fundamental causes	Predictive policing algorithms[1] are trained on data that reflects structural racism and criminalization of, e.g., homelessness and poverty. Groups that are more likely to interact with the police are more likely to be identified by policing algorithms as "at risk" for future offense.	*Please discuss the real-world inequities reflected in your training data and how they inform the problem formulation and intended purpose of your model.*	• *Restriction to particular settings or use cases* • *Human-in-the-loop deployment design* • *Multi-stakeholder engagement*

- Label Bias
- Aggregation Bias
- Learning Bias
- Representation Bias
- Evaluation Bias
- Human Use Bias

Angwin, J. Larson, S. Mattu, L. Kirchner. "Machine bias: There's software used across the country to predict future criminals And it's biased against blacks." ProPublica, 23 May2016; www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

Duke AI HEALTH

Label Bias

Bias Type	Example	Assessment	Mitigation Strategy
Label Bias Use of a biased proxy target variable in place of the ideal prediction target.	An algorithm[1] used to identify patients for high-risk care management services predict healthcare costs as a proxy for healthcare *need*. Despite having greater health needs, Black patients have lower average healthcare spending (due to structural barriers in access to care) and are thus less likely to be recognized by the algorithm as 'high risk.'	*Please discuss any proxies used as inputs or outputs. Provide a rationale and describe implications for use.*	• *Eliminating proxies (where possible) or choosing a proxy as close as possible to the intended idea or concept*

Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019 Oct 25;366(6464):447-453. doi: 10.1126/science.aax2342.

Human Use Bias

Bias Type	Example	Assessment	Mitigation Strategy
Human Use Bias Inconsistent user response to algorithm outputs for different subgroups.	A machine learning algorithm developed to help pathologists differentiate liver cancer types did not improve every pathologist's accuracy despite the model's high rate of correct classification. Instead, pathologists' accuracy was improved when the model's prediction was correct but decreased when the model's prediction was incorrect.	*Briefly describe how your algorithm fits into the clinical workflow. If it will replace an existing model or process, please include a comparison to baseline.*	• *Workflow design solutions* • *End user training* • *Post-deployment monitoring with chart review (required)* • *Collection of end user feedback and metrics of adoption*

Wang HE, et al. A bias evaluation checklist for predictive models and its pilot application for 30-day hospital readmission models. *J Am Med Inform Assoc*. 2022 Jul 12;29(8):1323-1333. doi: 10.1093/jamia/ocac065.

Duke AI HEALTH

Objectives

1. Establishing Context and Recognizing Challenges
2. Operationalizing AI Governance Principles
3. Navigating the Terrain of Large Language Models (LLMs)
4. Equipping Educators for AI-Driven Healthcare Technologies

Duke AI HEALTH

Proposed Applications of Generative AI

- Generating billing information
- Drafting administrative communications
- Automating clinical notes (summarization)
- EHR inbox responses (question and answer)
- Providing direct medical advice, mental health support, etc.

Many use cases have been proposed in popular media, ranging from relatively low-risk to very high risk.

Duke AI HEALTH

Limitations and Ethical Considerations

Accuracy	Large language models are trained to generate **plausible** results, not necessarily **factual** or **accurate** results.
Explainability and Transparency	**Explainability** refers to *how* an algorithm produces its output, whereas **transparency** refers to communication about, e.g. • Sources and types of data used for initial training and fine-tuning • Onward uses of prompt data • Labeling of algorithmically-generated outputs • Environmental impact
Recourse	What options are available to address algorithmic harms? Who has recourse?
Equitable access vs. equitable impact	How are the risks and benefits of generative AI distributed in the population?

Duke AI HEALTH

Example: Automated Clinical Notes

- **What is the problem being solved? Is generative AI a good solution for the task?**
  - What are the root causes of burnout? Are they addressed?
  - How might automated transcription of audio data affect patient trust?
- **What are the implications of the way the algorithm was trained?**
  - On what healthcare datasets has the model been trained or finetuned?
  - For whom is this likely to work more or less well?
- **Who is accountable for the note's accuracy and appropriateness?**
  - Is there a human reviewer? Is there adequate time for review?
- **Privacy and compliance**
  - If the model is exposed to patient data, how is that data being protected?
  - How might patient data be used to further train the model?
  - Who controls how future refinements of the model might be applied?

Duke AI HEALTH

Meeting the Challenge

"Ensure that AI technology serves humans, rather than taking over their responsibilities or replacing them. No matter how good an AI is, at some level, humans must be in charge."

– Michael J. Pencina, PhD

Chief Data Scientist, Duke Health

Vice Dean for Data Science, Duke University School of Medicine

Reference: https://www.newsobserver.com/opinion/article277237823.html#storylink=cpy

Duke AI HEALTH

Objectives

1. Establishing Context and Recognizing Challenges
2. Operationalizing AI Governance Principles
3. Navigating the Terrain of Large Language Models (LLMs)
4. Equipping Educators for AI-Driven Healthcare Technologies

Duke AI HEALTH

Competencies for the Use of AI-Based Tools by Health Care Professionals

AI-related Clinical Competencies for Health Care Professionals

Basic Knowledge of AI	Social and Ethical Implications of AI	Workflow Analysis for AI-based Tools	AI-enhanced Clinical Encounters	Evidence-based Evaluation of AI-based Tools

Practice-Based Learning and Improvement Regarding AI-based Tools

Russell, Regina G. et al. Competencies for the Use of Artificial Intelligence–Based Tools by Health Care Professionals. Academic Medicine 98(3):p 348-356

Duke AI HEALTH

Proposed Competencies for Identifying & Mitigating Algorithmic Bias

Competencies	Objectives
 **Trainees will explain what AI/ML algorithms are in the context of healthcare.**	• Describe how AI/ML technologies contribute to improved patient and health system outcomes such as diagnostics, service referrals, and staff scheduling. • Identify and differentiate between various applications of AI/ML technologies in health and healthcare, such as predictive clinical models, race-based calculators, and software such as medical devices.
 **Trainees will explain how AI governance and legal frameworks can impact equity.**	• Discuss the potential discriminatory effects and ethical implications of AI/ML technologies on care delivery and patient outcomes. • Summarize the legal frameworks, governance, and regulations pertaining to the use of AI/ML technologies including the National Institute of Standards and Technology Artificial Intelligence Risk Management Framework.
 **Trainees will learn ways of detecting and mitigating biases in AI/ML algorithms, across the algorithmic life cycle.**	• Detect and analyze potential biases in AI/ML algorithms and their training datasets, considering demographic, socioeconomic, and geographical factors. • Implement strategies to mitigate bias, promoting fairness in AI/ML algorithm. • Apply methods to assess the accuracy of AI/ML-generated outputs in diagnostics, treatment recommendations, and patient monitoring.

Duke-NCCU RC2 collaborators: Michael P. Cary, Jr., PhD, RN; Siobahn Grady, PhD; Ben Goldstein, PhD, MPH; Jacquelyn McMillian-Bohler, PhD, RN; Andrew Olson, MPP; Matthew Engelhard, MD, PhD; Jessica Smokoski, PhD; Nicoleta Economou-Zavlanos, PhD; Sophia Bessias, MPH, MSA; Irene Dougherty, PhD; and Charlene Harrington, PhD

Duke AI HEALTH

Training Opportunities & Resources

Short Courses and Workshops

- December 2023: Duke AI Health Electronic Health Records Study Design Workshop
- March 2024: Duke AI Health and School of Nursing to convene first-ever Duke Symposium on Algorithmic Equity and Fairness in Health

Formal Programs

- Duke Master of Management in Clinical Informatics
- Duke Master in Interdisciplinary Data Science
  - › Clinical mentor and capstone
- Duke AI Master in Engineering
  - › Product design and hands-on practice experience

Other Resources

- Duke AI Health offers a virtual seminar series in fall and spring semesters open to anyone in the world
- NIH
  - AI/ML Consortium to Advance Health Equity and Researcher Diversity (AIM-AHEAD)
  - Science Collaborative for Health disparities and AI bias Reduction (ScHARe)

Duke AI HEALTH

Research Symposium

SAVE THE DATE!

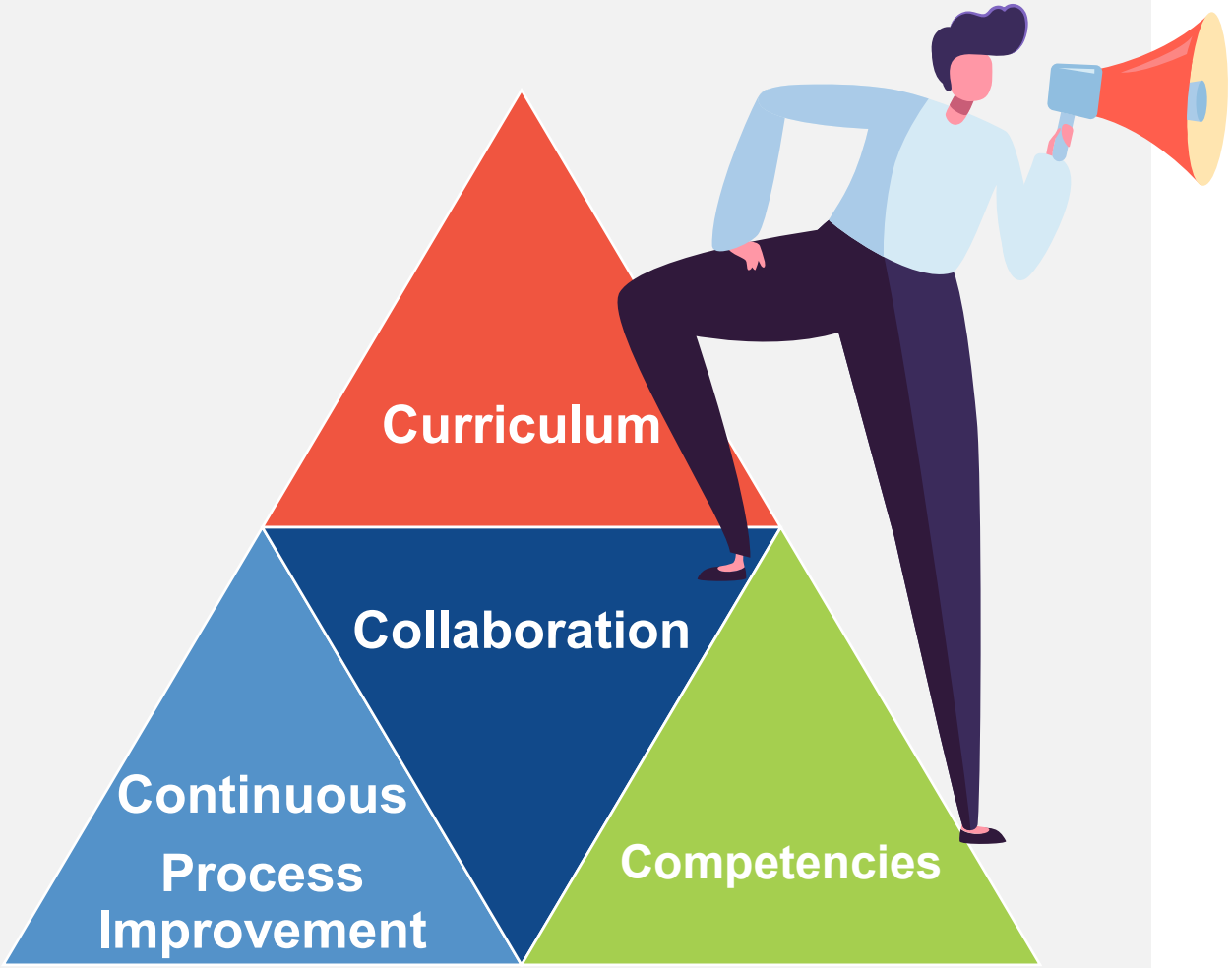
The Symposium is scheduled to take place in person at the Duke School of Nursing from **March 13-14, 2024**

MODERATOR

Michael Cary, PhD, RN
Elizabeth C. Clipp Term Chair of Nursing
Equity Scholar, AI Health

See you at the 2024 Duke Research Symposium!

Duke AI HEALTH



Call to Action for Transforming Healthcare Together

We invite you to join Duke AI Health in our commitment to advancing health equity and promoting responsible AI:

- Transformative Curriculum Design
- Build Competency-based Training Programs
- Strengthen Interdisciplinary Collaborations
- Continuous Process Improvement

Duke AI HEALTH