

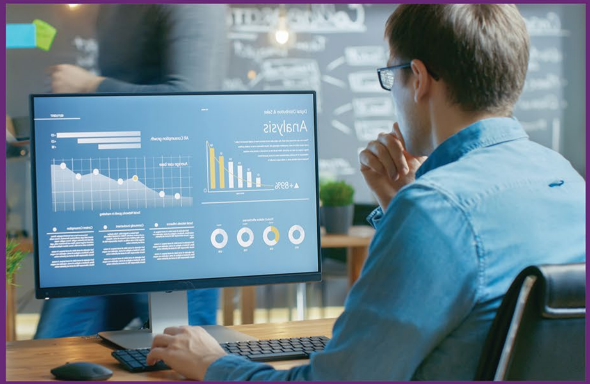
AGENCY FOR HEALTHCARE RESEARCH AND QUALITY

AHRQ

Impact of Healthcare Algorithms on Racial and Ethnic Disparities in Health and Healthcare

March 3, 2023
10:00 a.m.- 2:00 p.m. ET

NIH National Institute on Minority Health and Health Disparities



AGENCY FOR HEALTHCARE RESEARCH AND QUALITY

AHRQ

Welcome

Prashila Dullabh, MD
NORC

Racial Bias and Healthcare Algorithms
March 3, 2023
10:00–10:05 a.m. ET

NIH National Institute on Minority Health and Health Disparities

Zoom Housekeeping

Please use the Q&A feature to send comments or questions to the panelists.

Today's webinar is being recorded for notetaking purposes.

Live captions are available. Click the “more” icon on your screen and click “show subtitle”.

Update your Zoom name with your first & last name and organization (e.g., Jane Smith – NORC).

If you experience issues with Zoom, please contact Andrew Chiao OR Tyler Taylor from NORC, via the chat or email for assistance: chiao-andrew@norc.org, taylor-tyler@norc.org

National Institute on Minority Health and Health Disparities

Disclaimer

- Presentations do not necessarily represent the views of AHRQ or the U.S. Department of Health and Human Services (DHHS); therefore, please do not interpret any statement in this presentation as an official position of AHRQ or of DHHS.

- Additionally, presentations and presenters were selected to include diverse perspectives and do not necessarily represent the views of the consensus panel.

Consensus Panel Co-Chairs

AHRQ Agency for Healthca Research and Qualit

Lucila Ohno-Machado, MD, PhD, MBA
Yale School of Medicine

Marshall H. Chin, MD, MPH
University of Chicago

NIH National Institute on Minority Health and Health Disparities

Consensus Panel

Nasim Afsar, MD, MBA
Oracle Health

Malika Fair, MD, MPH
Association of American Medical Colleges

Shaheen Gauher, PhD
Elevance Health

Tina Hernandez-Boussard, PhD, MS, MPH
Stanford University

National Institute on Minority Health and Health Disparities

Consensus Panel

Maia Hightower, MD, MPH, MBA
University of Chicago Medicine

Bill Jordan, MD, MPH
American Medical Association

Tamra Moore, JD
King & Spalding

National Institute on Minority Health and Health Disparities

AGENCY FOR HEALTHCARE RESEARCH AND QUALITY

AHRQ

Welcome

Deborah Guadalupe Duran, PhD
NIMHD

Racial Bias and Healthcare Algorithms
March 3, 2023
10:00–10:05 a.m. ET

NIH National Institute on Minority Health and Health Disparities

Welcome Remarks

Steven Posnack, MS, MHS serves as the Deputy National Coordinator for Health Information Technology. In this role, he advises the national coordinator, leads the execution of ONC's mission, and represents ONC's interests at a national and international level. In conjunction with the national coordinator, Steve oversees ONC's federal coordination, regulatory policy, public-private initiatives, and the overall implementation of statutory authorities and requirements, such as those from the 21st Century Cures Act and HITECH Act.

Dr. LaShawn McIver, MD, MPH joined CMS as the Director of the Office of Minority Health in August 2020. She is a proven public health leader with experience in driving successful health initiatives and public policy efforts aimed at promoting health equity, improving health outcomes, increasing access to care, and promoting health system reform.

National Institute on Minority Health and Health Disparities

AGENCY FOR HEALTHCARE RESEARCH AND QUALITY

ONC Remarks

Stephen Posnack, MS, MHS
Deputy National Coordinator for Health IT
Office of the National Coordinator for Health Information Technology (ONC)

Racial Bias and Healthcare Algorithms
March 3, 2023
10:05–10:10 a.m. ET

National Institute
on Minority Health
and Health Disparities



AGENCY FOR HEALTHCARE RESEARCH AND QUALITY

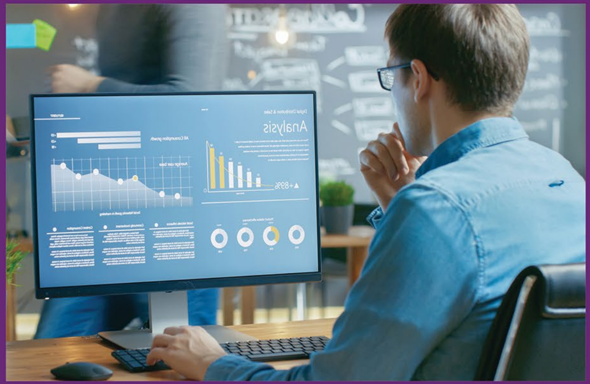
AHRQ

CMS Remarks

LaShawn McIver, MD, MPH
Director, Office of Minority Health
Centers for Medicare & Medicaid Services (CMS)

Racial Bias and Healthcare Algorithms
March 3, 2023
10:10–10:15 a.m. ET

NIH National Institute on Minority Health and Health Disparities


AGENCY FOR HEALTHCARE RESEARCH AND QUALITY

AHRQ

Evidence Review
Key Question 2

Brian Leas, MS, MA

University of Pennsylvania School of Medicine

Racial Bias and Healthcare Algorithms

March 3, 2023

10:15-10:35 a.m. ET

NIH National Institute on Minority Health and Health Disparities

Key Question 2

What is the effect of interventions, models of interventions, or other approaches **to mitigate racial and ethnic bias** in the development, validation, dissemination, and implementation of healthcare algorithms?

KQ 2: Clinical Topics

warfarin dosing

opioid misuse

postpartum depression

lung function

lung cancer screening

healthcare costs and utilization

kidney function

stroke risk

cardiovascular risk

organ donation

intensive care needs

KQ 2: Studies

Study Design

- 1 RCT
- 17 cohort or pre-post
- 15 simulations

Risk of Bias

- 5 Low ROB
- 23 Moderate ROB
- 5 High ROB

National Institute
on Minority Health
and Health Disparities

KQ 2: Mitigation Strategies

- Removing Input Variables
- Replacing Input or Outcome Variables
- Adding Input Variables
- Changing the Patient Mix Used for Development and Validation
- Developing Separate Algorithms by Race
- Refining Statistical and Analytical Techniques

KQ 2: Results

Mitigation Strategy	N, Studies	Algorithm
Removed race	15	eGFR for kidney function GLI spirometry equation for lung function Novel risk prediction algorithm for postpartum depression
Replaced race with biological indicators	4	eGFR for kidney function Kidney Donor Risk Index for kidney transplant suitability
Replaced biased healthcare outcome variable with unbiased variables	1	Novel risk prediction algorithm for complex healthcare needs
Added race	3	FRS for cardiovascular disease risk CHA_2DS_2-VASc for stroke risk
Added biological input variables	3	ASCVD for cardiovascular disease risk Novel risk prediction algorithm for cardiovascular disease COAG for warfarin dosing
Added health outcome variables	1	USPSTF-2020 for lung cancer risk
Added measures of SDOH	1	Novel risk prediction algorithm for complex healthcare needs
Recalibrated after improving population representation	4	ASCVD for cardiovascular disease risk Novel risk prediction algorithm for postpartum depression Donor Risk Index for liver transplant suitability
Stratified algorithms for Black and White patients	2	COAG for warfarin dosing Novel risk prediction algorithm for opioid misuse
Statistical techniques	2	Novel risk prediction algorithm for postpartum depression ASCVD for cardiovascular disease risk

KQ 2: Findings

- Aside from eGFR, substantial heterogeneity: patient populations, clinical conditions, healthcare settings, primary outcomes

- Traditional algorithms, traditional solutions

- Mitigation strategies **improve** algorithmic accuracy, but **inference** and **simulation** used to estimate effect on disparities

- Modeling may not fully reflect potential biases in algorithm translation, dissemination, and implementation

National Institute on Minority Health and Health Disparities

KQ 2: Conclusions

- We don’t know what we don’t know: might be unpublished studies that found no effect of mitigation interventions
- Further research is needed to quantify the real-world effects of modifying algorithms
- Mitigation effectiveness is largely context-specific and may depend on algorithm, clinical condition, population, setting, outcomes



AGENCY FOR HEALTHCARE RESEARCH AND QUALITY

AHRQ

Mitigating Bias in Algorithms

Christina Silcox, PhD
Duke-Margolis Center for Health Policy

Racial Bias and Healthcare Algorithms
March 3, 2023
10:35 – 10:43 a.m. ET

NIH National Institute on Minority Health and Health Disparities

Bias in Different Types of Algorithms

nature

Explore content ⌄ About the journal ⌄ Publish with us ⌄ Subscribe

nature > news > article

NEWS | 24 October 2019 | Update 26 October 2019

Millions of black people affected by racial bias in health-care algorithms

Study reveals rampant racism in decision-making software used by US hospitals — and highlights ways to correct it.

Heidi Ledford

nature

Explore content ⌄ About the journal ⌄ Publish with us ⌄ Subscribe

nature > news > article

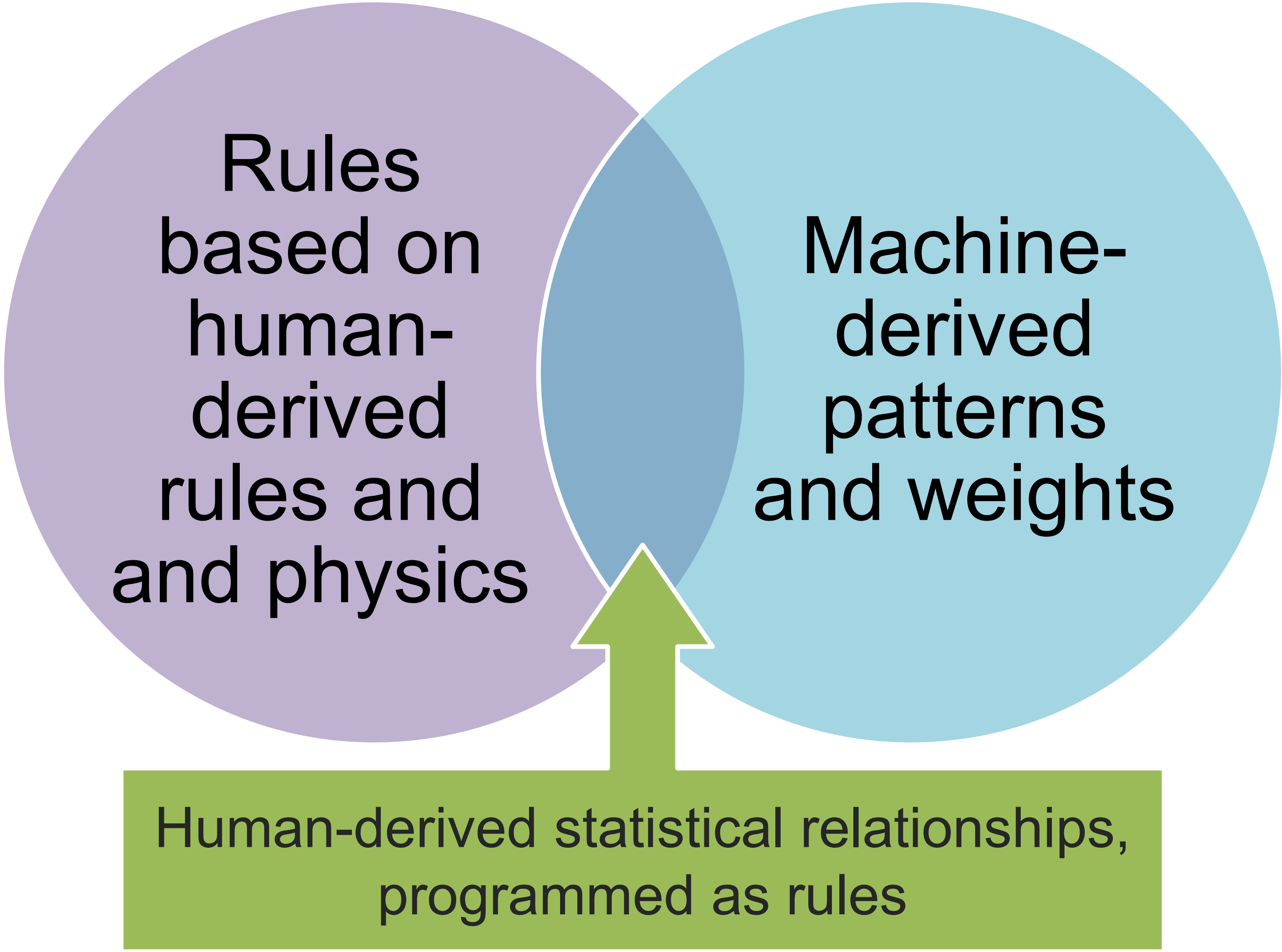
NEWS | 16 December 2020

Is a racially-biased algorithm delaying health care for one million Black people?

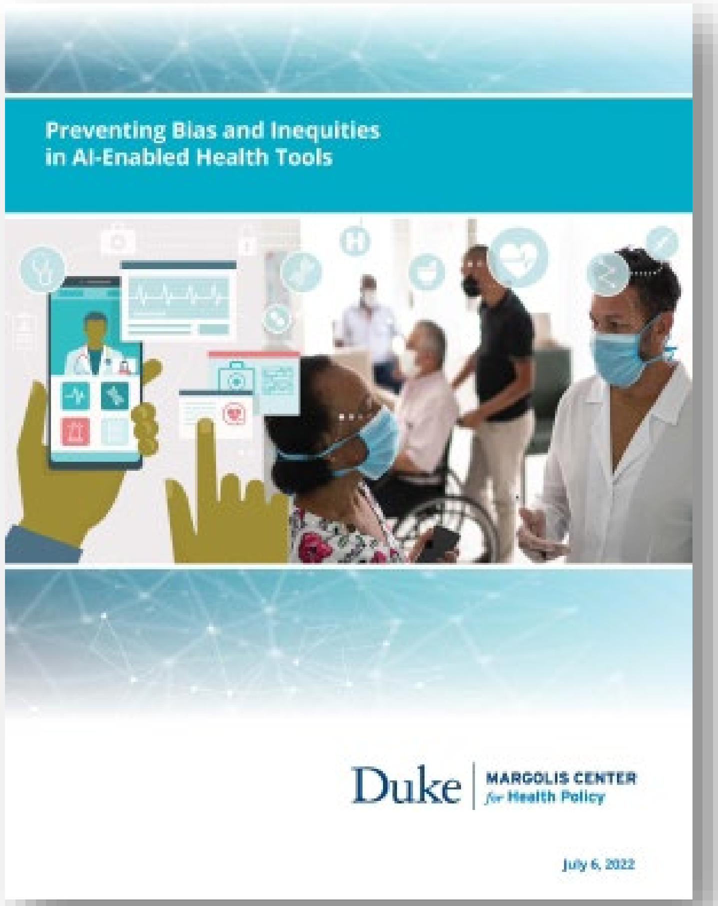
Sweeping calculation suggests it could be — but how to fix the problem is unclear.

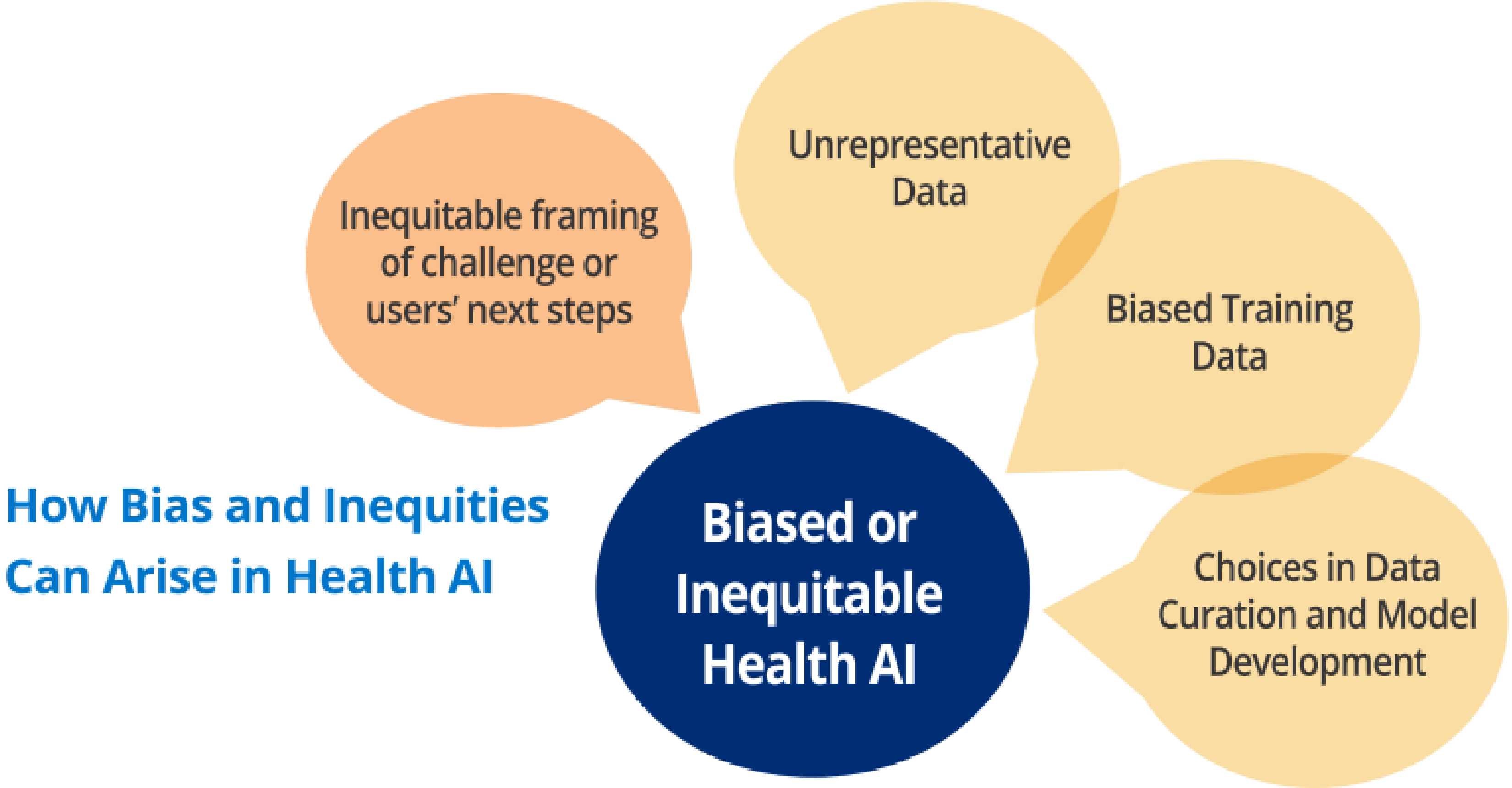
Jyoti Madhusoodanan

Overlap in Rules-Based and Machine Learning Algorithms



Sources of Bias and Inequity





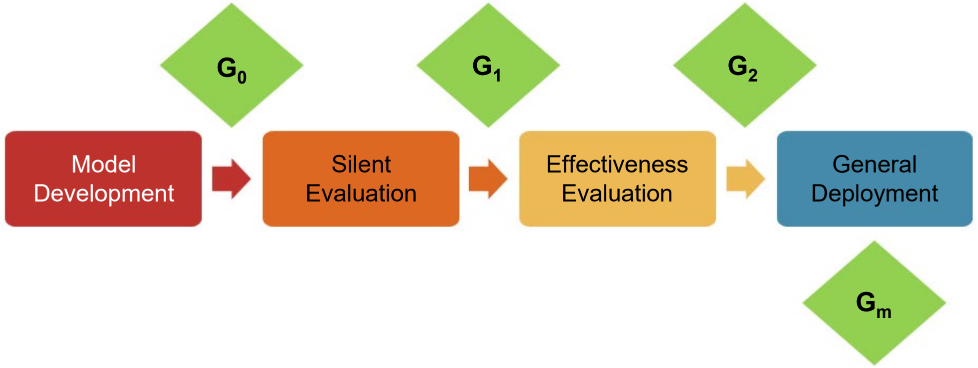
National Institute on Minority Health and Health Disparities

Preventing Bias and Inequities in AI-enabled Health Tools
https://healthpolicy.duke.edu/publications/preventing-bias-and-inequities-ai-enabled-health-tools

Implementing Fairness & Equity Principles

Duke Health's algorithmic oversight process translates ethical & quality principles into concrete evaluation criteria and submission requirements appropriate for each checkpoint throughout the lifecycle

Quality & Ethical Principles	Evaluation Criteria	Submission Material
Policies, Regulations etc.	Committee Approval	Development Teams
Fairness & Equity	Evidence & Future Plans	✓ Analytical problem framing and sampling (societal, representation, measurement bias) ✓ Justification re: sensitive variables as inputs, subgroup analysis on performance and impact metrics (algorithmic bias) ✓ Intended and unintended uses, workflow integration (human user bias)



Educating Our Community

National Institute on Minority Health and Health Disparities

Bedoya, A. D., et al. (2022): "A framework for the oversight and local deployment of safe and high-quality prediction models." Journal of the American Medical Informatics Association.
https://aihealth.duke.edu/algorithm-based-clinical-decision-support-abcds/



AGENCY FOR HEALTHCARE RESEARCH AND QUALITY

End-to-End Bias Evaluation Checklist for Predictive Models

Suchi Saria, PhD
John C. Malone Associate Professor
Computer Science, Stats, and Health Policy, Johns Hopkins University
Founder, Bayesian Health

Racial Bias and Healthcare Algorithms
March 3, 2023
10:43 – 10:51 a.m. ET

National Institute on Minority Health and Health Disparities

Algorithms are Sensitive to Healthcare Disparities

RESEARCH

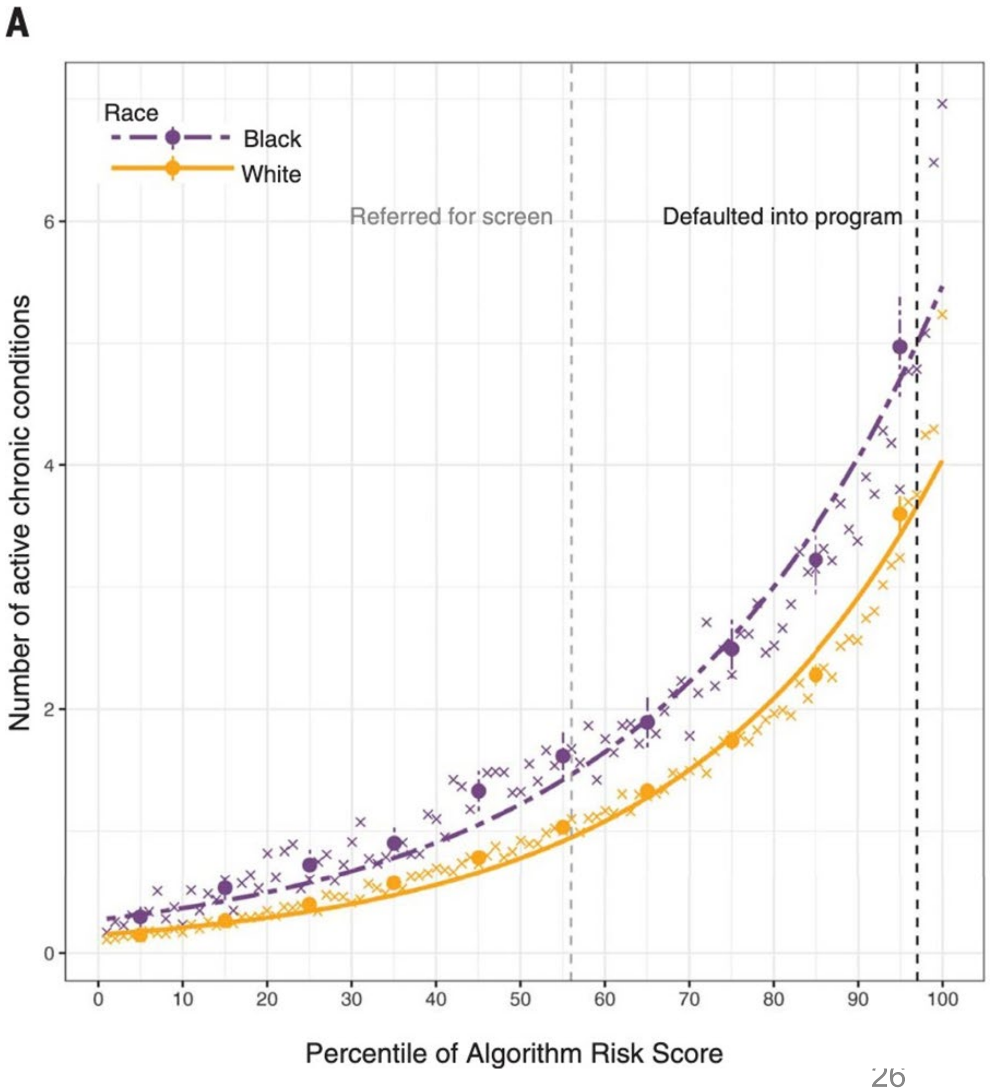
RESEARCH ARTICLE

ECONOMICS

Dissecting racial bias in an algorithm used to manage the health of populations

Ziad Obermeyer[1,2]*, Brian Powers[3], Christine Vogeli[4], Sendhil Mullainathan[5]*†

Health systems rely on commercial prediction algorithms to identify and help patients with complex health needs. We show that a widely used algorithm, typical of this industry-wide approach and affecting millions of patients, exhibits significant racial bias: At a given risk score, Black patients are considerably sicker than White patients, as evidenced by signs of uncontrolled illnesses. Remedying this disparity would increase the percentage of Black patients receiving additional help from 17.7 to 46.5%. The bias arises because the algorithm predicts health care costs rather than illness, but unequal access to care means that we spend less money caring for Black patients than for White patients. Thus, despite health care cost appearing to be an effective proxy for health by some measures of predictive accuracy, large racial biases arise. We suggest that the choice of convenient, seemingly effective proxies for ground truth can be an important source of algorithmic bias in many contexts.



If not developed and implemented carefully, algorithms can propagate and create healthcare disparities

National Institute on Minority Health and Health Disparities

Disparate Performance

- Many types of biases affect algorithm performance across subgroups.
  - ► Exist at all stages of model development and deployment
- This leads to disparities created by the algorithm’s use in the real world.

Sources of Bias

	Source of Bias	How the Bias Can Arise
Model definition and design	Label bias	Use of a biased proxy target lavariable in place of the ideal prediction target.
	Modeling bias	Use of a model, that due to its technical design, leads to inequitable outcomes.
Data Collection	Population bias	Poor performance in subsets of the deployment population due to non-representative training data.
	Measurement bias	Bias due to differences in how features are measured across subgroups.
Validation	Mission validation bias	Absence of validation studies that explicitly measure performance across subgroups
Deployment	Human use bias	Inconsistent user response to algorithm outputs for different subgroups.

Wang, H.E., Landers, M., Adams. R, Subbaswamy, A., Kharrazi, H., Gaskin, D.J. and Saria, S., 2022. A bias evaluation checklist for predictive models and its pilot application for 30-day hospital readmission models. *JAMIA*

National Institute on Minority Health and Health Disparities

Checklist for Bias Evaluation

- Case study: evaluating potential for bias in 30-day readmission risk models

Stage	Source of bias	LACE	HOSPITAL	ACG	HATRIX
1. Model definition and design	Label bias	RED	RED	RED	RED
	Modeling bias - general	RED	GREEN	RED	RED
	Modeling bias – key feature missing	RED	RED	GREEN	GREEN
	Modeling bias –accounting for bias	RED	RED	RED	RED
2. Data collection and acquisition	Population bias	GREEN	GREEN	YELLOW	GREEN
	Measurement bias - inputs	GREEN	GREEN	YELLOW	GREEN
	Measurement bias – prediction target	RED	RED	GREEN	GREEN
	Measurement bias - incompleteness	RED	RED	RED	RED
3. Validation	Missing validation bias	RED	RED	RED	RED
4. Deployment and model use	Human use bias – different interpretation	RED	RED	YELLOW	RED
	Human use bias – model use	YELLOW	YELLOW	YELLOW	YELLOW
	Human use bias – reduce uncertain	GREEN	GREEN	GREEN	GREEN

Figure 2. Model assessment heat map. An overall rating was given for each bias type based on the qualitative assessment of the checklist questions (details in Appendix 1.) RED indicates there is potential for concern. GREEN indicates there is limited potential for concern. YELLOW indicates the potential for concern is unclear or there is not enough information with which to draw a conclusion.

Wang, H.E., Landers, M., Adams, R., Subbaswamy, A., Kharrazi, H., Gaskin, D.J. and **Saria, S.**, 2022. A bias evaluation checklist for predictive models and its pilot application for 30-day hospital readmission models. *JAMIA*.



AGENCY FOR HEALTHCARE RESEARCH AND QUALITY

AHRQ

AI Code of Conduct (AICC) Project

Laura Adams

Senior Advisor, National Academy of Medicine

Racial Bias and Healthcare Algorithms

March 3, 2023

10:51-10:59 a.m. ET

NIH National Institute on Minority Health and Health Disparities

NAM Leadership Consortium's AI Initiatives

- ***Newest AI initiative***: Align the field around a "current best practice" healthcare AI Code of Conduct to be:
  - implemented
  - tested
  - validated
  - continually improved

- Identify each stakeholder's role in applying the Code at each stage of AI lifecycle

- Focus on equity, inclusion, and implementation vigilance

National Institute on Minority Health and Health Disparities

The Context

- Algorithms in healthcare are proliferating rapidly
- Evidence indicates the need for guidance, policy, and learning
- Aligning around a current best practice provides clarity, supports innovation, and promotes learning
- The NAM is in a unique position to convene a broad array of stakeholders to advance this work

Transparency and inclusion are key project design principles

- ► People support what they help create
- ► All of us are better than any one of us

NIH National Institute on Minority Health and Health Disparities

The NAM AI Code of Conduct (AICC) Project

- Build upon the existing body of work on AI principles/frameworks

- Then:
  - ► Identify areas of convergence and gaps in current frameworks
  - ► Seek expert stakeholder and public input
  - ► Publish NAM Commentary paper that includes a draft "best practice" code of conduct for public comment

- Incorporate feedback and dive deeper into such issues as:
  - ► Bias elimination and promotion of equity
  - ► Involvement of those affected by AI in the development of AI
  - ► Algorithmic vigilance post-implementation
  - ► Data linkage and sharing

National Institute on Minority Health and Health Disparities

NATIONAL ACADEMY OF MEDICINE

NAM Leadership Consortium

Collaboration for a Learning Health System

The NAM AICC Project (cont'd)

- Incorporate input into a NAM Special Publication which will include:
  - ► A current best practice healthcare AI Code of Conduct
  - ► A proposed methodology for Code implementation, testing, validation and continuous learning:
    - Collaborations with others are *essential* (e.g., Coalition for Health AI (CHAI) work in establishing algorithmic assurance/validation labs)
  - ► Role of each stakeholder at each AI lifecycle phase vis-à-vis the Code
  - ► Priorities for accelerating progress

Contact: ladams@nas.edu



AGENCY FOR HEALTHCARE RESEARCH AND QUALITY

Resources to Explore AI Bias Mitigation Strategies

Dr. Luca Calzoni, MD MS PhD Cand.
NIMHD ScHARE

Racial Bias and Healthcare Algorithms
March 3, 2023
10:59 – 11:07 a.m. ET

National Institute
on Minority Health
and Health Disparities

AI Bias Can Lead to Health Disparities

AI can be a transformative tool for improving care and population health, **when developed with Health Equity**

Many algorithms are biased by design, or trained on biased data

- *Example: an algorithm to distinguish between malignant and benign moles trained on light-skinned patients*

Biased algorithms can exacerbate existing inequities in socioeconomic status, race, ethnic background, disability, religion, gender or sexual orientation, **and lead to health disparities**

Harm to populations can also come:

- when Big Data is used for AI without consent
- when algorithms are not applied fairly across populations, or the socio-cultural context is missed due to lack of diversity in the AI workforce

National Institute on Minority Health and Health Disparities

Panch et al., 2019; FDA, 2017; Shachar, 2023

If bias is in the world, it will always be present in the data and learned by AI

The **public sector** can:

1. address society biases introduced or uncovered by AI
2. establish fairness standards
3. regulate AI deployed in health systems

but **regulation is not enough**, and **clinicians cannot bear all liability risks/ burdens of identifying biases**

To prevent biases from resulting in health disparities, **we need:**

- **implementation strategies** across the entire AI development cycle
- **critical thinking** and an ethical inquiry approach in AI users

Look Deeper with More Eyes

*"For the first time in history, we have a technology (AI) that is opening our eyes to **who we are**, and could allow us to play a conscious role in **who we want to become**.*

***Jennifer Aue -** IBM Director for AI Transformation*

Who we are

- Human biases are perpetuated or amplified in AI applications

Who we want to become

Develop AI with Equity to Prevent Health Disparities:

- Use models in **context**
- Ensure R's in AI:
  - **R**epeatability
  - **R**eplicability
  - **R**eproducibility
- Increase AI workforce **diversity** to **look deeper with more eyes**

ScHARe

AI Bias Mitigation Collaboratives

Join our effort to implement mitigation strategies in:

- Project design
- Data
- Algorithm development and training
- Implementation

Introducing the ScHARe Platform

ScHARe is a cloud-based population science data platform that offers researchers at all experience levels and disciplines:

- The ability to **collaboratively use AI tools** in a **secure setting**
- Access to SDoH and other population science large **datasets**

ScHARe fills three critical gaps:

1. Collaboratively advance **AI bias mitigation strategies and ethical inquiry** by increasing the use of diverse eyes and skills
2. Promote participation of **women and populations with health disparities** in data science
3. Leverage health disparities and healthcare outcomes **research opportunities** afforded by Big Data, AI and cloud computing

National Institute on Minority Health and Health Disparities

Register for ScHARe:

bit.ly/join-schare

Join our email list:
bit.ly/schare-news

Join the ScHARe Think-a-Thons

Think-a-Thons are **virtual meetings** for people conducting health disparities and healthcare outcomes research

- **Monthly sessions** (2 hours)
- Designed for new and experienced users
- Two types:
  1. **Instructional**
  2. **Research-focused**: teams collaborate around health disparities and healthcare outcomes research projects leading to publications
- Networking, mentoring and coaching opportunities

Join our Think-a-Thons:

bit.ly/think-a-thons

AI bias mitigation:
let's keep the dialogue open

Contact us:
schare@mail.nih.gov

National Institute on Minority Health and Health Disparities

Impact of Healthcare Algorithms on Racial and Ethnic Disparities in Health and Healthcare

Consensus Panel Discussion/Q & A

March 3, 2023

11:07 – 11:45 a.m. ET

Discussion Questions

- What works, what's missing, additional experiences and insights on bias mitigation strategies, including approaches to implementation?

- What guidance is needed to mitigate bias/what are the next steps, for different parts of AI lifecycle?
  - Guidance for algorithm development, testing, and updating?
  - Guidance for algorithm implementation?

Impact of Healthcare Algorithms on Racial and Ethnic Disparities in Health and Healthcare

Break

March 3, 2023

11:45 a.m.– 12:15 p.m. ET

Please take thirty minutes for lunch

AGENCY FOR HEALTHCARE RESEARCH AND QUALITY

AHRQ

Evidence Review

Contextual Question 2

Gary Weissman, MD, MSHP

University of Pennsylvania School of Medicine

Racial Bias and Healthcare Algorithms

March 3, 2023

12:15 – 12:35 p.m. ET

NIH National Institute on Minority Health and Health Disparities

Contextual Question 2

What are existing or emerging national or international **standards or guidance** for how algorithms should be developed, validated, implemented, and updated **to avoid introducing bias** that could lead to health and healthcare disparities?

CQ 2: Methodology

- Primary literature searches
- Supplemental searches for guidelines, frameworks, white papers
- AHRQ's Request for Information
- Technical Expert Panel and Key Informants

CQ 2: Who Develops Standards?

Academic researchers

- US
- UK
- Australia

Regulatory Agencies

- FDA
- NIST
- Canadian regulators

Think tanks

Corporations

Advocacy Groups

CQ 2: Examples

Good Machine Learning Practice for Medical Device Development: Guiding Principles

October 2021

The U.S. Food and Drug Administration (FDA), Health Canada, and the United Kingdom's Medicines and Healthcare products Regulatory Agency (MHRA) have jointly identified 10 guiding principles that can inform the development of Good Machine Learning Practice (GMLP). These guiding principles will help promote safe, effective, and high-quality medical devices that use artificial intelligence and machine learning (AI/ML).

Artificial Intelligence Risk Management Framework (AI RMF 1.0)

NIST NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY U.S. DEPARTMENT OF COMMERCE

Algorithmic Bias Playbook

Ziad Obermeyer
Rebecca Nissan
Michael Stern
Stephanie Eaneff
Emily Joy Bembeneck
Sendhil Mullainathan

June, 2021

Microsoft Responsible AI Standard, v2

GENERAL REQUIREMENTS

FOR EXTERNAL RELEASE

June 2022

CQ 2: Results

Table 14. Guidance, Standards, and Recommendations

Resources	Stakeholder	Summary of Content
Preventing bias and inequities in AI-enabled health tools.[99] **2022**	Academic researchers at Duke University Margolis Center for Health Policy	Authors identified 4 areas of algorithmic bias: 1) Inequitable framing of the healthcare challenge 2) Unrepresentative training data 3) Biased training data 4) Insufficient care with choices in data selection, curation, preparation, and model development They also offer recommendations for key stakeholders: Developers should recognize the potential for harm, follow good machine learning practices, work with diverse teams, and develop an understanding of the problem being solved, the data being used, potential differences across subgroups, and how the algorithm is likely to be used. Purchasers and users should test algorithms in their populations immediately and over time, focusing on patient outcomes. Health systems/payers/other owners of large health datasets should prioritize standardization reduce bias in subjective descriptions, and note where their data may differ across groups. FDA and other agencies should ensure that devices that use AI perform well for all subgroups, require clear, accessible labeling, and build systems to monitor for biased outcomes.
The medical algorithmic audit.[100] **2022**	Academic researchers based primarily in United Kingdom	Presents rationale (based on fairness and justice) and describes components of an algorithmic audit tailored to medicine. Expands on work of Raji by emphasizing intended use, intended impact, exploratory error analysis, subgroup testing, and adversarial testing in the context of healthcare.
Who audits the auditors? Recommendations from a field scan of the algorithmic auditing ecosystem.[101] **2022**	Algorithmic Justice League	Not specific to healthcare, focuses on AI. Presents 6 recommendations for policymakers: 1) Require the owners and operators of AI systems to engage in independent algorithmic audits against clearly defined standards 2) Notify individuals when they are subject to algorithmic decision-making systems 3) Mandate disclosure of key components of audit findings for peer review 4) Consider real-world harm in the audit process, including through standardized harm incident reporting and response mechanisms 5) Directly involve the stakeholders most likely to be harmed by AI systems in the algorithmic audit process 6) Formalize evaluation and, potentially, accreditation of algorithmic auditors.
Microsoft responsible AI standard, v2: general requirements.[102] **2022**	Microsoft	Microsoft's detailed standards for AI algorithms. Shaped around 6 core goals: accountability, transparency, fairness, reliability and safety, privacy and security, and inclusiveness. Numerous principles relevant to healthcare disparities, including: F2.1) Identify and prioritize demographic groups, including marginalized groups, that may be at risk of being differentially affected by the system based on intended uses and geographic areas where the system will be deployed. Include: 1) groups defined by a single factor and 2) groups defined by a combination of factors. F2.2) Evaluate all data sets to assess inclusiveness of identified demographic groups and collect data to close any gaps. F2.1) Reassess the system design, including the choice of training data, features, objective function, and training algorithm, to pursue the goals of minimizing differences between the rates at which resources and opportunities are allocated to identified demographic groups, paying particular attention to those that exceed the target maximum difference, while recognizing that doing so may appear to affect system performance and it is seldom clear how to make such tradeoffs. F2.1.1) For North America, use Best Practices for Age, Gender Identity, and Ancestry to help identify demographic groups and methods for collecting demographic information. F2.1.2) Work with user researchers to understand variations in demographic groups across intended uses and geographic areas. F2.1.3) Work with domain-specific subject matter experts to understand the facts that impact performance of your system and how they vary across identified demographic groups in this domain. F2.1.4) Work with members of identified demographic groups to understand risks of and impacts associated with differences between the rates at which resources and opportunities are allocated. F3.1) Identify and prioritize demographic groups, including marginalized groups, that may be at risk of being subject to stereotyping, demeaning, or erasing outputs of the system. Include: 1) groups defined by a single factor, and 2) groups defined by a combination of factors. F3.5) Reassess the system design, including the choice of training data, features, objective function, and training algorithm, to pursue the goal of minimizing the potential for stereotyping, demeaning, and erasing the identified demographic groups.
A proposal for identifying and managing bias in artificial intelligence.[103] **2021**	National Institute of Standards and Technology (NIST)	NIST has been developing the groundwork for consensus standards on bias in AI. The proposal is organized along 3 key stages: Pre-Design, Design and Development, and Deployment. Within each stage, the authors describe challenges that can introduce bias and suggest multiple potential solutions.
Algorithmic Bias Playbook[104] **2021**	Academic researchers at the University of Chicago Booth School of Medicine and the University of California Berkley School of Public Health	Describes a 4-step process (with multiple sub-steps) for researchers and institutions investigating any type of algorithm. Focuses heavily on harms associated with label bias. Step 1: Inventory algorithms 1A) Talk to relevant stakeholders about how and when algorithms are used. 1B) Designate a "steward" to maintain and update the inventory. Step 2: Screen for bias 2A) Articulate the ideal target (what the algorithm should be predicting) vs. the actual target (what it is actually predicting). 2B) Analyze and interrogate bias. Step 3: Retrain biased algorithms (or throw them out) 3A) Try retraining the model on a label closer to the ideal target. 3B) Consider alternative options. 3C) Consider suspending or discontinuing use of the algorithm. Step 4: Set up structures to prevent future bias 4A) Implement best practices for organizations working with algorithms.
Good machine learning practice for medical device development: guiding principles.[105] **2021**	US Food and Drug Administration, Health Canada, Medicines and Healthcare Products Regulatory Agency	Brief overview of 10 principles for medical device development driven by machine learning but broadly applicable to algorithms. 1) Multidisciplinary expertise is leveraged throughout the total product life cycle. 2) Good software engineering and security practices are implemented. 3) Clinical study participants and data sets are representative of the intended patient population. 4) Training data sets are independent of test sets. 5) Selected reference datasets are based on best available methods. 6) Model design is tailored to the available data and reflects the device's intended use.

National Institute on Minority Health and Health Disparities

CQ 2: Themes

Recent guidance focuses more on AI, less on traditional algorithms

Recommendations often focus on these issues:

- Representative datasets
- Diverse, multidisciplinary teams
- Transparency
- Accountability
- Fairness

National Institute on Minority Health and Health Disparities

CQ 2: Unanswered Questions

- Role of self-regulation vs. external requirements
- Third-party auditing
- Standards for who?
  - Algorithm developers
  - End-users
- "Off-label" use
- Scope of concern
  - Traditional algorithms
  - AI
  - Imaging
- Beyond race and ethnicity

National Institute on Minority Health and Health Disparities

CQ 2: References

Costanza-Chock S, Raji ID, Buolamwini J. Who audits the auditors? Recommendations from a field scan of the algorithmic auditing ecosystem. 2022 2022; New York (NY). 1268075: Association for Computing Machinery; pp. 1571–83.

Good machine learning practice for medical device development: guiding principles. Silver Spring (MD): U.S. Department of Health and Human Services, Food and Drug Administration; 2021. https://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles.

Liu X, Glocker B, McCradden MM, et al. The medical algorithmic audit. Lancet Digital Health. 2022 May;4(5):e384-e97. doi: 10.1016/S2589-7500(22)00003-6.

Locke T, Parker V, Thoumi A, et al. Preventing bias and inequities in AI-enabled health tools. Washington (DC): Duke-Margolis Center for Health Policy; 2022 Jul. https://healthpolicy.duke.edu/publications/preventing-bias-and-inequities-ai-enabled-health-tools.

CQ 2: References

Microsoft responsible AI standard, v2. General requirements. Redmond (WA): Microsoft Corporation; 2022. https://blogs.microsoft.com/wp-content/uploads/prod/sites/5/2022/06/Microsoft-Responsible-AI-Standard-v2-General-Requirements-3.pdf.

Obermeyer Z, Nissan R, Stern M, et al. Algorithmic bias playbook. Chicago (IL): The University of Chicago Booth School of Business, The Center for Applied Artificial Intelligence; 2021 Jun

Schwartz R, Down L, Jonas A, et al. A proposal for identifying and managing bias in artificial intelligence. Draft NIST special publication 1270. Gaithersburg (MD): National Institute of Standards and Technology; 2021 Jun. doi: 10.6028/NIST.SP.1270-draft.



AGENCY FOR HEALTHCARE RESEARCH AND QUALITY

AHRQ

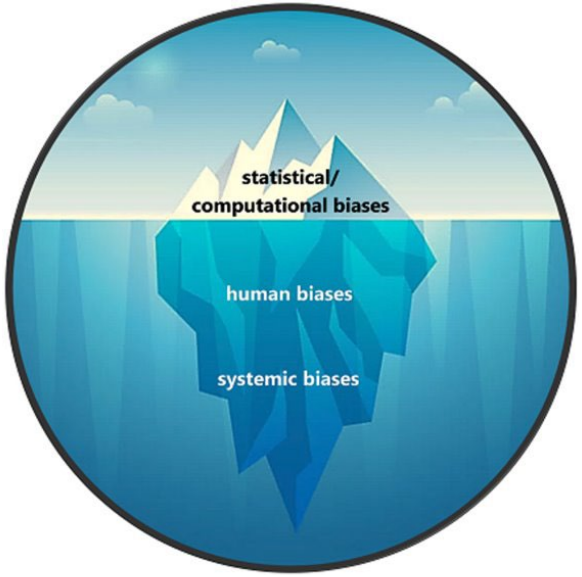
Managing AI Bias with the NIST AI Risk Management Framework: A Socio-Technical Approach

Reva Schwartz
Principal Investigator – Bias in AI
National Institute of Standards and Technology

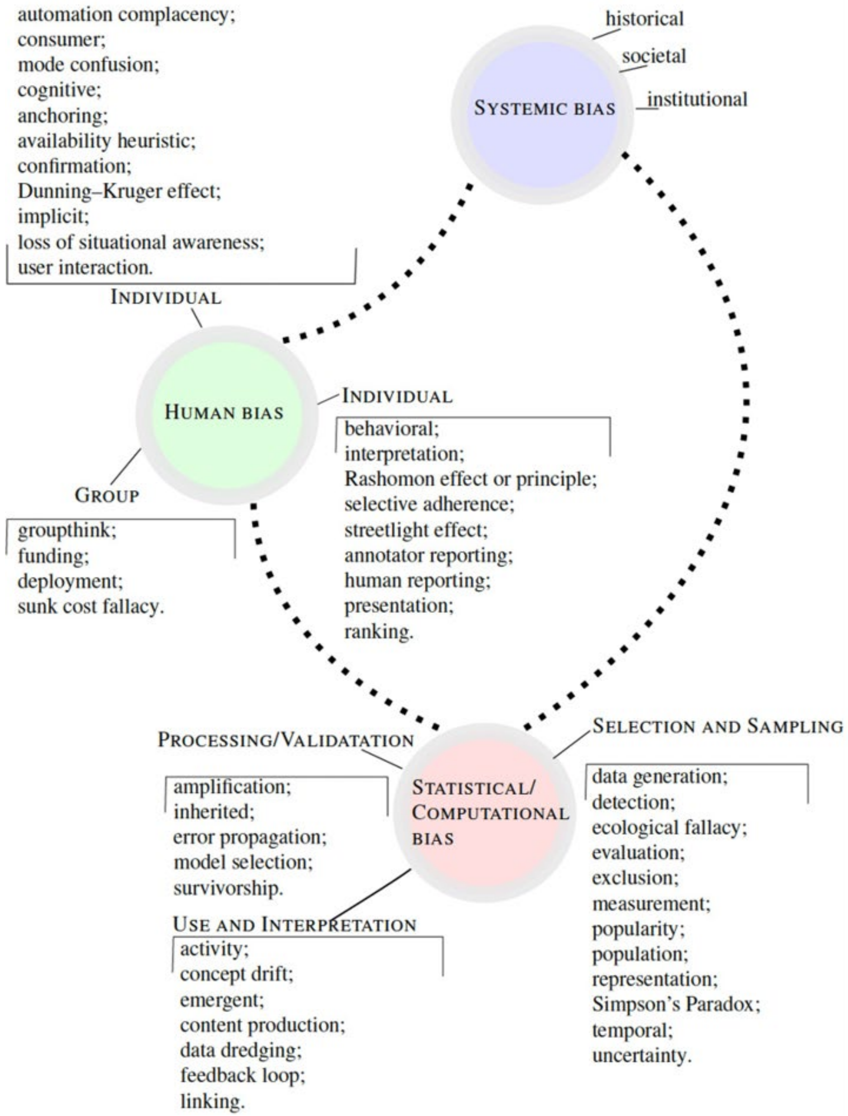
Racial Bias and Healthcare Algorithms
March 3, 2023
12:35 - 12:43 p.m. ET

NIH National Institute on Minority Health and Health Disparities

Taxonomy of AI Bias



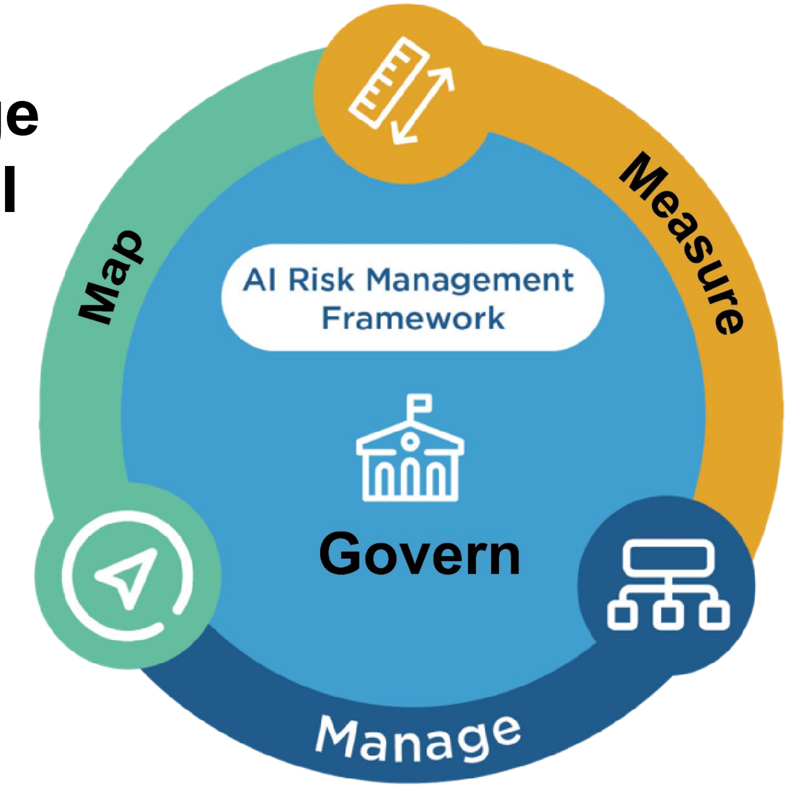
Current focus on computational/statistical bias obfuscates the other two categories



National Institute
on Minority Health
and Health Disparities

What is the AI RMF?

Voluntary resource for organizations designing, developing, deploying, or using AI systems to manage AI risks and promote trustworthy and responsible AI



Flexibly applied

Rights-preserving

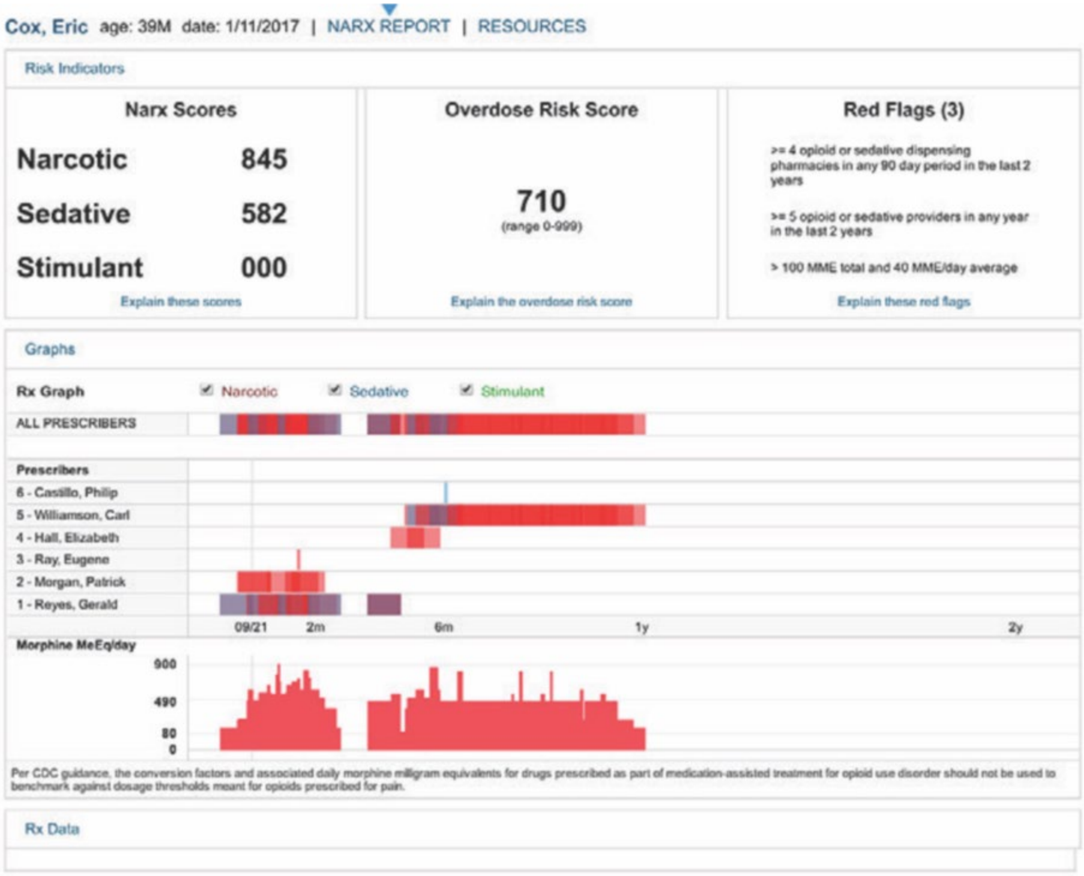
Measurable

How can the AI RMF help organizations manage the risks from AI bias?

- Operationalizes trustworthiness and societal values
- AI Governance and organizational culture
- Socio-technical approach
- Risks and impacts focused

Example: Why a socio-technical framing matters for managing bias

AHRQ Agency for Healthca Research and Qualit



Example of NarxCare's overdose risk score.

https://technophilosoph.com/en/2021/09/28/the-problem-with-an-ai-based-overdose-risk-score-for-pain-medications/

See also: https://www.wired.com/story/opioid-drug-addiction-algorithm-chronic-pain/

- model uses proxy for narcotic use (referred to as "addictiveness") that is inherently unobservable
- resulting in
  - people being denied medication that they genuinely needed,
  - bias - women – who are more likely to experience chronic pain – were more likely to be wrongly identified as "highly addictive"

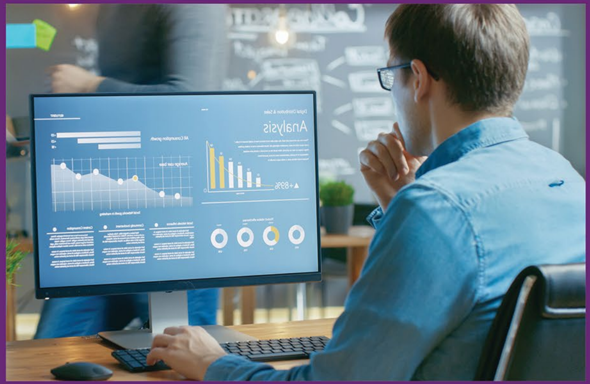
NIH National Institute on Minority Health and Health Disparities

THANK YOU

Contact us via email at aiframework@nist.gov

For more info on the NIST AI RMF, visit https://www.nist.gov/itl/ai-risk- management-framework

National Institute on Minority Health and Health Disparities


AGENCY FOR HEALTHCARE RESEARCH AND QUALITY

AHRQ

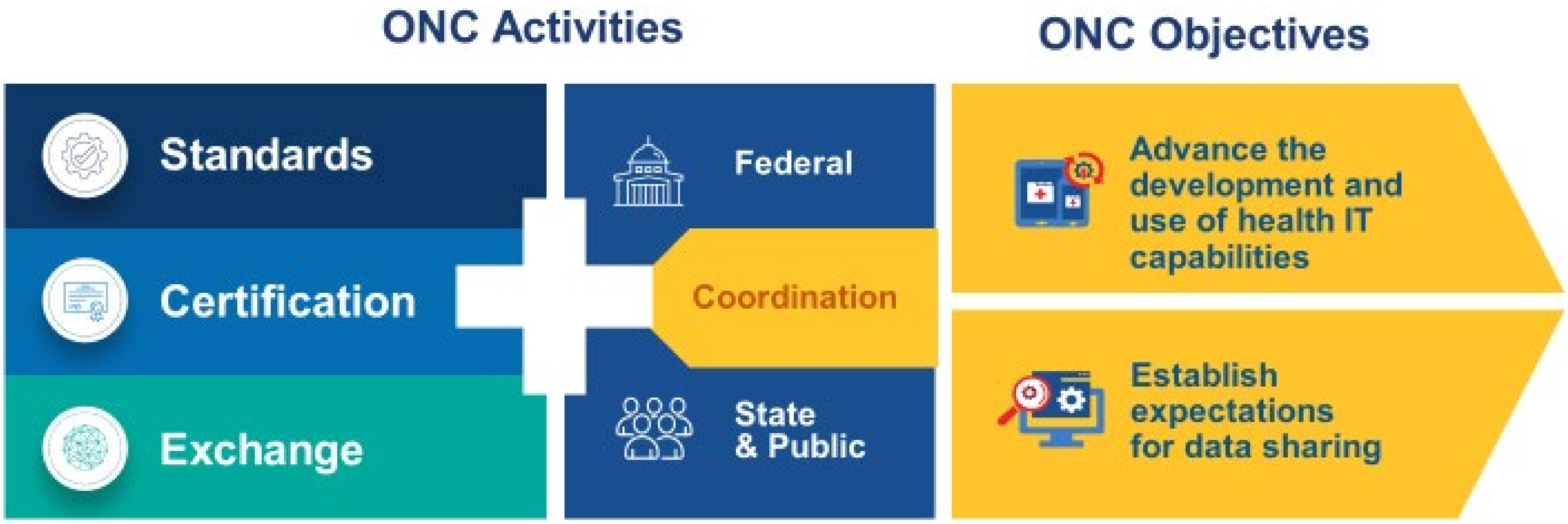
Understanding AI + Health IT

Stephen Konya
ONC

Racial Bias and Healthcare Algorithms
March 3, 2023
12:43 - 12:51 p.m. ET

NIH National Institute on Minority Health and Health Disparities

ONC – Understanding AI + Health IT

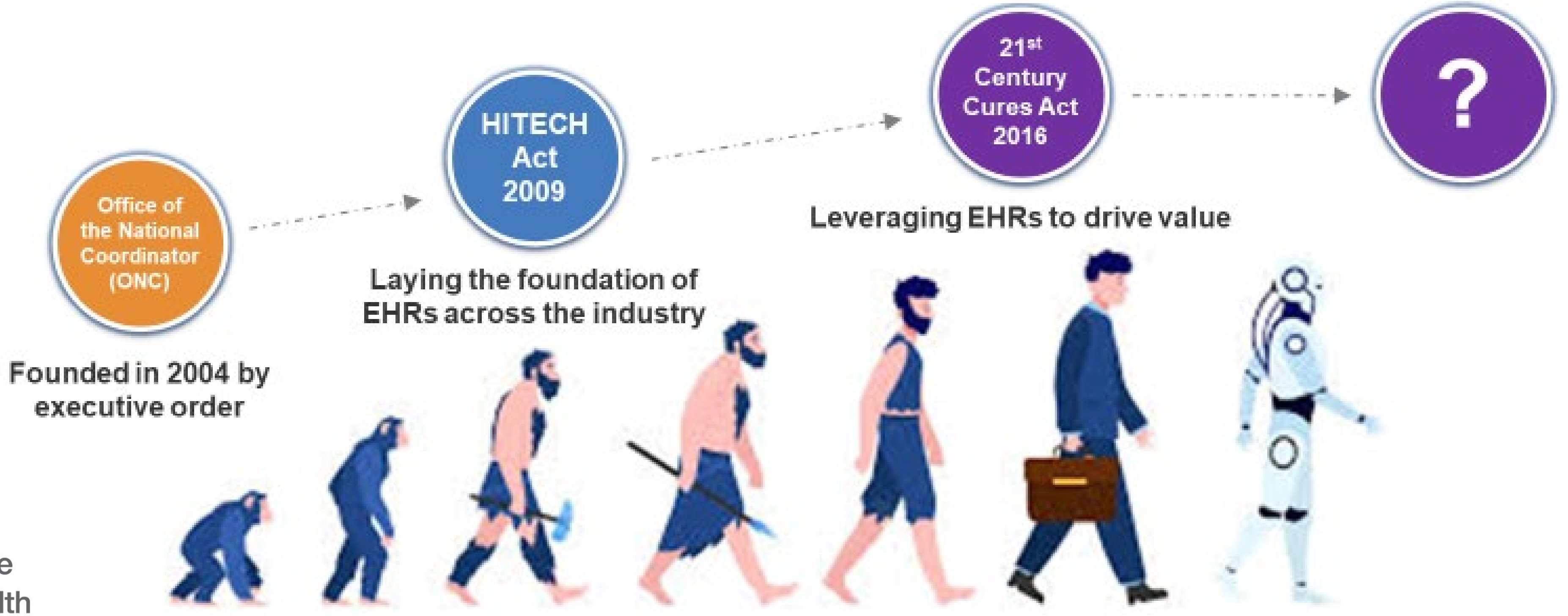


National Institute on Minority Health and Health Disparities

ONC – Understanding AI + Health IT

The Evolution of Health IT & Digital Health (in the US)

ONC is charged with formulating the **federal government's health IT strategy** to advance national goals for better and safer health care through **an interoperable nationwide health IT infrastructure**



ONC – Understanding AI + Health IT

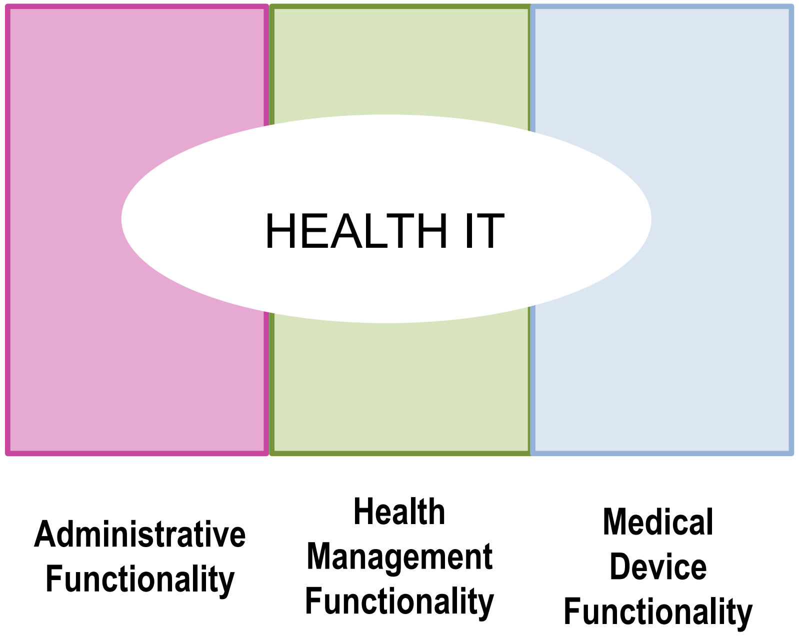
FDASIA Report (2014) – FDA, FCC, & ONC

Proposed strategy and recommendations

- Are based on the premise that risk and corresponding controls should focus on health IT functionality – not the platform(s) on which such functionality resides or the product name/ description of which it is a part.
- Seek to advance a framework that is relevant to current functionalities and technologies yet sufficiently flexible to accommodate the future and rapid evolution of health IT.

Agencies' proposed strategy identifies three categories of health IT:

1) administrative health IT functions
2) health management health IT functions, and
3) medical device health IT functions

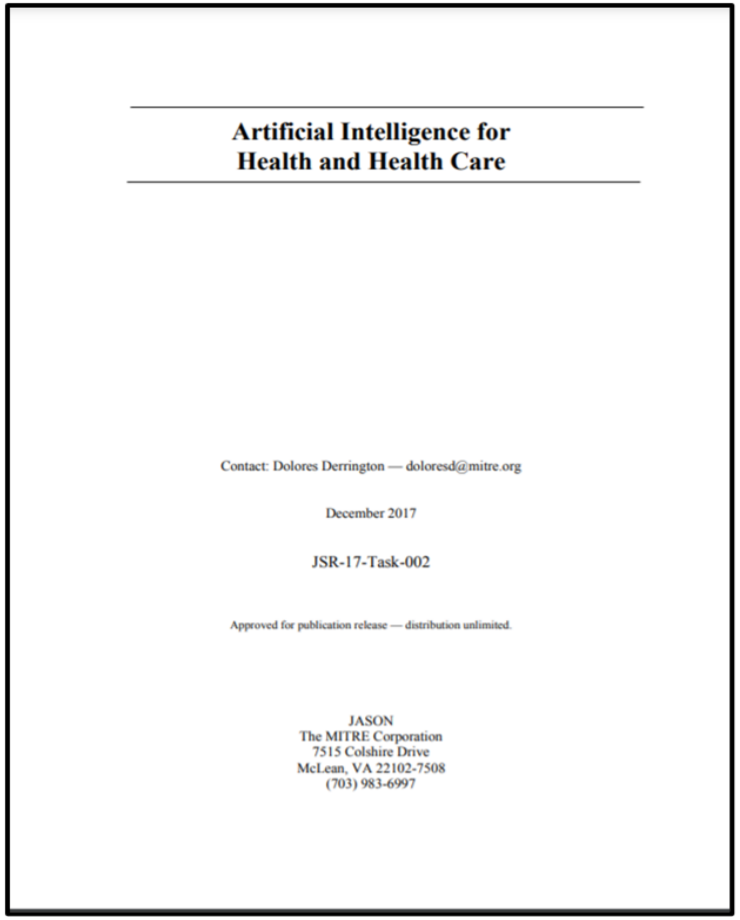


Source: FDASIA Health IT Report – Proposed Strategy & Recommendations for a Risk-Based Framework (ONC, FDA, FCC)

ONC – Understanding AI + Health IT

JASON Report (2017) – ONC, AHRQ, and RWJF:
“Artificial Intelligence for Health and Health Care”

- JASON supports the collection and curation of new health data sources for AI applications as well. For example:
  - ► Capturing smartphone data
  - ► Integrating social and environmental data
  - ► Supporting AI competitions
- The recommendations in the new report underline the importance of ONC’s efforts toward interoperable and standardized health data and AHRQ’s efforts to effectively use those data to improve the quality and safety of patient care.
- These efforts will improve capabilities to exchange and appropriately use high-quality health data – critical elements in powering AI efforts in health and healthcare.


Artificial Intelligence for
Health and Health Care

Contact: Dolores Derrington — doloresd@mitre.org

December 2017

JSR-17-Task-002

Approved for publication release — distribution unlimited.

JASON
The MITRE Corporation
7515 Colshire Drive
McLean, VA 22102-7508
(703) 983-6997

Health IT Buzz Blog:
https://www.healthit.gov/buzz-blog/interoperability/hype-reality-artificial-intelligence-ai-transform-health-healthcare

ONC – Understanding AI + Health IT

An ONC Artificial Intelligence Showcase (2022) -

"Seizing the Opportunities and Managing the Risks of Use of AI in

30+ showcase presentations grouped into the following 3 categories;

- *Advancing Responsible Ai in Health IT – Guiding Principles*
- *Transparency and Accountability*
- *Evaluating Data Input Needs & Real-World Performance*

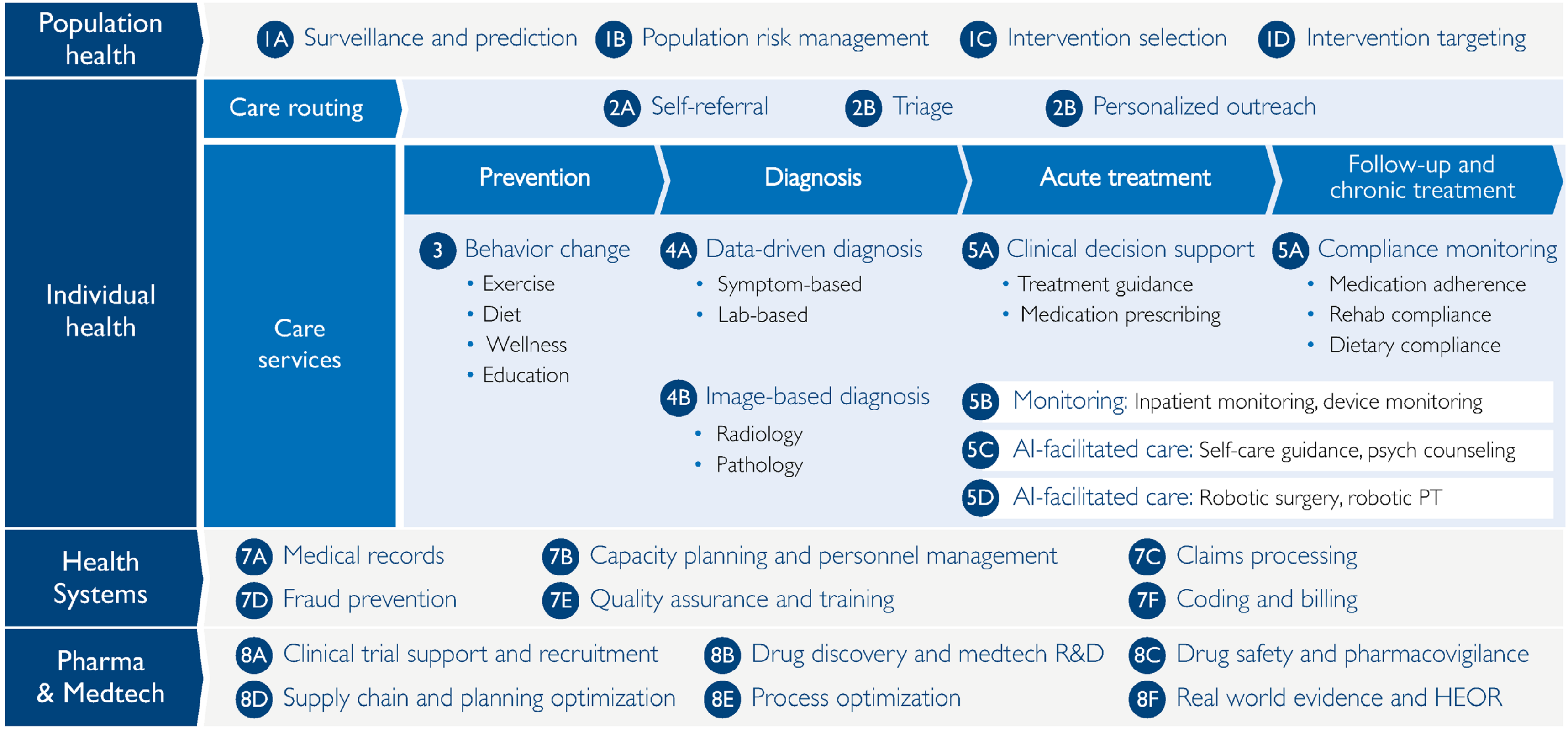
➢ View Agenda [PDF - 246 KB]
➢ View Presentation Slides [PDF - 12 MB]
➢ **Watch Event Recording**

National Institute on Minority Health and Health Disparities

More details here:
https://www.healthit.gov/news/events/onc-artificial-intelligence-showcase-seizing-opportunities-and-managing-risks-use-ai

ONC – Understanding AI + Health IT

Figure 2: Framework of all AI Use Cases in Healthcare

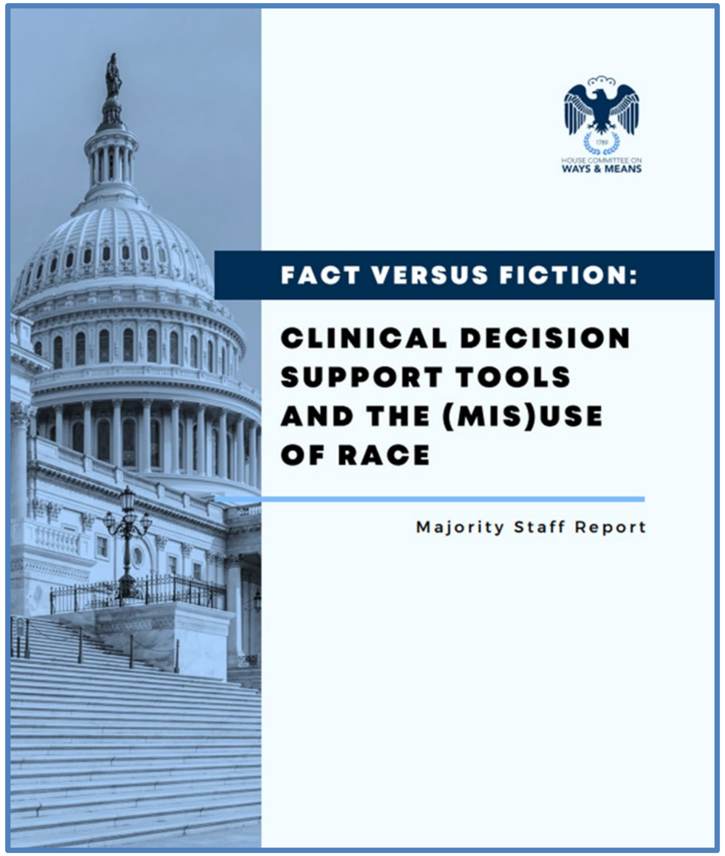


National Institute on Minority Health and Health Disparities

*Figure from USAID's "Artificial Intelligence in Global Health: Defining a Collective Path Forward" https://www.usaid.gov/cii/ai-in-global-health

ONC – Understanding AI + Health IT

House Ways and Means Committee Report:
“Clinical Decision Support Tools (CDST) and the (Mis)use of Race”



(access full report here)

Key Findings:

- Respondents acknowledged the unacceptable nature of findings that CDSTs produce avoidable differences for patients of color
  - ► One-third of respondents said they are not planning to reevaluate use of race and ethnicity in clinical algorithms
- Raised the absence of a central hub of accountability as a barrier to addressing these complex issues across scientific and medical professions
  - ► Some recommended leadership from largest and most influential organizations (e.g., the Centers for Medicare & Medicaid Services) must assemble stakeholders to develop standards, guidance, and best practices for using race in CDSTs.
- Emphasized role of bias in CDST development and care delivery, suggesting solution lies upstream (e.g., at the level of health technology research and development and through clinician education)
- Strategies must be enacted to proactively correct and confront the challenges of the misuse of race and ethnicity in CDSTs

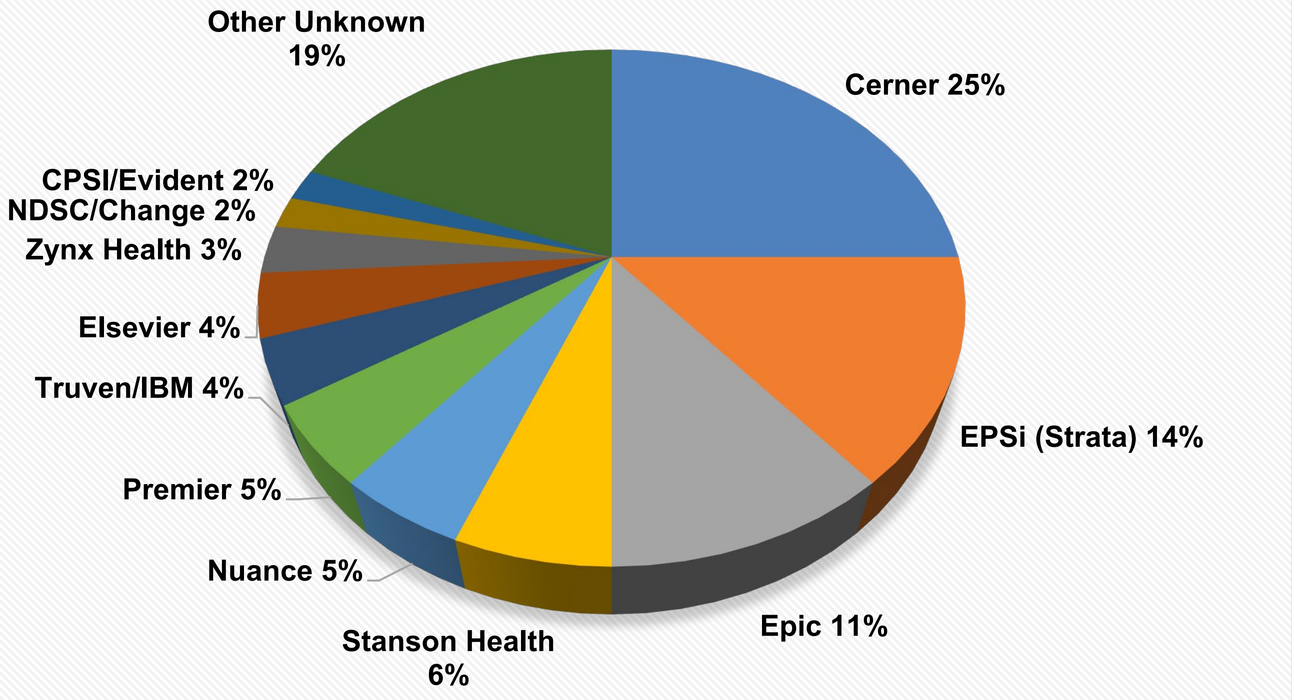
ONC – Understanding AI + Health IT

#1 - A Lack of Transparency

"Many CDSTs are proprietary, making it difficult for independent researchers to evaluate and validate these tools to ensure they function as intended and do not disadvantage certain patients."

Clinical Decision Support – 2018, REACTION DATA (2019), https://www.reactiondata.com/report/clinical-decision-support-2018/.

Figure 5. CDST Vendor Market Share



National Institute on Minority Health and Health Disparities

#2 - Relationship Between CDSTs and Health Equity

- While CDSTs have vastly improved medicine, they also remain vulnerable to implicit and explicit biases inherent to big data
- There is significant potential for the more advanced tools to address racial inequalities if care is taken to acknowledge the underlying data's susceptibility to bias and proactively "clean" the data.
- Because big data replicates or amplifies human biases, adding its elements to CDST architecture with questionable foundations can yield unintended, yet avoidable consequences

Source: House Ways and Means Committee Report: "Clinical Decision Support Tools (CDST) and the (Mis)use of Race"

ONC – Understanding AI + Health IT

#3 – CDSTs# Race/Equity Related Findings:
NEJM Study (June 2020)

- Racial correction in clinical algorithms is harmful for a range of conditions, from childbirth to cancer.
- Study authors concluded that race had been misinterpreted or misused in multiple CDSTs, resulting in worse outcomes for people of color.
- Incorporating race data into clinical algorithms can entrench disparities by potentially producing different treatment approaches for individuals that are not based on precision medicine but are simply chosen because of race/ethnicity, historical differences in outcomes based on race, discrimination, racism, and biases about race.

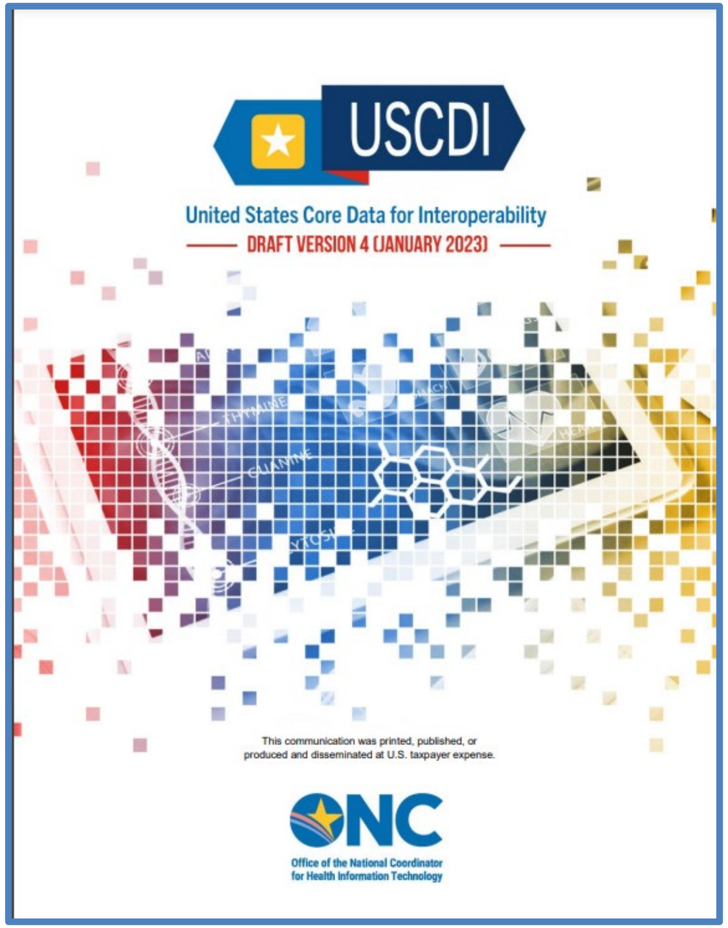
Source: House Ways and Means Committee Report: "Clinical Decision Support Tools (CDST) and the (Mis)use of Race"

National Institute on Minority Health and Health Disparities

ONC – Understanding AI + Health IT

Agency for Healthcare Research and Quality

US Core Data for Interoperability (USCDI):
The Minimum Dataset of the Health Care Delivery System



- **ONC standard for minimum dataset required for interoperability**
  - Defines required data elements and vocabulary standards
  - Agnostic to format
- **Updated on annual cycle with federal agency and industry input**
  - Updates based on multiple criteria including standards maturity and public/industry priority

NIH National Institute on Minority Health and Health Disparities

ONC – Understanding AI + Health IT

USCDI

USCDI Version 3 (July 2022)

With the [publication of Draft USCDI v4](), ONC is accepting feedback on its content until April 17, 2023.

ONC plans on releasing a final USCDI v4 in July 2023.

National Institute on Minority Health and Health Disparities

Allergies and Intolerances
- Substance (Medication)
- Substance (Drug Class)
- Reaction

Assessment and Plan of Treatment
- Assessment and Plan of Treatment
- SDOH Assessment

Care Team Member(s)
- Care Team Member Name
- Care Team Member Identifier
- Care Team Member Role
- Care Team Member Location
- Care Team Member Telecom

Clinical Notes
- Consultation Note
- Discharge Summary Note
- History & Physical
- Procedure Note
- Progress Note

Clinical Tests
- Clinical Test
- Clinical Test Result/Report

Diagnostic Imaging
- Diagnostic Imaging Test
- Diagnostic Imaging Report

Encounter Information
- Encounter Type
- Encounter Diagnosis
- Encounter Time
- Encounter Location
- Encounter Disposition

Goals
- Patient Goals
- SDOH Goals

Health Insurance Information ★
- Coverage Status ★
- Coverage Type ★
- Relationship to Subscriber ★
- Member Identifier ★
- Subscriber Identifier ★
- Group Number ★
- Payer Identifier ★

Health Status/ Assessments
- Health Concerns ➡
- Functional Status ★
- Disability Status ★
- Mental / Cognitive Status ★
- Pregnancy Status ★
- Smoking Status ➡

Immunizations
- Immunizations

Laboratory
- Test
- Values/Results
- Specimen Type ★
- Result Status ★

Medications
- Medications
- Dose ★
- Dose Unit of Measure ★
- Indication ★
- Fill Status ★

Patient Demographics/ Information
- First Name
- Last Name
- Middle Name (Including middle initial)
- Suffix
- Previous Name
- Date of Birth
- Date of Death ★
- Race
- Ethnicity
- Tribal Affiliation ★
- Sex
- Sexual Orientation
- Gender Identity
- Preferred Language
- Current Address
- Previous Address
- Phone Number
- Phone Number Type
- Email Address
- Related Person's Name ★
- Related Person's Relationship ★
- Occupation ★
- Occupation Industry ★

Problems
- Problems
- SDOH Problems/Health Concerns
- Date of Diagnosis
- Date of Resolution

Procedures
- Procedures
- SDOH Interventions
- Reason for Referral ★

Provenance
- Author Organization
- Author Time Stamp

Unique Device Identifier(s) for a Patient's Implantable Device(s)
- Unique Device Identifier(s) for a patient's implantable device(s)

Vital Signs
- Systolic blood pressure
- Diastolic blood pressure
- Heart Rate
- Respiratory rate
- Body temperature
- Body height
- Body weight
- Pulse oximetry
- Inhaled oxygen concentration
- BMI Percentile (2 - 20 years)
- Weight-for-length Percentile (Birth - 24 Months)
- Head Occipital-frontal Circumference Percentile (Birth - 36 Months)

★ New Data Classes and Elements ➡ Data Element Reclassified

ONC – Understanding AI + Health IT

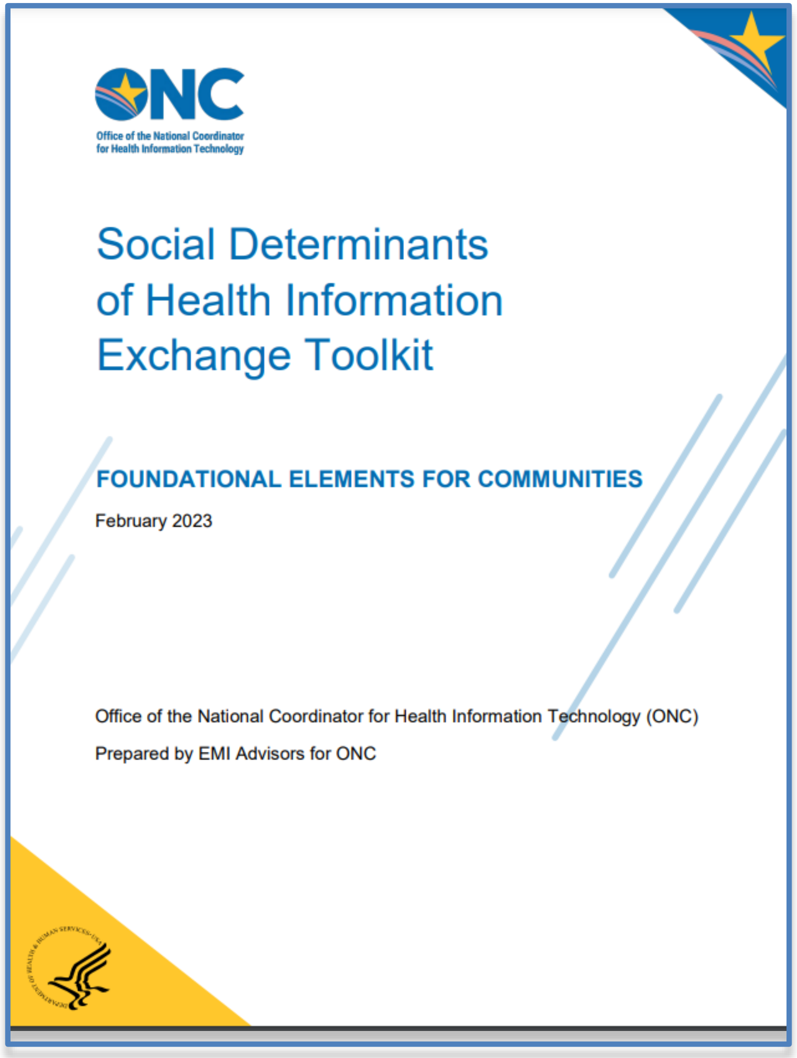
New Releases:

SDOH Toolkit and Learning Forum Sessions for the Health IT Community

1) Health IT Buzz [Blog post]

2) [Register for the 2023 series], covering;

- Community Level Governance
- Values, Principles, and Privacy
- Implementation, Measurement and Evaluation
- SDOH Information Exchange Learning Forum Summary


ONC
Office of the National Coordinator for Health Information Technology

Social Determinants of Health Information Exchange Toolkit

FOUNDATIONAL ELEMENTS FOR COMMUNITIES

February 2023

Office of the National Coordinator for Health Information Technology (ONC)

Prepared by EMI Advisors for ONC

National Institute on Minority Health and Health Disparities

ONC – Understanding AI + Health IT

FHITCC - AI in Health IT Survey – January 2022

- Through its role as the National Coordinator for Health IT, and by leveraging the FHITCC, ONC conducted an internal interagency assessment surrounding the use of Artificial Intelligence (Ai) in the field of health IT
- The survey aimed to achieve the following;
  - compile a robust list of federal activities and initiatives currently underway
  - collect insights on key areas of interest and focus priorities for those agencies
  - and attempt to identify specific point of contacts for Ai related work located within each agency
- This activity was designed to compliment several other federal activities that were/are also intended to gain a better understanding of the current and planned federal landscape with respect to the responsible use of AI in Health IT. (*i.e., activities being led by NAIIO/OSTP, OMB, OCIO/OCAIO, HRSA, AHRQ, ASPE, OCR, FDA, NIH, FTC, NIST, VA, etc.)*

National Institute on Minority Health and Health Disparities

Federal Agency Representation

22 Different Agencies Responded		
ACF	FDA	OCIO / OCAIO
AHRQ	FTC*	OCR
CDC	HRSA	ONC
CMS	IHS	SSA*
DHA*	NASA*	USAID*
DoE*	NIH	VA*
DoS*	NIST*	
DoT*	OASH	

*Non-HHS Agencies

ONC – Understanding AI + Health IT

Based on your knowledge, which of the following use cases for the deployment of Ai in healthcare, are considered as areas of potential interest for your agency? (check all that apply)

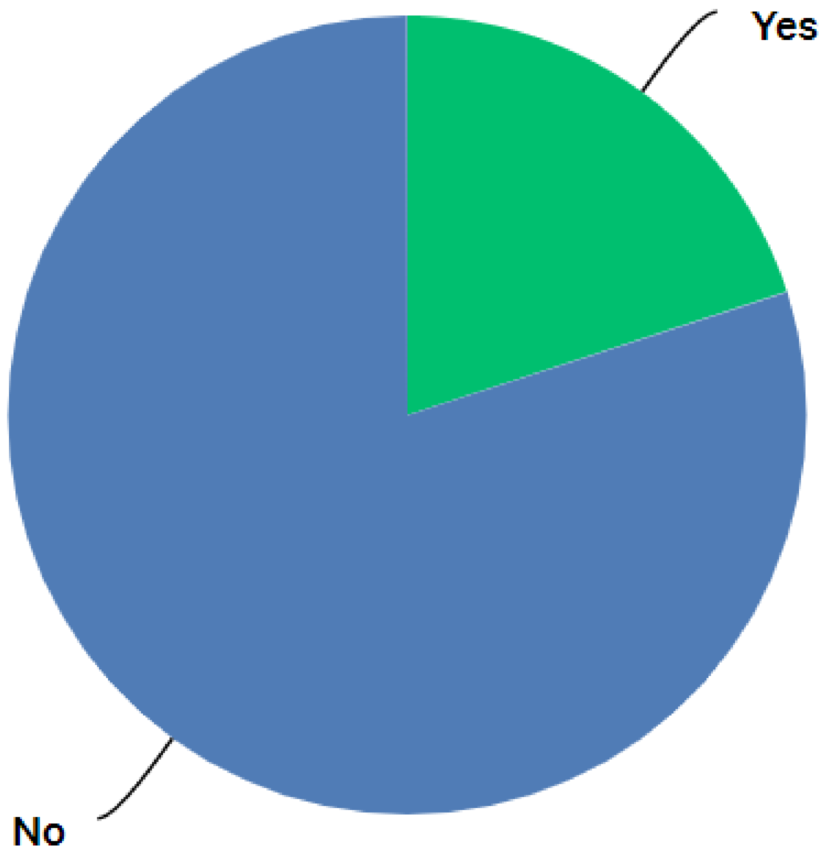
ANSWER CHOICES		RESPONSES
Research		70.59%
Bias / Equity		64.71%
Clinical Care / Clinical Decision Support Tools (CDSTs)		64.71%
Natural Language Processing (NLP) / Voice Tech / Conversational Ai / Chatbots		58.82%
Predictive Analytics (clinical)		58.82%
Medical Device / Diagnostics		52.94%
Population / Public Health		52.94%
Human / Social Services		47.06%
Imaging		47.06%
Predictive Analytics (business)		47.06%
Other (please specify)	Responses	41.18%
Administrative / Operational Functions		35.29%
Genomics / Precision Medicine		29.41%

National Institute on Minority Health and Health Disparities

ONC – Understanding AI + Health IT

Has your agency publicly issued any official guidance, policy statements, regulatory requirements, and/or strategy relating to the design, development or use of Ai within healthcare?

Yes	23%
No	77%



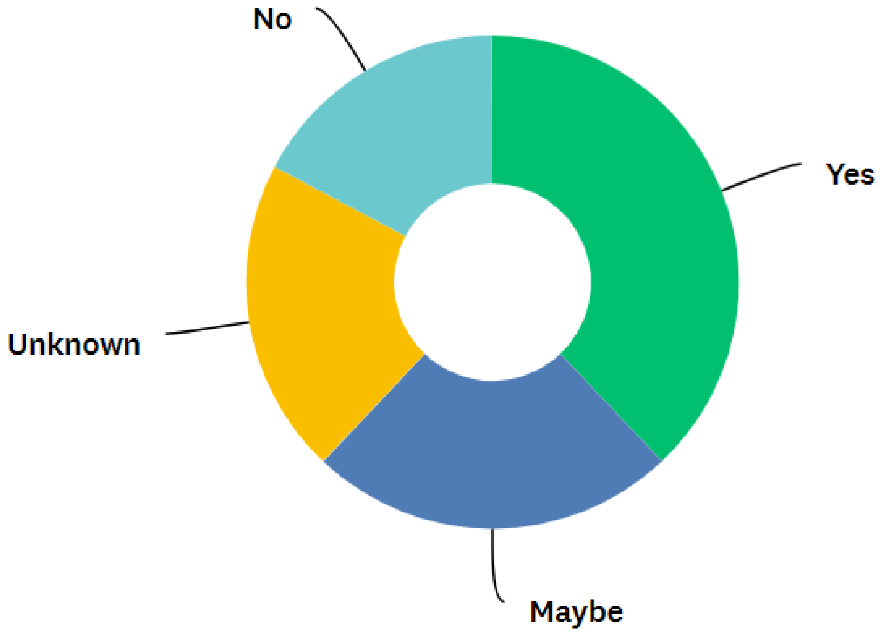
National Institute on Minority Health and Health Disparities

Health Equity by Design / AI Bias

As part of ONC's focus on ensuring health equity by design with regard to the adoption and use of technology, we are specifically interested in identifying work underway by federal agencies to try and mitigate the potential for systemic biases that could be exacerbated by the use of algorithms in health care and human service settings.

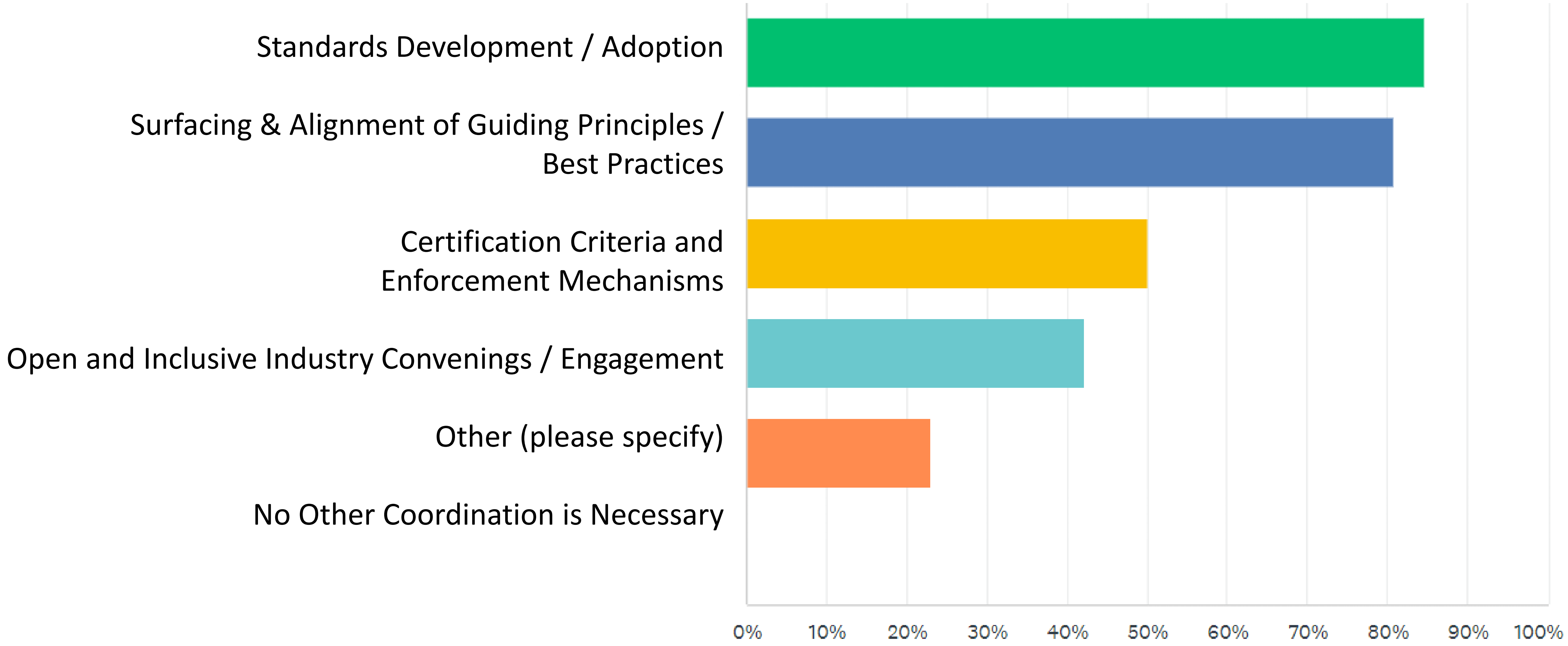
Please indicate if your agency is currently, or plans to pursue work in this area:

Yes	38%
Maybe	24%
Unknown	21%
No	17%



ONC – Understanding AI + Health IT

Please indicate your interest in the following areas of potential need for federal coordination when it comes to the adoption and use of Ai specifically in the practice of health care.



ONC – Understanding AI + Health IT

ONC's Health IT Buzz Blog Series: Artificial Intelligence & Machine Learning

- Back to the Future: What Predictive Decision Support Can Learn from DeLoreans and The Big Short (Dec. 2022)
- Two Sides of the AI/ML Coin in Health Care (Oct. 2022)
- Minimizing Risks and Maximizing Rewards from Machine Learning (Sep. 2022)
- Getting the Best out of Algorithms in Health Care (Jun. 2022)

National Institute
on Minority Health
and Health Disparities

Thank you!

Subscribe to our weekly eblast
at healthit.gov for the latest updates!

Email: Stephen.Konya@hhs.gov

Phone: 202-690-7151

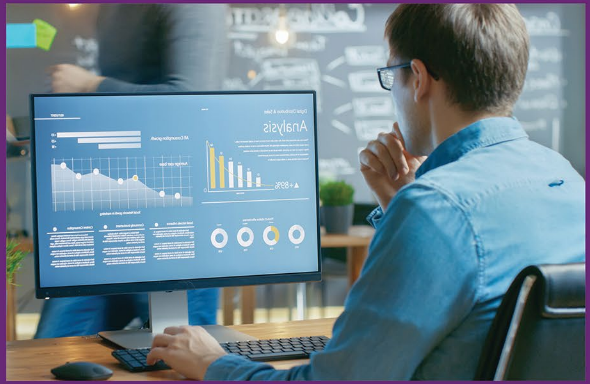
Health IT Feedback Form:
https://www.healthit.gov/form/healthit-feedback-form

Twitter: @onc_healthIT

LinkedIn: Office of the National Coordinator for Health Information Technology

Youtube:
https://www.youtube.com/user/HHSONC

National Institute
on Minority Health
and Health Disparities



AGENCY FOR HEALTHCARE RESEARCH AND QUALITY

AHRQ

Artificial Intelligence/Machine Learning (AI/ML)-Enabled Medical Devices

Matthew Diamond MD, MPH
Chief Medical Officer, Digital Health Center of Excellence
Center for Devices & Radiological Health (CDRH), US FDA

Racial Bias and Healthcare Algorithms
March 3, 2023
12:51- 12:59 p.m. ET

NIH National Institute on Minority Health and Health Disparities

FDA's Collaborative Patient-Centered Approach to AI/ML-Enabled Medical Devices

Agency for Healthcare Research and Quality

Recent Milestones

2019

Published AI/ML-SaMD Discussion Paper

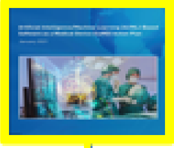
First joined Collaborative Community related to AI/ML

Evolving Role of AI in Radiological Imaging

Public Workshop on AI/ML in Radiological Imaging

2020

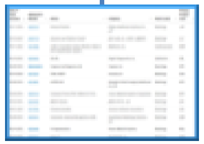
PEAC Mtg on Patient Trust in AI/ML Devices



Published AI/ML Medical Device Software Action Plan

2021

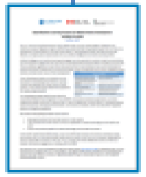
Posted List of Currently Marketed AI/ML Devices



Transparency of AI/ML-Enabled Medical Devices

Public Workshop on Transparency of AI/ML Devices

Published GMLP Guiding Principles



2

2022

Final Document

Machine Learning-enabled Medical Devices: Key Terms and Definitions

Contributed to IMDRF's Key Terms and Definitions

FINAL GUIDANCE

Clinical Decision Support (CDS) Software

Published Final Guidance: Clinical Decision Support

Coming Soon

AI/ML Guidance

DRAFT GUIDANCE

Marketing Submission Recommendations for a Change Control Plan

We're working collaboratively with stakeholders to build a proactive, patient-centered approach to AI/ML-enabled devices that promotes health equity.

NIH National Institute on Minority Health and Health Disparities

AI/ML Medical Device Software Action Plan

- Holistic, patient-centered approach to AI/ML-enabled devices
- Invitation for collaboration with broad set of stakeholders
- Five Aims encompassing regulatory policy and science

Five Aims:

- ❑ Update the proposed AI/ML framework
- ❑ Strengthen FDA's role in harmonizing GMLP
- ❑ Foster a patient-centered approach
- ❑ Support development of regulatory science methods
- ❑ Advance real-world performance pilots

National Institute on Minority Health and Health Disparities

Agency for Healthcare Research and Quality

Patient-Centered Approach Incorporating Transparency to Users

AI/ML-enabled devices have unique considerations that necessitate a proactive patient-centered approach:

- that takes into account issues including usability, equity, trust, and accountability
- that promotes transparency to all users and to patients more broadly

Patient Engagement Advisory Committee (PEAC) Meeting held Oct 2020

Workshop on Transparency of AIML-enabled devices held Oct 2021

National Institute on Minority Health and Health Disparities

Regulatory Science Supporting FDA's Strategic Priority to Promote Health Equity

- Collaborate on regulatory science efforts to develop evaluation methods for AI/ML-enabled medical software, especially related to algorithm bias and robustness.
- Ensure important performance considerations – including with respect to race, ethnicity, disease severity, gender, age, and geographical considerations – are addressed throughout the total product lifecycle
- Facilitate more rapid and continuous improvement of AI/ML-enabled device performance across diverse populations
- Ongoing research being conducted in collaboration with Centers for Excellence in Regulatory Science and Innovation (CERSIs) and within FDA's OSEL.
- In Collaborative Communities we work together to achieve common objectives on medical device challenges and promote health equity.

CERSI Johns Hopkins University

CERSI UCSF-Stanford

NIH National Institute on Minority Health and Health Disparities

We provide stakeholders ongoing clarity through guidance, including on:

AHRQ Agency for Healthca Research and Qualit

- Collection of Race and Ethnicity Data in Clinical Trials
- Evaluation of Sex-Specific Data in Medical Device Clinical Studies
- Patient Engagement in the Design and Conduct of Medical Device Clinical Studies
- Evaluation and Reporting of Age-, Race-, and Ethnicity-Specific Data in Medical Device Clinical Studies
- Clinical Decision Support Software
- Software as a Medical Device (SAMD): Clinical Evaluation
- Clinical Performance Assessment: Considerations for Computer-Assisted Detection Devices Applied to Radiology Images and Radiology Device Data in Premarket Notification Submissions

NIH National Institute on Minority Health and Health Disparities

Internationally Harmonized GMLP Guiding Principles

- **Good Machine Learning Practice (GMLP):** accepted practices in AI/ML algorithm design, development, training, and testing that can facilitate the quality development and assessment of AI/ML-enabled technologies

- Ten guiding principles issued by US FDA, MHRA (UK) and Health Canada to promote global harmonization in efforts for the identification of best practices and the creation of standards

- Intended to help inform the develop-ment of GMLP and encourage broad stakeholder engagement

U.S. FOOD & DRUG ADMINISTRATION | Health Canada Santé Canada | Medicines & Healthcare products Regulatory Agency

Good Machine Learning Practice for Medical Device Development: Guiding Principles	
Multi-Disciplinary Expertise are Leveraged Throughout the Total Product Life Cycle	Good Software Engineering and Security Practices are Implemented
Clinical Study Participants and Data Sets are Representative of the Intended Population	Training Data Sets are Independent of Test Sets
Selected Reference Datasets are Based Upon Best Available Methods	Model Design is Tailored to the Available Data and Reflects the Intended Use of the Device
Focus is Placed on the Performance of the Human-AI Team	Testing Demonstrates Device Performance during Clinically Relevant Conditions
Users are Provided Clear, Essential Information	Deployed Models are Monitored for Performance and Re-training Risks are Managed

National Institute on Minority Health and Health Disparities

Further Questions or Feedback:

www.fda.gov/digitalhealth

DigitalHealth@fda.hhs.gov

Matthew Diamond, MD, PhD

Chief Medical Officer

CDRH Digital Health Center of Excellence

Office of Strategic Partnerships & Technology Innovation (OST)

Center for Devices and Radiological Health (CDRH), U.S. Food and Drug Administration

matthew.diamond@fda.hhs.gov

(301) 332-5126

National Institute on Minority Health and Health Disparities

AGENCY FOR HEALTHCARE RESEARCH AND QUALITY

AHRQ

Digital Equity - A Global Perspective

Dr. Bilal A Mateen, MBBS, MPH
Wellcome Trust

Racial Bias and Healthcare Algorithms
March 3, 2023
12:59 – 1:07 p.m. ET

NIH National Institute on Minority Health and Health Disparities

A Global Perspective on Bias in AI

Characteristics of trustworthy data science

Technical, institutional and social characteristics need to be considered at every stage of a development process

Technical

The properties and effectiveness of data science and digital tools

- Tools must be effective in solving the problems they were created to address, in the intended context or environment
- Software should be easy to maintain, open by default, and can evolve to meet changing needs

Institutional

The policy and regulatory environment

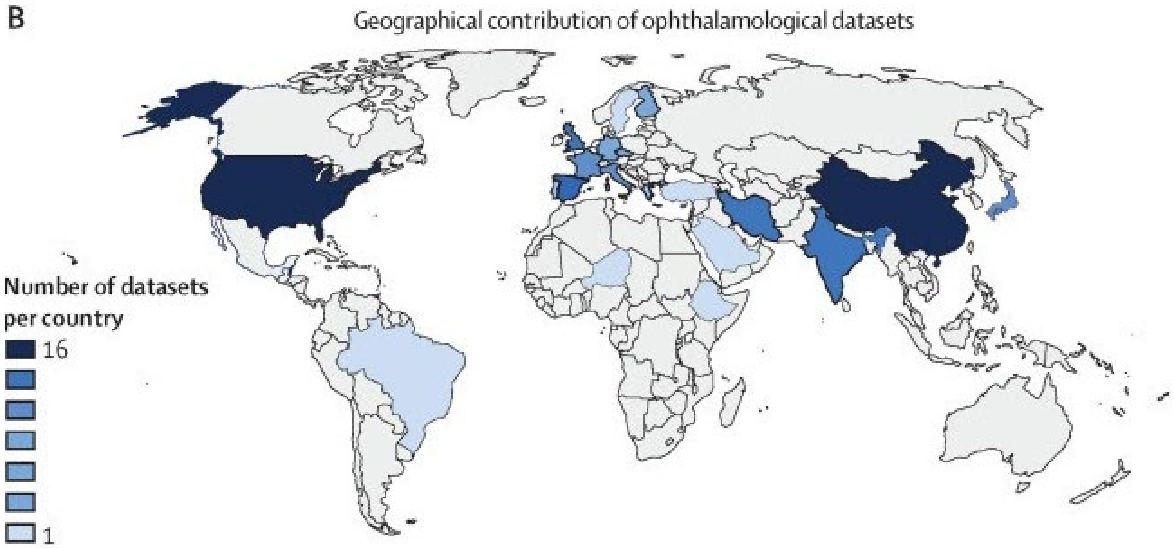
- Policy needs to facilitate responsible and sustainable innovation
- Data standards should support efficient data exchange and analysis, including across borders
- Governance must balance the need for adequate control over data, with access for science and research

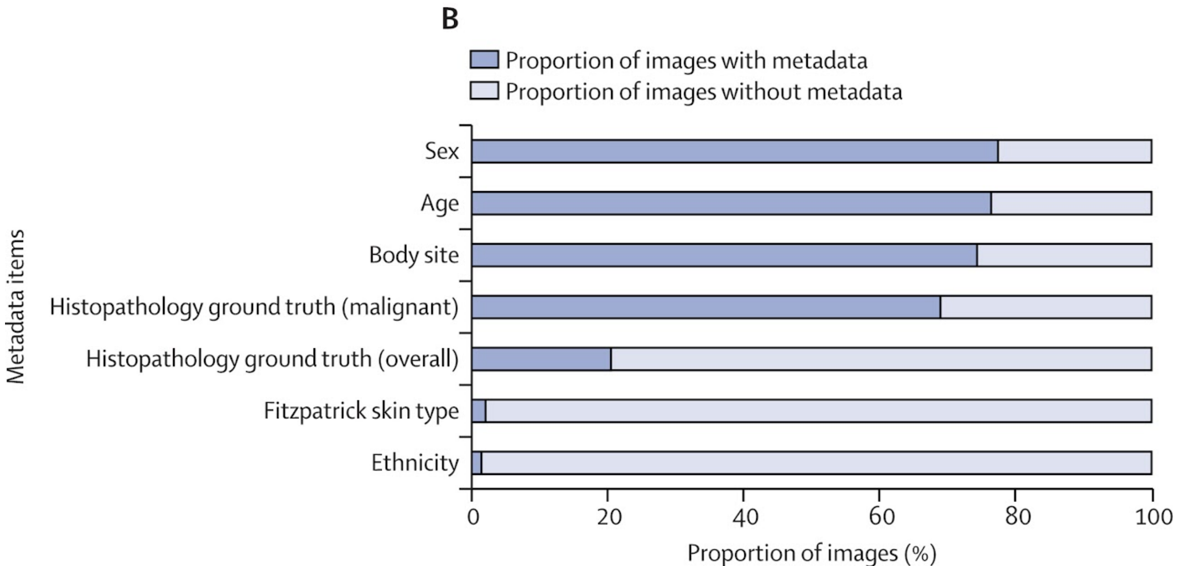
Social

The role of communities that develop, use and are affected by data science and digital tools

- Datasets and digital tools need to be designed for – and with – diverse populations, so they don't only work for some groups at the expense of others
- People should have sufficient levels of visibility and control over data about their health
- Research cultures need to value people working on data science and research software

A Structural Problem Starting with the Data we Use





	GDPPR modal	GDPPR recency	ECIA	HES modal	HES recency
White British	98	97	97	97	97
Bangladeshi	96	96	95	93	92
Pakistani	91	91	90	87	87
Chinese	92	92	89	83	82
Indian	88	87	86	83	82
Black African	82	81	80	77	75
Black Caribbean	81	79	76	74	73
Arab	70	68	66	No data	No data
White Irish	70	64	64	65	62
Mixed White and Black Caribbean	60	58	62	66	64
Other White	54	48	44	52	49
Mixed White and Asian	52	49	50	48	47
Other Asian	59	54	49	39	37
Mixed White and Black African	33	29	34	40	38
Other Mixed	20	18	15	14	12
Other Black	15	14	14	13	12
Any Other Ethnic Group	14	13	8	6	5
Traveller	6	5	5	No data	No data

Requiring Regulatory, Academic & Private Sector Collaboration

Agency for Healthcare Research and Quality

Correspondence

https://doi.org/10.1038/s41591-022-01987-w

Tackling bias in AI health datasets through the STANDING Together initiative

Check for updates

To the Editor – As of June 2022, a wide range of Artificial Intelligence (AI) as a Medical Device (AIaMDs) have received regulatory clearance internationally, with at least 343 devices cleared by the US Food and Drug Administration (FDA)[1]. Despite the enormous potential of AIaMDs, their rapid growth in healthcare has been accompanied by concerns that AI models may learn biases engrained in medical practice and exacerbate health inequalities. This has been exemplified by several AI systems that have shown the ability of algorithms to systematically misrepresent and exacerbate health problems in minority groups[2,3]. This raises concerns that, without appropriate safeguarding, AI models prioritize sample size. There are concerns that many health datasets do not adequately represent minority groups; however, the extent of this problem is unknown because many datasets do not provide demographic information, such as on ethnicity and race. Publicly available datasets for skin cancer and eye imaging have shown inconsistent and incomplete demographic reporting, and are disproportionately collected from a small number of high-income countries[6,7]. For skin cancer datasets, the reporting of key demographic information, such as ethnicity and skin tone, even when clinically relevant, was only present in 2% of datasets[7].

Under-representation in datasets can affect the fairness of AI systems by two observations and labels were constructed. These concerns have motivated calls for better documentation practices and the creation of tools such as 'Datasheets for Datasets' and 'Healthsheets'[8,9].

The aforementioned problems are becoming increasingly recognized by regulators of medical devices. In October 2021, The US FDA, Health Canada and the UK Medicines and Healthcare products Regulatory Agency (MHRA) jointly published 10 guiding principles for good machine learning practice. This specifically states that data should be representative of the intended population in order to "manage any bias, promote appropriate and generalizable performance across the intended patient population, assess usability

Equity in medical devices independent review: terms of reference

April 2022

Purpose

The purpose of the review is to establish the extent and impact of potential ethnic and other unfair biases in the design and use of medical devices and to make recommendations for more equitable solutions.

6.3.1 Diversity and inclusion in development

6.3.1.1 The supplier shall document how they ensure they and their suppliers accommodate the diversity of users impacted by the product. If this is not possible, the supplier shall provide a justification as to why.

NOTE 1 Assessment of potential biased processes and impacts could be limited when development teams are homogenous, i.e. non-diverse. It is assumed that the diversity of the team reflects current good practice for inclusion criteria.

6.3.1.2 Suppliers shall undertake and document a process to ensure that the users identified and engaged in **5.1** are representative of the intended users

NOTE 2 Within the intended groups of users stakeholders should be diverse (e.g. across ethnicity, gender, socioeconomic status, age, and geography). For example, for a product which treats conditions found in older populations, engagement might not require stakeholders with diverse ages. But it is still necessary to engage stakeholders across other social categories (e.g. different genders and ethnicities) within older populations to meet the requirements for b). Suppliers might need to consider and account for the range of digital skills and engagement among different groups of people when seeking a wide range of participants.

6.3.1.3 Suppliers shall document both its justification for the chosen approach and any specific framework of methodology followed and the outcomes.

6.3.2 Assessment of model bias risks

6.3.2.1 Developers shall undertake a risk assessment to identify risks of bias in model development and deployment that might result in inequitable outcomes.

NOTE Possible approaches to complete the algorithmic risk and impact assessment include:

a) an extension of standard project risk management processes;

b) Assessment List for Trustworthy AI [12] (specifies a Human Rights Impact Assessment);

c) the Ada Lovelace Algorithmic Impact Assessment (AIA) [13] or the Black Box report [14] (which divides the audit for bias and assessment of impact);

d) the Canadian government AIA [15] (outlines an additional approach for a quantitative method); and

e) the NHS England Equality and Health Inequality Assessment (EHIA) form [16].

6.3.2.2 The risk assessment (see **6.3.2.1**) shall document the degree of potential impact and harm for sub-groups and outline mitigations.

6.3.2.3 The risk assessment shall include:

a) explanations for the mechanism referred to in **8.1.4** for monitoring of real-world impact on sub-groups post-deployment**;** and

b) details of when the assessment is to be reviewed, e.g. at key stages of model maintenance, new use-cases, updates, and decommissioning.

And Careful Consideration of Technical Solutions

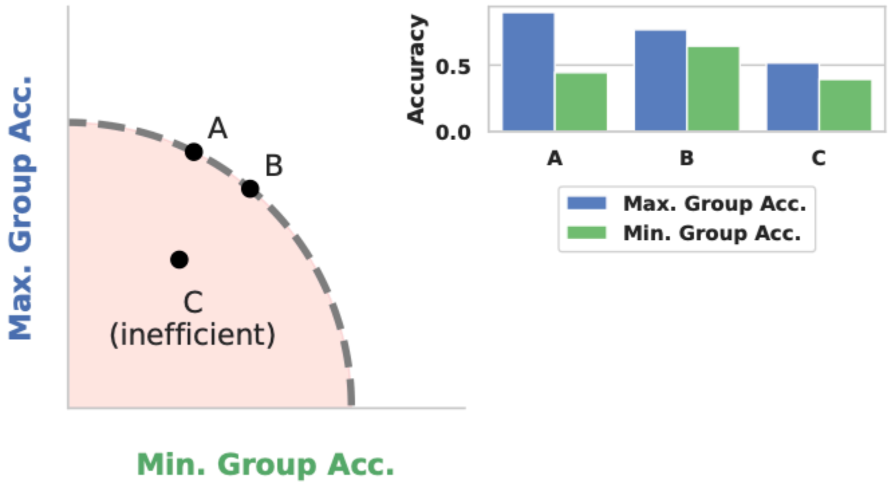


Figure 1. **Pareto curve.** We depict the typical trade-off assumed by most fairness studies in computer vision. By adjusting fairness, the accuracies of a classifier on its best- and worst-performing groups form a *Pareto curve* (dotted gray line). Points A and B are maximally efficient configurations because they lie on the curve; B is fairer since it yields a lower accuracy difference between the groups (see bar plot on right). Point C is as fair as B, but is inefficient because it reduces the accuracies of both groups. In our experiments, we find that accuracy-based fairness constraints applied to deep neural networks tend to achieve inefficient configurations like C.

White Test Patients				
Privacy Level	Average White Influence	Average Black Influence	Most Helpful Ethnicity	Most Harmful Ethnicity
None	0.29 ± 2.40	0.71 ± 1.40	White	White
Low	−0.22 ± 0.70	−0.03 ± 0.17	White	White
High	−0.11 ± 3.94	0.03 ± 1.35	White	White
Black Test Patients				
Privacy Level	Average White Influence	Average Black Influence	Most Helpful Ethnicity	Most Harmful Ethnicity
None	0.48 ± 1.39	0.44 ± 2.19	**Black**	White
Low	−0.23 ± 0.75	−0.03 ± 0.18	White	White
High	−0.40 ± 4.10	0.12 ± 1.45	**White**	White

Table 4: Group influence summary statistics across all privacy levels for white (majority) and Black (minority) training patients on both white and Black test patients in MIMIC-III. Privacy changes the most helpful group from Black patients to the majority white patients and minimizes their helpful influence. This needs careful consideration as the use of ethnicity is still being investigated in medical practice.

Impact of Healthcare Algorithms on Racial and Ethnic Disparities in Health and Healthcare

Consensus Panel Discussion/Q & A

March 3, 2023

1:07– 1:40 p.m. ET

Discussion Questions

- What's missing: gaps in experience and insights related to algorithmic standards/stewardship, including pitfalls?
- Guidance needed: which standards are needed and possible and by whom? How to develop standards and gain adoption?

Impact of Healthcare Algorithms on Racial and Ethnic Disparities in Health and Healthcare

Closing Remarks

Christine Chang, MD

Associate Director, EPC Division, AHRQ

Racial Bias and Healthcare Algorithms

March 3, 2023

1:40-1:45 p.m. ET

National Institute on Minority Health and Health Disparities

What's Next?

- Panel deliberations
- Webinar presenting panel recommendations: May 15, 2023
- For inquiries and to be added to the distribution list
  AIDisparities@norc.org

Thank you!

National Institute
on Minority Health
and Health Disparities