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

March 2, 2023
10:30 a.m.- 3:40 p.m. ET
Welcome

Prashila Dullabh, MD
NORC

Racial Bias and Healthcare Algorithms
March 2, 2023
10:30–10:35 a.m. ET
Zoom Housekeeping

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

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

Today’s webinar is being recorded for notetaking purposes.

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
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.
Welcome Remarks

Dr. Robert Otto Valdez, PhD, MHSA was appointed Director of AHRQ in February 2022. He was previously the Robert Wood Johnson Foundation (RWJF) Professor Emeritus of Family & Community Medicine and Economics at the University of New Mexico (UNM).

Dr. Eliseo Perez-Stable, MD is Director of the National Institute on Minority Health and Health Disparities (NIMHD) at the National Institutes of Health (NIH). He oversees NIMHD’s annual budget to advance the science of minority health and health disparities research.

Dr. RDML Felicia Collins, MD, MPH, FAAP is the Deputy Assistant Secretary for Minority Health. As the Director of the Office of Minority Health (OMH), she leads the office in its mission to improve the health of racial and ethnic minority populations through the development of health policies and programs that help eliminate health disparities.
Racial Bias and Healthcare Algorithms
March 2, 2023
10:40 – 10:45 A.M. ET
Populations with Health Disparities

- Racial and ethnic minority populations in census
- Less privileged socio-economic status
- Underserved rural residents
- Sexual and gender minorities
- Social disadvantage that results in part from being subject to discrimination or racism, and being underserved in health care
- A health outcome that is worse in these populations compared to a reference population group defines a health disparity
## National Institute on Minority Health and Health Disparities Research Framework

### Levels of Influence*

<table>
<thead>
<tr>
<th>Domains of Influence (Over the Lifecycle)</th>
<th>Individual</th>
<th>Interpersonal</th>
<th>Community</th>
<th>Societal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biological</td>
<td>Biological Vulnerability and Mechanisms</td>
<td>Caregiver-Child Interaction Family Microbiome</td>
<td>Community Illness Exposure Herd Immunity</td>
<td>Sanitation Immunization Pathogen Exposure</td>
</tr>
<tr>
<td>Behavioral</td>
<td>Health Behaviors Coping Strategies</td>
<td>Family Functioning School/Work Functioning</td>
<td>Community Functioning</td>
<td>Policies and Laws</td>
</tr>
<tr>
<td>Physical/Built Environment</td>
<td>Personal Environment</td>
<td>Household Environment School/Work Environment</td>
<td>Community Environment Community Resources</td>
<td>Societal Structure</td>
</tr>
<tr>
<td>Sociocultural Environment</td>
<td>Sociodemographics Limited English Cultural Identity Response to Discrimination</td>
<td>Social Networks Family/Peer Norms Intellectual Discrimination</td>
<td>Community Norms Local Structural Discrimination</td>
<td>Social Norms Societal Structural Discrimination</td>
</tr>
<tr>
<td>Health Care System</td>
<td>Insurance Coverage Health Literacy Treatment Preferences</td>
<td>Patient-Clinician Relationship Medical Decision-Making</td>
<td>Availability of Services Safety Net Services</td>
<td>Quality of Care Health Care Policies</td>
</tr>
<tr>
<td>Health Outcomes</td>
<td>Individual Health</td>
<td>Family/ Organizational Health</td>
<td>Community Health</td>
<td>Population Health</td>
</tr>
</tbody>
</table>

*Health Disparity Populations: Race/Ethnicity, Low SES, Rural, Sexual/Gender Minority
Other Fundamental Characteristics: Sex/Gender, Disability, Geographic Region

National Institute on Minority Health and Health Disparities. 2019
AI/Algorithm Applications

“For the first time in history, we have technology (AI) that is Opening our eyes to who we are, is changing us as we speak, and could allow us to play a conscious role in who we want to become.”

Jennifer Aue
IBM Director for AI Transformation
AI professor at the University of Texas

Who We Are: Human Biases exist in AI/Algorithm Applications

NIH National Institute on Minority Health and Health Disparities

Look Deeper with More Eyes
“For the first time in history, we have technology (AI) that is Opening our eyes to who we are, is changing us as we speak, and could allow us to play a conscious role in who we want to become.”

**Who We Want to Become:** Ethical AI/Algorithms

**Use Models in Context**

- Epic’s overhaul of a flawed algorithm shows why AI oversight is a life-or-death issue
- The Problem
  - An Epic EHR system developer trained a machine learning model to predict sepsis using a certain population’s data
  - When the model was reused with a new population, the performance was substantially worse than the original results suggested

**Ensure R’s in AI/Algorithms**

- **Repeatability** – same result with same data
- **Replicability** – someone else achieves same result with same data
- **Reproducibility** – same result with different data (generalizable)

**Models in Context – Know Populations**

National Institute on Minority Health and Health Disparities

**Develop AI/Algorithms with Health Equity to Prevent Health Disparities**

Jennifer Aue
IBM Director for AI Transformation
AI professor at the University of Texas

Look Deeper with More Eyes
NIMHD Goals in Data Science and Cloud Computing

• Increase **workforce of underrepresented women and populations with health disparities** in data science and cloud computing
• Utilize social determinants of health and population science **big data** in research to understand and improve health outcomes and reduce disparities
• Develop **ethical AI** utilizing bias mitigation strategies across the continuum of design, data selection, algorithm development and training, and implementation to ensure health equity
Cloud-based **social determinants of health and population science data** platform designed to accelerate research in health disparities and health outcomes, and to develop AI bias mitigation strategies.

Register for ScHAe:  [https://www.nimhd.nih.gov/resources/schare/](https://www.nimhd.nih.gov/resources/schare/)
Social Determinants of Health Measures

• PhenX Toolkit on SDOH measures: https://www.phenxtoolkit.org/collections/view/6
• Demographics including family background
• Urban or rural residence or geographic region
• Cultural identity, religiosity, spirituality
• Language proficiency, Literacy, numeracy
• Structural determinants: housing, green space, broadband, economic opportunity, transportation, schools, healthy food access, public safety, political
HHS-OMH Remarks

RDML Felicia Collins, MD, MPH, FAAP
Deputy Assistant Secretary for Minority Health
Director, Office of Minority Health
U.S. Department of Health and Human Services

Racial Bias and Healthcare Algorithms
March 2, 2023
10:45–10:50 a.m. ET
Introduction and Purpose of the Meeting

Anjali Jain, MD
AHRQ

Racial Bias and Healthcare Algorithms
March 2, 2023
10:50 a.m.- 11:10 a.m. ET
"Of all the forms of inequality, injustice in healthcare is the most shocking and inhumane."

DR. MARTIN LUTHER KING, JR.
Background

• AHRQ received a request from Congress to review the evidence on the potential of algorithms to contribute to disparities in health care for racial and ethnic minorities

• In response, AHRQ:
  o issued a request for information (RFI) in the federal register
  o commissioned an evidence review with the aim of informing guidance to mitigate bias in healthcare algorithms
RFI questions were intended to:

- Identify algorithms in use with potential for racial/ethnic bias
- Discover existing approaches to identifying or mitigating bias in algorithms
- Characterize awareness of algorithms and bias among patients, providers, and others
- Identify standards for algorithm development, validation, and updating
Responses to the RFI

• **42 respondents**
  – 485 pages of responses

• **Respondents included**
  – 18 clinical and professional associations
  – 9 groups focused on health technology, including algorithm developers
  – 7 universities
  – 4 federal and state agencies (non-AHRQ)
  – 1 payer
  – 4 individuals
Insights from the RFI

• Responses analyzed using qualitative analysis
• Respondents named 18 algorithms with potential for bias
• Major themes from responses included:
  o Addressing racial bias in healthcare algorithms is urgent and important
  o Algorithms are in widespread use and have a potentially large impact
  o Bias and disparities can result from algorithms whether or not they explicitly include race
  o Great heterogeneity and lack of standardization in how race and social determinants of health data are collected and defined
  o Bias can be introduced at all stages of algorithm development and implementation
  o Organizations making efforts to assess bias related to algorithms and improve inequities
  o Clinicians and patients often unaware of algorithm use and potential for bias
  o Algorithms should be discussed as part of shared decision making between patient and provider

• Evidence review awarded to ECRI/Penn May 2021

• Review conducted since May 2021, includes input from experts and stakeholders as key informants & technical experts

• Draft report posted for public comment, February 9, 2023
  o Comments on the report can be submitted until 11:59 p.m. ET on March 9, 2023 at the link below

  https://effectivehealthcare.ahrq.gov/products/form/racial-disparities-health-healthcare-draft-comments

• 2 Key questions and 4 contextual questions

  o **Key Question 1.** What is the effect of healthcare algorithms on racial and ethnic differences in access to care, quality of care, and health outcomes?

  o **Key Question 2.** What is the effect of interventions or approaches to mitigate racial and ethnic bias in the development, validation, dissemination, and implementation of healthcare algorithms?
Impact of Healthcare Algorithms on Racial and Ethnic Disparities in Health and Healthcare

- 4 contextual questions to explore practical aspects of algorithmic use and bias, addressed through supplemental literature reviews and conversations with experts and key stakeholders
  - **CQ 1** examines the problem’s scope within healthcare.
  - **CQ 2** describes recently emerging standards and guidance on how racial and ethnic bias can be prevented or mitigated during algorithm development and deployment.
  - **CQ 3** explores stakeholder awareness and perspectives about the interaction of algorithms and racial and ethnic disparities in health and healthcare.
  - **CQ 4** involved an in-depth analysis of a sample of six algorithms to better understand how their design and implementation might contribute to disparities.
“Not everything that is faced can be changed, but nothing can be changed until it is faced.”

~James Baldwin
Introduction of Keynote Speaker

Arlene Bierman, MD, MS
AHRQ

Racial Bias and Healthcare Algorithms
March 2, 2023
11:10 a.m.- 11:15 a.m. ET
Dissecting Racial Bias

Ziad Obermeyer, MD
UC Berkeley
Blue Cross of California Distinguished Associate Professor of Health Policy and Management

Racial Bias and Healthcare Algorithms
March 2, 2023
11:15 a.m.- 11:45 a.m. ET
Many great uses of algorithms in health
  - Risk prediction: What will happen
  - Diagnosis: Likelihood that patient has a disease
  - ...

Many worries about disparities in these algorithms

depression, or opioid misuse; and warfarin dosing. We found evidence that algorithms can: a) reduce disparities (i.e., revised Kidney Allocation System, prostate cancer screening tools), b) perpetuate or exacerbate disparities (e.g., estimated glomerular filtration rate [eGFR] for kidney function measurement, cardiovascular disease risk assessments), and/or c) have no effect on racial or ethnic disparities (e.g., HEART Pathway). Further algorithms that perpetuated or

What makes the difference?
Biased vs. unbiased algorithms

• A common concern: **Race as a predictor**
  ► a big problem if “hard-coded,” e.g., assumptions about Black lung capacity

• Today: A different concern—and a way to debias algorithms

Figure 2. Conceptual Model for Understanding Racial and Ethnic Biases Introduced During Algorithm/Clinical Decision-Making Tool Development, Translation, Dissemination, and Implementation

(a) Algorithm Development Phase

<table>
<thead>
<tr>
<th>Algorithm development steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem Formulation</td>
</tr>
<tr>
<td>Data Selection and Management</td>
</tr>
<tr>
<td>Model Training / Development</td>
</tr>
<tr>
<td>Validation / Performance Evaluation</td>
</tr>
</tbody>
</table>
Example 1: Targeting extra help for complex patients

• Complex, chronically ill patients have high costs, poor care
  ▶ Innovation: ‘high-risk care management’
  ▶ But expensive – so targeting critical

• Algorithms are used everywhere for this
  ▶ Specific software we study: 70 million patients/year (US)
  ▶ Market estimates: 150-200 million patients/year (US)

• Common goal: Find patients who are going to get sick
  ▶ As measured by future health care costs
  ▶ So we can target help now
We studied ‘racial bias’

- Principle: Same score → Treated the same
  - Should have same needs

- Color of their skin should not matter

- But it does
  - Black patients have worse realized health
  - At every algorithm score

Obermeyer, Powers, Vogeli, Mullainathan, Science 2019
Dissecting the bias

- We’d like to understand where the algorithm is going wrong
- One clue: where it is going right
- Algorithm predicts total health costs well for Black and White patients

Obermeyer, Powers, Vogeli, Mullainathan, Science 2019
• Algorithm is accurately predicting cost

• Black patients have lower costs at the same health status
  1. White patients have better access to health care
  2. The health system treats Black patients differently

• Result: biased health prediction
  ▶ With or without race adjustment
  ▶ In this case: No race adjustment
Finding better targets for prediction

• Insight: We have other proxy variables besides cost
  ▶ Total cost vs. avoidable cost vs. health outcomes

• We worked with developer to re-train algorithm on health
  ▶ Huge benefits for equity: 84% less bias
  ▶ Better fit with business purpose

• Suggests finding better proxies is a high-value activity
  ▶ Practical: Same dataset, same pipeline, different label

Obermeyer, Powers, Vogeli, Mullainathan, Science 2019
Our ‘playbook’—inspired by work over past 2 years

• Bad news: We found bias almost everywhere we looked
  ► Population health resource allocation
  ► Clinical disease prediction
  ► Operational decisions

• Good news: Almost all fixable
  ► By retraining on less biased label

DOWNLOAD THE PLAYBOOK
Example 2: Pain is concentrated in most disadvantaged

- But story isn’t as simple as it looks

- Typical exercise in literature, e.g., for knee osteoarthritis:
  - Two patients, similar x-rays
  - Compare pain scores

- Black, lower-income, lower-education: still have more pain
  - At every level of x-ray graded disease severity
Some explanations from the literature

• If it’s not in their knees…

• Maybe it’s in their heads?
  ► Stress makes similar stimuli more painful
  ► Psychosomatic factors
  ► Coping skills

• Or in the medical system
  ► Access to therapies
Concrete clinical scenario

- Implication of literature
  - Black patients’ pain not reflected in disease severity

- Leads to allocation of non-knee-based treatments

- But what do we mean by ‘disease severity’?
  - How do we measure it?

Pain → X-ray → Looks normal → ‘Not in the knee’
Current SOTA
Measuring osteoarthritis severity

- Objective grading scales, based on x-ray appearance
- Most common: Kellgren-Lawrence, 1957 (KLG)

- Original studies on coal miners in Lancashire, England
  - No mention of subjects’ race, sex
A good job for an algorithm?

- Human radiologists may overlook causes of pain in disadvantaged groups
- We’d like an algorithm to help—but…
  - Typical approach: train to match human performance
- Exactly what we don’t want to do!
Finding a better target for prediction

Learn from the radiologist

Kellgren-Lawrence = 2/4

Listen to the patient

Pain = 9/10
Finding the data: Not straightforward

• Easy to find: x-rays + radiologist interpretation
  ► Sitting on every hospital’s PACS system

• Much harder to find: x-rays + patient pain experience

• But once we have data: a very straightforward ML problem

  ► If pain is predictable from knee image
  – …Pain is in the knee (not in the head, coping, …)
Algorithm closes nearly half the pain gap

- Adjusting for standard severity measure: –9%

- Adjusting for algorithmic severity measure: –43%
  - 4.7x more than standard measure
  - 95% CI: 3.2-11.8

- Similar results for
  - Income: 2.0x
  - Education: 3.6x

Koos score: Max 100, severe ≤86

The stakes are high

- Take patients with severe pain
- Simulate swapping in algorithm severity, not radiologist
- Double fraction of Black knees eligible for surgery
Summary

• Algorithmic bias is often **decided early**
  ▶ How we **ask the question** for algorithms to answer
  ▶ Not how the algorithm **answers the question**

• **Suggests problem formulation** is a critical area
  ▶ This is understudied
  ▶ Because it’s difficult: What are we trying to do?
Consensus Panel Co-Chair Comments

Marshall Chin, MD, MPH
University of Chicago

Racial Bias and Healthcare Algorithms
March 2, 2023
11:45 a.m.- 11:50 a.m. ET
Consensus Panel for Racial Bias and Healthcare Algorithms

- **Consensus Panel Composition:**
  - 2 co-chairs
  - 7 panelists
  - Diverse perspectives represented

- **Consensus Panel Role: Identify and formulate:**
  - Guiding principles for racial/ethnic bias prevention, identification, and mitigation
  - Potential solutions, approaches and resources to address such bias
  - Actionable next steps for stakeholders

- **Panel will present findings at a virtual public meeting on May 15, 2023**
Impact of Healthcare Algorithms on Racial and Ethnic Disparities in Health and Healthcare

Break
March 2, 2023
11:50 a.m.– 12:00 p.m. ET

Please take ten minutes for a break
Evidence Review
Methods, Key Question 1, and Contextual Question 1

Kelley Tipton, MPH
ECRI

Shazia M. Siddique, MD, MSHP
University of Pennsylvania School of Medicine

Racial Bias and Healthcare Algorithms
March 2, 2023
12:00-12:20 p.m. ET
Use of Race and Ethnicity in Healthcare Algorithms

Key Questions (KQs)

• KQ 1: What is the effect of healthcare algorithms on racial and ethnic differences in access to care, quality of care, and health outcomes?

• KQ 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?
Contextual Questions (CQs)

• CQ 1: How widespread is the inclusion of input variables based on race and ethnicity in healthcare algorithms?

• CQ 2: What are existing and 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 3: To what extent are patients, providers (e.g., clinicians, hospitals, health systems), payers (e.g., insurers, employers), and policymakers (e.g., healthcare and insurance regulators, state Medicaid directors) aware of the inclusion of input variables based on race and ethnicity in healthcare algorithms?

• CQ 4: Select a sample of approximately 5-10 healthcare algorithms that have the potential to impact racial and ethnic disparities in access to care, quality of care, or health outcomes and are not included in KQs 1 or 2. For each algorithm, describe the type of algorithm, its purpose (e.g., screening, risk prediction, diagnosis, etc.), its developer and intended end-users, affected patient population, clinical condition or process of care, healthcare setting, and information on outcomes, if available.
Conceptual Model for Understanding Racial and Ethnic Biases Introduced During Algorithm/Clinical Decision-Making Tool Development, Translation, Dissemination, and Implementation

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algorithm</td>
<td>A mathematical formula or model that combines different input variables or factors to inform a calculation or an estimate, such as an estimate of disease or risk of a particular health outcome.</td>
</tr>
<tr>
<td>Algorithmic bias</td>
<td>Differential performance of an algorithm in different groups (such as racial or ethnic groups) due to intrinsic attributes of the algorithm.</td>
</tr>
<tr>
<td>Risk of bias (ROB)</td>
<td>The likelihood that a study’s reported results are misleading due to methodologic issues in study design.</td>
</tr>
</tbody>
</table>
Analytic Framework/PICOTS for KQs

Population
- Patients whose healthcare could be affected by algorithms

Intervention and Comparators

KQ 1
- **Intervention:**
  - Algorithms that have been, or are currently being, used for screening, risk prediction, diagnosis, prognosis, treatment, or resource allocation
- **Comparators:**
  - No algorithm
  - Same algorithm with or without race or ethnicity
  - Same algorithm with or without other input variables that may contribute to bias
  - Different algorithm designed for the same purpose
  - No comparator

KQ 2
- **Intervention:**
  - Interventions, models of interventions, or approaches to mitigate bias associated with use of algorithms
- **Comparators:**
  - Original algorithm, dataset, or approach
  - Alternative mitigation strategies

Outcomes

Access to Care
- Patient use of or eligibility for healthcare services
- Patient use of population health services
- Direct costs to patients

Quality of Care
- Appropriateness of diagnosis, treatment, and/or monitoring (e.g., diagnostic/prognostic accuracy)
- Timeliness of care
- Patient experience/satisfaction
- Hospital readmission
- Hospital length of stay

Health Outcomes
- Mortality/survival
- Morbidity
- Quality of life
- Functional status

Settings

**Hospital:** Inpatient, emergency department, observation unit

**Ambulatory:** Post-acute care, primary, specialty, rehabilitation care sites, long-term care

**Non-clinical site:** Home care (telemedicine, self-care)
Overview of Project Methods

• Systematic literature search of Embase, MEDLINE, PubMed, Cochrane Library, and grey literature (1/1/2011 to 1/12/2022)
  ▶ Updated search performed through 2/7/2023

• Used predefined criteria and dual review to screen all records for KQ 1 and KQ 2; selected eligible full-length research studies published in English for one or both KQs

• Assessed studies’ methodologic ROB using ROBINS-I and piloted an appraisal supplement to assess racial and ethnic equity-related ROB

• Completed a narrative synthesis, catalogued study characteristics and outcome data

• CQs addressed through supplemental searches, review of RFI responses, and discussions with SMEs, TEP, and KIs

• External peer review completed; report posted for public comment on 2/9/2023
Study Flow Diagram

Articles identified through database literature searches (n=8,500)

Abstracts screened (n=4,462)

Full-text articles assessed (n=260)

Articles excluded at title level (n=4,038)

Articles excluded at abstract level (n=4,202)

Articles excluded at full-text level (n=218)

Key Question 1 (n=218):
- Does not examine a clinical algorithm or algorithm-based tool (n=90)
- Does not examine the effect of a clinical algorithm or algorithm-based tool on racial or ethnic differences (n=37)
- Does not report an outcome of interest (n=34)
- Does not report outcomes by race or ethnicity (n=4)
- Derivation study with only internal validation (n=6)
- Non-US population (n=26)
- Not a full-length primary study (n=21)

Key Question 2 (n=218):
- Does not examine the effect of an intervention to mitigate bias of a clinical algorithm or algorithm-based tool (n=171)
- Does not report an outcome of interest (n=21)
- Does not report outcomes by race or ethnicity (n=5)
- Not a full-length primary study (n=21)

42 studies included*

12 for Key Question 1
33 for Key Question 2

*3 studies were included for both
Classification of Studies by Key Question

• KQ 1: included studies evaluated an algorithm’s effect on health or healthcare outcomes stratified by racial and ethnic groups

• KQ 2: included studies intended to develop an intervention or strategy to mitigate
  ▶ racial and ethnic algorithmic bias OR
  ▶ a known racial and ethnic disparity associated with an algorithm

• Studies included in both KQ 1 and 2 described
  ▶ a racial and ethnic disparity associated with an algorithm, AND
  ▶ an intervention on the algorithm to mitigate the disparity
KQ 1 Results: Overview

• 12 included studies
  ► Algorithms reduce disparities (n=4 studies)
  ► Algorithm with no effect on disparities (n=1 study)
  ► Algorithms that perpetuate or exacerbate disparities (n=7 studies)
• KQ 2 with further evidence of algorithms that perpetuate or exacerbate disparities, thereby warranting mitigation strategies
• Studies were appraised at moderate-to-high risk of bias
### KQ 1: Algorithms Shown to Reduce Disparities

<table>
<thead>
<tr>
<th>Clinical Assessment</th>
<th>Number of Studies</th>
<th>Algorithm(s)</th>
<th>Comparator</th>
<th>Includes race or ethnicity? (Y/N)</th>
<th>Primary outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kidney Transplant Suitability</td>
<td>1 [Zhang 2018]</td>
<td>Kidney Allocation System (KAS)</td>
<td>Pre-implementation of KAS</td>
<td>Y</td>
<td>Waitlisting rate</td>
</tr>
<tr>
<td>Lung Transplant Suitability</td>
<td>1 [Wille 2013]</td>
<td>Lung Allocation Score (LAS)</td>
<td>Pre-implementation of LAS</td>
<td>N</td>
<td>Death while on waitlist or ineligibility due to morbidity while on waitlist</td>
</tr>
<tr>
<td>Prostate Cancer Risk Risk</td>
<td>2 [Presti 2021]</td>
<td>KPPC RC</td>
<td>Compared KPPC RC models</td>
<td>Y</td>
<td>Biopsies avoided and clinically significant prostate cancers missed</td>
</tr>
<tr>
<td></td>
<td>[Carbanaru 2019]</td>
<td>PCPT</td>
<td>PBCG</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

KPPC RC=Kaiser Permanente prostate cancer risk calculator; N=no; PCPT=Prostate Cancer Prevention Trial algorithm; PBCG=Prostate Biopsy Collaborative Group algorithm; Y=yes

**Takeaway:** Existing disparities were identified prior to algorithm development and implementation. These algorithms were implemented as part of an intentional effort to tackle disparities.
### KQ 1: Algorithms with No Effect on Disparities

<table>
<thead>
<tr>
<th>Clinical Assessment</th>
<th>Number of Studies</th>
<th>Algorithm</th>
<th>Comparator</th>
<th>Includes race or ethnicity? (Y/N)</th>
<th>Primary outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emergency Department Triage</td>
<td>1 [Snively 2021]</td>
<td>HEART Pathway</td>
<td>Pre-implementation of HEART Pathway</td>
<td>N</td>
<td>30-day death or myocardial infarction</td>
</tr>
</tbody>
</table>

**Takeaway:** The HEART Pathway did not significantly impact death or MI rates for BIPOC individuals. However, non-white patients and women were more likely to be classified as low risk and discharged early. Longer term implications have not been assessed.

Of note, non-adherence to the pathway was higher for women, but non-significant for non-White individuals, providing insight on pragmatic challenges of algorithm implementation.
### KQ 1: Algorithms Shown to Perpetuate Disparities

<table>
<thead>
<tr>
<th>Clinical Assessment</th>
<th>Number of Studies</th>
<th>Algorithm</th>
<th>Comparator</th>
<th>Includes race or ethnicity? (Y/N)</th>
<th>Primary Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severity of Illness Scores Applied to Crisis Standards of Care</td>
<td>3</td>
<td>SOFA and LAPS2</td>
<td>Compared models/tiering systems</td>
<td>N</td>
<td>In-hospital mortality</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SOFA, OASIS, APACHE IVa</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SOFA tiering systems</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Ashana 2021]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Sarkar 2021]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Miller 2021]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

APACHE IVa=Acute physiology and chronic health evaluation; LAPS2=Laboratory-based Acute Physiology Score version 2; N=no; NR=not reported; OASIS=Oxford Acute Severity of Illness Score; SOFA=Sequential Organ Failure Assessment

**Takeaway:** Applying severity of illness scores outside of its original intended application (e.g. Crisis Standards of Care for the COVID-19 pandemic) results in less resources for BIPOC (Black and Hispanic) individuals, thereby leading to disparities.
### KQ 1: Algorithms Shown to Perpetuate Disparities

<table>
<thead>
<tr>
<th>Clinical Assessment</th>
<th>Number of Studies</th>
<th>Algorithm</th>
<th>Comparator</th>
<th>Includes race or ethnicity? (Y/N)</th>
<th>Primary Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severity of Illness Scores Applied to Crisis Standards of Care</td>
<td>3</td>
<td>SOFA and LAPS2</td>
<td>Compared models/tiering systems</td>
<td>N</td>
<td>In-hospital mortality</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SOFA, OASIS, APACHE IVa</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SOFA tiering systems</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>[Ashana 2021]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>[Sarkar 2021]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>[Miller 2021]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lung Cancer Risk</td>
<td>2</td>
<td>USPSTF-2013</td>
<td>PLCOm2012</td>
<td>N (Y for comparator)</td>
<td>Lung cancer screening eligibility</td>
</tr>
<tr>
<td></td>
<td></td>
<td>[Pasquenelli 2021]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>[Han 2020]</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Takeaway:** Both studies found that USPSTF-2013 resulted in higher proportions of Black patients being ineligible for lung cancer screening. However, this is not a pre-post study. Downsides of potential over-screening were not assessed.
## KQ 1: Algorithms Shown to Perpetuate Disparities

<table>
<thead>
<tr>
<th>Clinical Assessment</th>
<th>Number of Studies</th>
<th>Algorithm</th>
<th>Comparator</th>
<th>Includes race or ethnicity? (Y/N)</th>
<th>Primary Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severity of Illness Scores Applied to Crisis Standards of Care</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>In-hospital mortality</td>
</tr>
<tr>
<td></td>
<td>[Ashana 2021]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[Sarkar 2021]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[Miller 2021]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lung Cancer Risk</td>
<td>2</td>
<td>SOFA and LAPS2</td>
<td></td>
<td></td>
<td>Lung cancer screening eligibility</td>
</tr>
<tr>
<td></td>
<td>[Pasquenelli 2021]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>[Han 2020]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Opioid Misuse Risk</td>
<td>1</td>
<td>Natural language processing classifier</td>
<td>None</td>
<td>NR</td>
<td>Referral for education, treatment options, and care pathways</td>
</tr>
<tr>
<td></td>
<td>[Thompson 2021]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High-Risk Care Management</td>
<td>1</td>
<td>Commercial risk prediction calculator</td>
<td>None</td>
<td>N</td>
<td>Eligibility for a care management program</td>
</tr>
<tr>
<td></td>
<td>[Obermeyer 2019]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Takeaway:** Algorithms that do not include race can lead to disparities: Obermeyer studied an algorithm which predicted healthcare costs, as a proxy for healthcare needs. This is flawed because the association between costs and health differs across racial and ethnic groups.
Further evidence from KQ 2: Algorithms Perpetuate Disparities

<table>
<thead>
<tr>
<th>Clinical Category</th>
<th>Algorithm</th>
<th>Key Question</th>
<th>Study</th>
<th>Study Design(^d)</th>
<th>Disparities in Health outcome(^b)</th>
<th>Disparities in Access(^b)</th>
<th>Disparities in Quality(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Kidney function measurement</strong></td>
<td>eGFR(^c)</td>
<td>KQ 2</td>
<td>Ahmed 2021(^2)</td>
<td>Modelling(^d)</td>
<td>*</td>
<td>↑</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>eGFR(^c)</td>
<td>KQ 2</td>
<td>Inker 2021(^2)</td>
<td>Modelling(^d)</td>
<td>*</td>
<td>*</td>
<td>↑</td>
</tr>
<tr>
<td></td>
<td>eGFR(^c)</td>
<td>KQ 2</td>
<td>Casal 2021(^4)</td>
<td>Modelling(^d)</td>
<td>*</td>
<td>↑</td>
<td>↑</td>
</tr>
<tr>
<td></td>
<td>eGFR(^c)</td>
<td>KQ 2</td>
<td>Duggal 2021(^5)</td>
<td>Modelling(^d)</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>eGFR(^c)</td>
<td>KQ 2</td>
<td>Hoenig 2022(^6)</td>
<td>Modelling(^d)</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>eGFR(^c)</td>
<td>KQ 2</td>
<td>Inker 2021(^6)</td>
<td>Modelling(^d)</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>eGFR(^c)</td>
<td>KQ 2</td>
<td>Mahmud 2022(^7)</td>
<td>Modelling(^d)</td>
<td>↑</td>
<td>↑</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>eGFR(^c)</td>
<td>KQ 2</td>
<td>Miller 2021(^8)</td>
<td>Modelling(^d)</td>
<td>*</td>
<td>*</td>
<td>↑</td>
</tr>
<tr>
<td></td>
<td>eGFR(^c)</td>
<td>KQ 2</td>
<td>Panchal 2022(^9)</td>
<td>Modelling(^d)</td>
<td>↑</td>
<td>↑</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>eGFR(^c)</td>
<td>KQ 2</td>
<td>Shi 2021(^1)</td>
<td>Modelling(^d)</td>
<td>↑</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>eGFR(^c)</td>
<td>KQ 2</td>
<td>Tsai 2021(^1)</td>
<td>Modelling(^d)</td>
<td>↑</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>eGFR(^c)</td>
<td>KQ 2</td>
<td>Yap 2021(^1)</td>
<td>Modelling(^d)</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>eGFR(^c)</td>
<td>KQ 2</td>
<td>Zelnick 2021(^1)</td>
<td>Modelling(^d)</td>
<td>↑</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>eGFR(^c)</td>
<td>KQ 2</td>
<td>Coresh 2019(^1)</td>
<td>Modelling(^d)</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td><strong>Kidney transplant allocation</strong></td>
<td>Kidney Donor Index</td>
<td>KQ 2</td>
<td>Julian 2017(^1)</td>
<td>Modelling(^d)</td>
<td>*</td>
<td>*</td>
<td>↑</td>
</tr>
<tr>
<td></td>
<td>Revised KAS(^c)</td>
<td>KQ 1</td>
<td>Zhang 2018(^1)</td>
<td>Pre-post</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td><strong>Severity of illness scores for Crisis Standards of Care</strong></td>
<td>SOFA</td>
<td>KQ 1</td>
<td>Miller 2021(^2)</td>
<td>Modelling(^d)</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>SOFA, LAPS2</td>
<td>KQ 1 and 2</td>
<td>Ashana 2021(^2)</td>
<td>Modelling(^d)</td>
<td>*</td>
<td>↑</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>APACHE Iva, OASIS, SOFA</td>
<td>KQ 1</td>
<td>Sarkar 2021(^1)</td>
<td>Modelling(^d)</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td><strong>Prostate Cancer Risk</strong></td>
<td>PCPT(^c)</td>
<td>KQ 1</td>
<td>Carbanaru 2019(^1)</td>
<td>Modelling(^d)</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>KPCC RC(^c)</td>
<td>KQ 1</td>
<td>Presti 2021(^1)</td>
<td>Modelling(^d)</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

\(^a\)Not reported

**Summary Evidence Map**

- **Direction of Effect:** (arrow direction)
  - ↑ Increase
  - ↓ Decrease
  - ↔ No effect

[Image of Summary Evidence Map]
## Further evidence from KQ 2: Algorithms Perpetuate Disparities

### Evidence Map (Continued)

<table>
<thead>
<tr>
<th>Clinical Category</th>
<th>Algorithm</th>
<th>Key Question</th>
<th>Study Design</th>
<th>Disparities in Health outcome</th>
<th>Disparities in Access</th>
<th>Disparities in Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liver transplantation</td>
<td>Donor Risk Index</td>
<td>KQ 2</td>
<td>Shores 2013</td>
<td>*</td>
<td>*</td>
<td>↑</td>
</tr>
<tr>
<td></td>
<td>ASCVD</td>
<td>KQ 2</td>
<td>Weale 2021</td>
<td>*</td>
<td>*</td>
<td>↑</td>
</tr>
<tr>
<td></td>
<td>Modified ASCVD</td>
<td>KQ 2</td>
<td>Topel 2018</td>
<td>↑</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>ASCVD</td>
<td>KQ 2</td>
<td>Fairman 2020</td>
<td>↑</td>
<td>↑</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Pooled cohort equations</td>
<td>KQ 2</td>
<td>Yadlowsky 2018</td>
<td>*</td>
<td>↑</td>
<td>♬</td>
</tr>
<tr>
<td></td>
<td>Framingham risk score</td>
<td>KQ 2</td>
<td>Fox 2016</td>
<td>*</td>
<td>↑</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Framingham risk score</td>
<td>KQ 2</td>
<td>Drawz 2012</td>
<td>*</td>
<td>*</td>
<td>↑</td>
</tr>
<tr>
<td>Cardiovascular risk</td>
<td>ASCVD</td>
<td>KQ 2</td>
<td>Weale 2021</td>
<td>*</td>
<td>*</td>
<td>↑</td>
</tr>
<tr>
<td></td>
<td>ASCVD</td>
<td>KQ 2</td>
<td>Topel 2018</td>
<td>*</td>
<td>*</td>
<td>↑</td>
</tr>
<tr>
<td></td>
<td>ASCVD</td>
<td>KQ 2</td>
<td>Fairman 2020</td>
<td>↑</td>
<td>↑</td>
<td>*</td>
</tr>
<tr>
<td>Lung Cancer Screening</td>
<td>USPSTF-2013</td>
<td>KQ 1</td>
<td>Pasquinelli 2021</td>
<td>*</td>
<td>↑</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>USPSTF-2013</td>
<td>KQ 1</td>
<td>Han 2020</td>
<td>*</td>
<td>*</td>
<td>↑</td>
</tr>
<tr>
<td></td>
<td>USPSTF-2020</td>
<td>KQ 1</td>
<td>Landy 2021</td>
<td>↑</td>
<td>↑</td>
<td>*</td>
</tr>
<tr>
<td>Lung Transplant Allocation</td>
<td>Lung Allocation System</td>
<td>KQ 1</td>
<td>Wille 2013</td>
<td>*</td>
<td>↓</td>
<td>*</td>
</tr>
<tr>
<td>Lung Function</td>
<td>GLI Spirometry Equation</td>
<td>KQ 2</td>
<td>Baugh 2022</td>
<td>*</td>
<td>*</td>
<td>↑</td>
</tr>
<tr>
<td></td>
<td>GLI Spirometry Equation</td>
<td>KQ 2</td>
<td>Elmaleh-Sachs 2021</td>
<td>*</td>
<td>*</td>
<td>↑</td>
</tr>
<tr>
<td>Anticoagulation</td>
<td>Warfarin dosing algorithms</td>
<td>KQ 2</td>
<td>Kimmel 2013</td>
<td>↑</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Warfarin dosing algorithms</td>
<td>KQ 2</td>
<td>Lindi 2015</td>
<td>↑</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>CHA2DS2-VASc</td>
<td>KQ 2</td>
<td>Kabra 2016</td>
<td>*</td>
<td>*</td>
<td>↑</td>
</tr>
<tr>
<td>Emergency Department Triage</td>
<td>HEART Pathway</td>
<td>KQ 1</td>
<td>Snavely 2021</td>
<td>**</td>
<td>*</td>
<td>↑</td>
</tr>
<tr>
<td>Other</td>
<td>Novel algorithm for high-risk care management</td>
<td>KQ 1 and 2</td>
<td>Obermeyer 2019</td>
<td>*</td>
<td>↑</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Natural language processing algorithm</td>
<td>KQ 1 and 2</td>
<td>Thompson 2021</td>
<td>*</td>
<td>*</td>
<td>↑</td>
</tr>
</tbody>
</table>

Direction of Effect: (arrow direction)

- ↑ Increase
- ↓ Decrease
- ↔ No effect

*Not reported*
KQ 1 Summary

• The effect of algorithms is complex, and some have been shown to perpetuate or exacerbate disparities, some reduce disparities, and others have no effect
• Additionally, an algorithm may exacerbate disparities for one outcome, but reduce disparities for another outcome
• Many algorithms in clinical use perpetuate or exacerbate racial and ethnic disparities (e.g. eGFR, ASCVD)
• Disparities can be reduced, regardless of whether race and ethnicity are utilized in the algorithm, when disparities are outlined and used to inform algorithm development (e.g. KAS, prostate CA screening)
• Most of the evidence focused on non-AI algorithms
Contextual Question 1: Extent of inclusion of input variables based on race and ethnicity in algorithms?

• We examined 45 algorithms
  ► 17 include race and ethnicity
  ► 5 include measures that may serve as proxies for race and ethnicity (e.g., SDOH, healthcare costs)
  ► Clinical category, setting, and purpose varied
  ► Developers included clinical research teams, organizations setting healthcare policy, health plans, EHR vendors

• We examined additional resources (e.g., websites)
  ► MDCalc – 14 of 700+ algorithms include race and ethnicity
Contextual Question 1: Extent of inclusion of input variables based on race and ethnicity in algorithms?

- Excluded >800 studies due to study design and outcome reporting
  - Many included similar algorithms included in our review
  - Some conducted in specialties not included in our review
- Algorithms likely affect every medical specialty, healthcare setting, and patient population
- Tip of the iceberg – review was limited in scope and may not fully represent larger environment
Creating Fair, Reliable and Useful Models

Nigam H. Shah, MBBS, PhD
Professor of Medicine and Biomedical Data Science
Chief Data Scientist, Stanford Healthcare
Associate Dean for Research, Stanford School of Medicine

Racial Bias and Healthcare Algorithms
March 2, 2023
12:20-12:28 p.m. ET
Stanford Medicine Program for AI in Healthcare

The Model

Al-guided work

Policy & capacity

The action
Supporting Algorithmic Equity in a Public Healthcare System: a Case Study in Opioid Safety

Suzanne Tamang, PhD
Veterans Affairs

Racial Bias and Healthcare Algorithms
March 2, 2023
12:36-12:44 p.m. ET
STORM: Family of Decision Support Tools to Support Safe Care of Patients Exposed to Opioids

Includes: Predictive analytics for risk stratification, flexible population management, summary information on risk mitigation implementation for targeting QI and education, recommendation and tracking of risk mitigation, and patient level care review.
• In 2019, PERC worked with *Data Science for Social Good* and the *FDA’s Office of Minority Health & Health Equity* to develop a performance evaluation framework on de-identified data (2014-15)

• Using a diverse set of stakeholders, and visually driven model “diagnostics”, we quantified differences in performance, by gender, age, race/ethnicity
  ▶ **AUROC**
  ▶ **PR Curves**
  ▶ **Calibration**
  ▶ **False-negative and false-positive parity rates**

• We found evidence of **algorithmic bias**, but also salient challenges interpreting results of under-represented minority groups (e.g., American Indian/Alaskan Native, Asian) and “interactions” (e.g., female and >65, female and Black or African-American).
Example #1 of Racial & Ethnic Bias: Calibration

• **Calibration** is defined as the following property:
  
  • “If we assign some group a risk of $x$, the actual outcome incidence rate should also be $x$”

• For example, if we assign a group of people a risk of 40%, the actual overdose/suicide-related incidence rate should also be 40%.
• In 2021, PERC applied the framework to STORM-2 (2014-1015)
• STORM-2 is three models:
  ► No opioids in the observation window
  ► Discontinued during the observation window
  ► Actively on opioids on the index date
• Extended PERC framework to include:
  ► Per true-positive plot: for each true positive, how many false-negatives and false-positives are detected?
  ► False Omission Rate: Given a negative prediction, the FOR tells you the probability that the true value is positive.
Example #2 of Racial & Ethnic Bias: FOR

- False Omission Rate for ActiveRx

Esther Meerwijk PhD, Data Scientist, Ci2i, Palo Alto VA

STORM mandated case review cutoff: 0.0609, "VERY HIGH"
Where Are We in 2023?

• Fostering and engaging a VA Community of Practice for modeling and monitoring
  ▶ STORM, REACH VET, CAN, Rockies NLP
  ▶ NAII Datasheets and Model Cards

• Next steps for suicide and overdose prediction models
  ▶ Apply framework to more recent data (2016-2020) and new subgroups
  ▶ Comparing methods for mitigating bias *(Duncan McElfresh PhD, HSR&D Fellow)*
    - Regression calibration – apply a subgroup specific transformation
    - Subgroup-specific models – fit separate models
    - Subgroup-specific cut points – define different high-risk cutoffs
  ▶ Develop dashboard to monitor performance over time
    - Empirically inform recalibration of model, predictors to include in STORM and alternative prediction algorithms

Suzanne Tamang, PhD
Acknowledgements

PERC Team
• Esther Meerwijk, PhD
• Duncan McElfresh, PhD
• Jodie Trafton, PhD
• Susana Martins, MD
• Amy Robinson, PharmD
• Elizabeth Oliva, PhD

Partners
• Craig Kreisler, PhD
• Joseph Erdos, MD
• Michael Jonathan Stringer
• Christine Lee PharmD
• John Scott, MD
• Jonathan Nebeker, MD

• OMHSP
• PERC
• PERC Platform Support
• HDAP
• BISL
• OIT
Impact of Healthcare Algorithms on Racial and Ethnic Disparities in Health and Healthcare

Consensus Panel Discussion/Q & A

March 2, 2023
12:45– 1:30 p.m. ET
Discussion Questions

• What’s missing, in terms of other experience and insights from the audience or related topics that were not covered in this session?

• What guidance is needed to mitigate bias/what are the next steps, for different parts of AI lifecycle, implementation perspective?
  ► When/what/where/how to use algorithms?
  ► Addressing bias in existing algorithms?
Impact of Healthcare Algorithms on Racial and Ethnic Disparities in Health and Healthcare

Break
March 2, 2023
1:30 – 2:00 p.m. ET

Please take thirty minutes for lunch
Evidence Review

Contextual Question 3

Brian Leas, MS, MA
University of Pennsylvania School of Medicine

Racial Bias and Healthcare Algorithms
March 2, 2023
2:00-2:20 p.m. ET
To what extent are patients, providers (e.g., clinicians, hospitals, health systems), 
payers (e.g., insurers, employers), and 
policymakers (e.g., healthcare and 
insurance regulators, state Medicaid 
directors) aware of the inclusion of 
variables based on race and ethnicity in 
healthcare algorithms?
CQ 3: Methodology

• Primary literature searches

• AHRQ’s Request for Information

• Technical Expert Panel and Key Informants

• Feedback from peer reviewers
CQ 3: Key Informants and Technical Expert Panel

12 Key Informants (KIs)
10-member Technical Expert Panel (TEP)

- Experts in research and practice
  - Healthcare algorithm development, use, and auditing
  - Health and healthcare disparities; health equity; race and ethnicity in healthcare

- Healthcare providers
  - Clinicians, health systems, academic medical centers, public and community health, specialty societies

- Patient advocates

- Payers (commercial and government)

- Vendors of health IT systems and healthcare algorithms

- Federal agencies
CQ 3: Patient Perspectives

Challenges

• Limited awareness and understanding
  ▶ How algorithms are used in healthcare
  ▶ How race and ethnicity interacts with health and healthcare
• Literacy (health, science, tech)
• Views shaped by personal/family experiences

Opportunities

• Patient-centered care and shared decisionmaking
• Personalized medicine and genetics
CQ 3: Provider Perspectives

Individual Clinicians
- Limited understanding
  - Know how and when to use algorithms
  - Don’t understand development, implementation, sources of bias
- Deference and trust
  - Regulators, societies, health systems, EHRs

Hospitals and health systems
- Focused on implementation, not potential sources of bias
- Adapt EHR products to patient population, incentives, priorities (“off-label” use)
- Minimal transparency
Medical education is an opportunity to address many concerns

- Critical thinking about algorithms
- Use of clinical practice guidelines and EHR tools
- Human genetics
- Race, ethnicity, biology
- Disparities and equity
- Population health
CQ 3: Payers

• Not highly focused on disparities

• Just following the data

• Minimal transparency

• Decentralized operations, disjointed regulations
CQ 3: Policymakers

• All sectors anticipating federal guidance
• Substantial activity in last 3 years

Challenges
• Multiple agencies with overlapping stakes
• Who should guidance/regulation address?
  ► EHR vendors, algorithm and AI developers, auditors, payers, providers
• How to address proprietary data and systems?
• Limited evidence!
Addressing Racial Bias in Healthcare Algorithms: Steps You Can Take Today

Crystal Grant, PhD
Technology Fellow, Speech, Privacy, and Technology Project, American Civil Liberties Union

Racial Bias and Healthcare Algorithms
March 2, 2023
2:20-2:30 p.m. ET
Assume the Healthcare Algorithm is Biased.

- Garbage in, Garbage out. Bias in, bias out.
- The data on which algorithms are trained reflects all sociocultural and environmental realities of racism in America’s present and past and its effects on people’s biology.
  - There is no genetic basis of race. Race is a social construct with real-world effects.
- While techniques exist that attempt to mitigate these biases in the training data, they too present limitations.

The richest Black women have infant mortality rates at about the same level as the poorest white women.

<table>
<thead>
<tr>
<th>Infant deaths per 100,000 for mothers who are...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black</td>
</tr>
<tr>
<td>White</td>
</tr>
<tr>
<td>Rate for richest Black mothers</td>
</tr>
<tr>
<td>437</td>
</tr>
<tr>
<td>350</td>
</tr>
<tr>
<td>Rate for poorest white mothers</td>
</tr>
</tbody>
</table>

Childbirth Is Deadlier for Black Families Even When They’re Rich, Expansive Study Finds
• Algorithm developers are not subject matter experts in patient care. Yet, in creating a healthcare tool, they are making what amount to clinical and medical decisions.
Assume the Algorithm Has not Had Adequate Oversight or Regulation.

• Many healthcare algorithms undergo no oversight and do not require FDA approval.

• Among tools regulated by FDA, in obtaining approval/clearance:
  ► Assessments of performance bias across racial or ethnic groups are not required
  ► If provided, this data isn’t made accessible to the public or researchers
  ► Overuse of the 510(k) clearance process claiming substantial similarity may lead to less rigorous testing than is ideal for influential health algorithms

• After approval or clearance, degradation in the performance of an algorithm when deployed in RWD can occur, yet the FDA doesn’t penalize those who fail to conduct post-market studies.
Conclusion: Steps You Can Take Today

• If we assume the healthcare algorithm we plan to use is biased, incorrect, and under-regulated:
  ► **Administrators**: Demand more transparency from vendors on how a tool was built, results from bias testing, interrogate why certain outputs result given certain inputs. Partner with researchers to conduct ongoing reviews.
  ► **Clinicians**: Question an algorithm that uses patients’ race to assume biological information about them; stay alert for “anecdotal” bias in tools.
  ► **Researchers**: Push federal regulatory bodies to make data from algorithm developers available. Assess whether performance of a tool at approval/clearance holds up in use with RWD, and if any biases emerge.

Email: cgrant@aclu.org, Twitter: @itscrystalgrant

ACLU WHITE PAPER: AI IN HEALTH CARE MAY WORSEN MEDICAL RACISM
Strategies to Address Algorithmic Bias in Medicine

Helen Burstin, MD, MPH, MACP
CEO, Council of Medical Specialty Societies

Racial Bias and Healthcare Algorithms
March 2, 2023
2:30-2:40 p.m. ET
CMSS Member Societies
Widespread Issue in Clinical Algorithms*

- Cardiology
- Nephrology
- Hematology/Oncology
- Neurology
- Hepatology
- Endocrinology
- Infectious diseases
- Obstetrics
- Pulmonary medicine
- Transplant medicine
- Urology
- Addiction medicine
- Surgery
- Mental health

* Specialties represented in 45 algorithms included in AHRQ report
Draft Recommendations: Specialty Societies

• Promote stakeholder awareness (including patients) of potential algorithmic risk
• Work with policymakers to review clinical algorithms, and address those that result in racial and ethnic inequities
• Ensure that algorithms included in clinical guidelines and recommendations statements are assessed from a health equity lens and that methods are adequately reported
• Invest in further research to assess the effect of algorithms on racial and ethnic disparities before widespread implementation
Recognize that any change in eGFR reporting must consider the multiple social and clinical implications, be based on rigorous science, and be part of a national conversation about uniform reporting of eGFR across health care systems.

 Attempt to incorporating concerns of patients and the public, especially in marginalized and disadvantaged communities, while rigorously assessing the underlying scientific and ethical issues embedded in current practice.

 Working towards an unbiased approach to assessment of kidney function so that laboratories, clinicians, patients, and public health officials can make informed decisions to ensure equity and personalized care for patients with kidney diseases.

 Keep laboratories, clinicians, and other kidney health professionals apprised.

 Identify any potential long-term implications of removing race from the eGFR formula.
“Race-based medicine has been pervasively interwoven into the fabric of health care delivery in the United States for more than 400 years. Race is a historically derived social construct that has no place as a biologic proxy.

In addition to valid measures of social determinants of health, the effects of racism require consideration in clinical decision-making tools in ways that are evidence informed and not inappropriately conflated with the limiting phenotype of race categorization.

This policy statement addresses the elimination of race-based medicine part of a broader commitment to dismantle the structural as and systemic inequities that lead to racial health disparities.”
Obstetrics: Implementation Approach

• Vaginal Birth after C-section (VBAC) Calculator
  ► VBAC Calculator revised
    – MFM Network, May 2021
  ► Analysis with and without race and ethnicity
    – Am J Ob Gyn, Dec 2021
  ► Updated VBAC online calculator from MFM does not include race/ethnicity; added new variable related to treatment for chronic hypertension
  ► Further clinician and patient education and dissemination
Potential Next Steps (1)

• Develop standards regarding inclusion of race in clinical research that support development of clinical guidelines and algorithms
• Support research that assesses the impact of race in clinical algorithms, recognizing importance of context, intentionality, and outcomes
• Support research that assesses the impact of other drivers, including SDOH and structural racism
• Effectively communicate and educate patients and clinicians on the potential impact of race in clinical algorithms
Potential Next Steps (2)

• Cross-specialty learning to develop best approaches to assess/remove race in clinical algorithms, assess long-term implications, and effective dissemination/implementation strategies

• Cross-disciplinary partnerships to develop AI/ML data sets that could support prospective assessment of race in clinical algorithms

• Broad stakeholder engagement that leads to changes in clinical research standards and clinical practice
Impact of Healthcare Algorithms on Racial and Ethnic Disparities in Health and Healthcare

Consensus Panel Discussion/Q & A

March 2, 2023
2:40– 3:25 p.m. ET
Discussion Questions

• What works, what’s missing in terms of related topics, experience, and insights, including trust issues related to algorithmic biases?

• What guidance is needed to mitigate bias/what are the next steps, for different parts of AI lifecycle?
  ▶ Approaches to increasing awareness and building trust among health professionals and communities, especially vulnerable groups and minorities?
  ▶ Approaches to involving patients and clinicians more fully in these efforts?
Impact of Healthcare Algorithms on Racial and Ethnic Disparities in Health and Healthcare

Closing Remarks

Craig Umscheid, MD, MS
Senior Science Advisor and Director, EPC Division, AHRQ
Racial Bias and Healthcare Algorithms
March 2, 2023
3:25-3:40 p.m. ET
Thank you!