Stanford MEDICINE

What Can Go Wrong?

Artificial Intelligence in Biomedical and Behavioral Research

Dr. Tina Hernandez-Boussard, Ph.D., MPH, FACMI

Associate Dean of Research
Professor of Medicine, Biomedical Data Science, Surgery, Epidemiology & Population Health

Promise of Data-Driven Medicine

- Transformative potential in diagnosis, treatment, and outcomes.
- Can help decision-making be faster and more accurate.
- Generative AI is the new frontier with many possibilities

Double-Edged Sword

AI APPLICATIONS IN HEALTHCARE

BIOMEDICAL RESEARCH

- Automated experiments
- Automated data collection
- Gene function annotation
- Literature mining

TRANSLATIONAL RESEARCH

- Biomarker discovery
- Drug-target prioritization
- Genetic variant annotation

MEDICAL PRACTICE

- Disease diagnosis
- Treatment selection
- Patient monitoring
- Risk stratification models

When AI in Healthcare Goes Wrong

- **Bias in Algorithmic Decision Making**

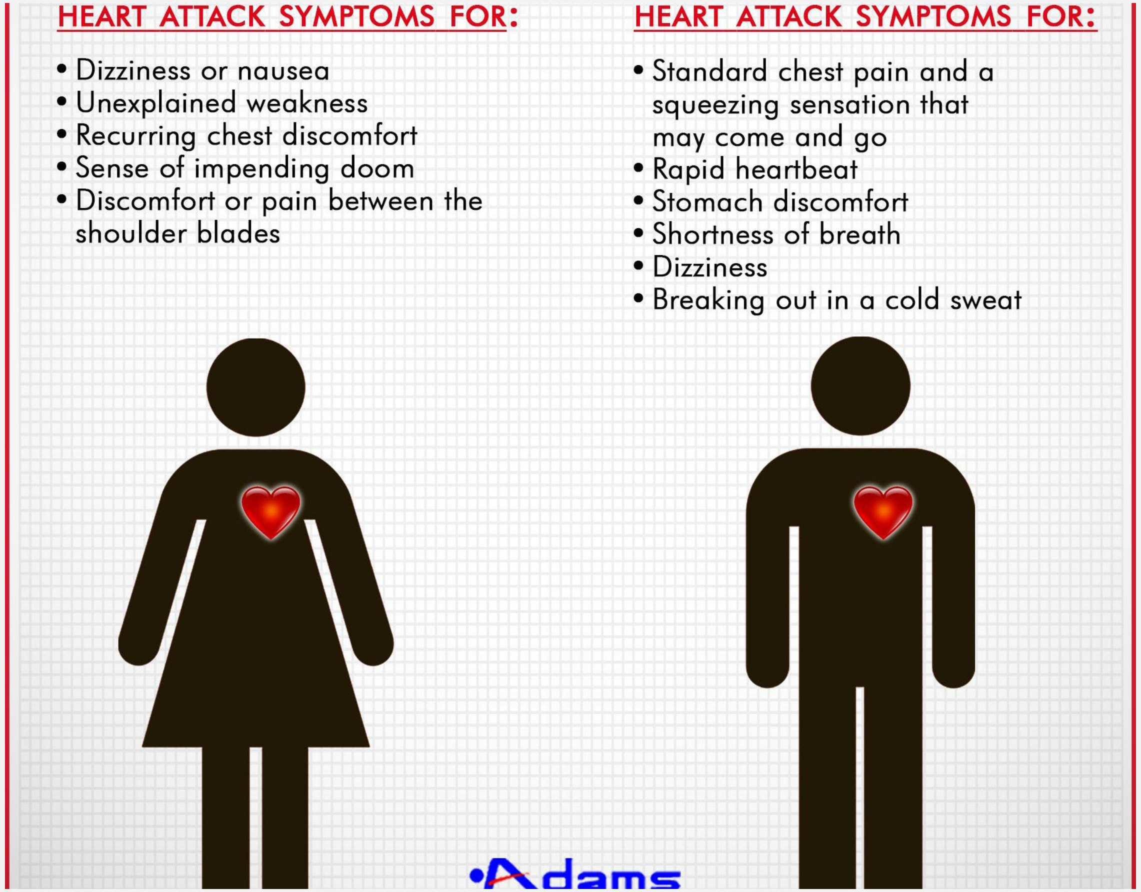
  Obermeyer Science 2019 – Allocation of resources based on cost not need

- **Privacy Breaches and Data Misuse**
  - Hospital Data Breaches
  - Unauthorized Data Sharing

- **Faulty or Inaccurate Medical Devices and Software**
  - Pulse Oximeter and skin pigmentation bias

- **Overreliance on Technology**
  - Art of Medicine

Where AI can Go Wrong



Predicting COVID-19 Outcomes

- **COVID-19 presents a disproportionate impact on minorities**
  - Economic and social circumstances
  - Existing health disparities

Bias at Warp Speed: How AI may Contribute to the Disparities Gap in the Time of COVID-
Röösli E, Rice B, Hernandez-Boussard T. J Am Med Inform Assoc. 2020 PMID: 3280500

Ethical and Societal Implications of AI in Healthcare

- **Public Trust and Perception**
  - Growing concerns about data privacy and security
  - Lack of standards to communicate data use and re -use
- **Healthcare Commercialization**
  - Commercial use of health data by technology companies
  - Ethical concerns over profit motives overriding patient welfare.
- **Regulatory and Legal Landscape**
  - Evolving legal frameworks to address novel challenges
  - Limited policies to protect individuals and ensure equitable AI
- **Cultural and Ethical Norms**
  - Social Determinants of Health in health data and data -driven healthcare.
  - Balance technological innovation with ethical considerations and societal values.
- **Inequities in Healthcare Access and Outcomes: Digital Divide**
  - Risk of widening health inequalities due to uneven distribution of technology in health
  - Community health centers, rural health settings, certain patient groups

Stanford MEDICINE

Real world examples of societal harm caused by AI in healthcare

Race-Based Medicine

Medical practice guided by algorithms that include race and/or ethnicity

- Race as a proxy for biological differences
- Race fails to account for other factors associated with health outcomes
- Reinforces stereotypes

Hernandez-Boussard T, Siddique SM, Bierman AS, Hightower M, Burstin H. Promoting Equity In Clinical Decision Making: Dismantling Race-Based Medicine. Health Aff (Millwood). 2023 Oct;42(10):1369-1373.

Race-Based Medicine Examples

Kidney Disease -eGFR

- eGFR (estimated glomerular filtration rate) used to diagnose & monitor kidney disease *adjusted* for race
  - Assumed Black patients had inherently higher creatinine levels.

Race is a social construct with greater variation within race than between races

diagnosed later with more severe disease
  - reduced chances of being placed on kidney transplant waitlists

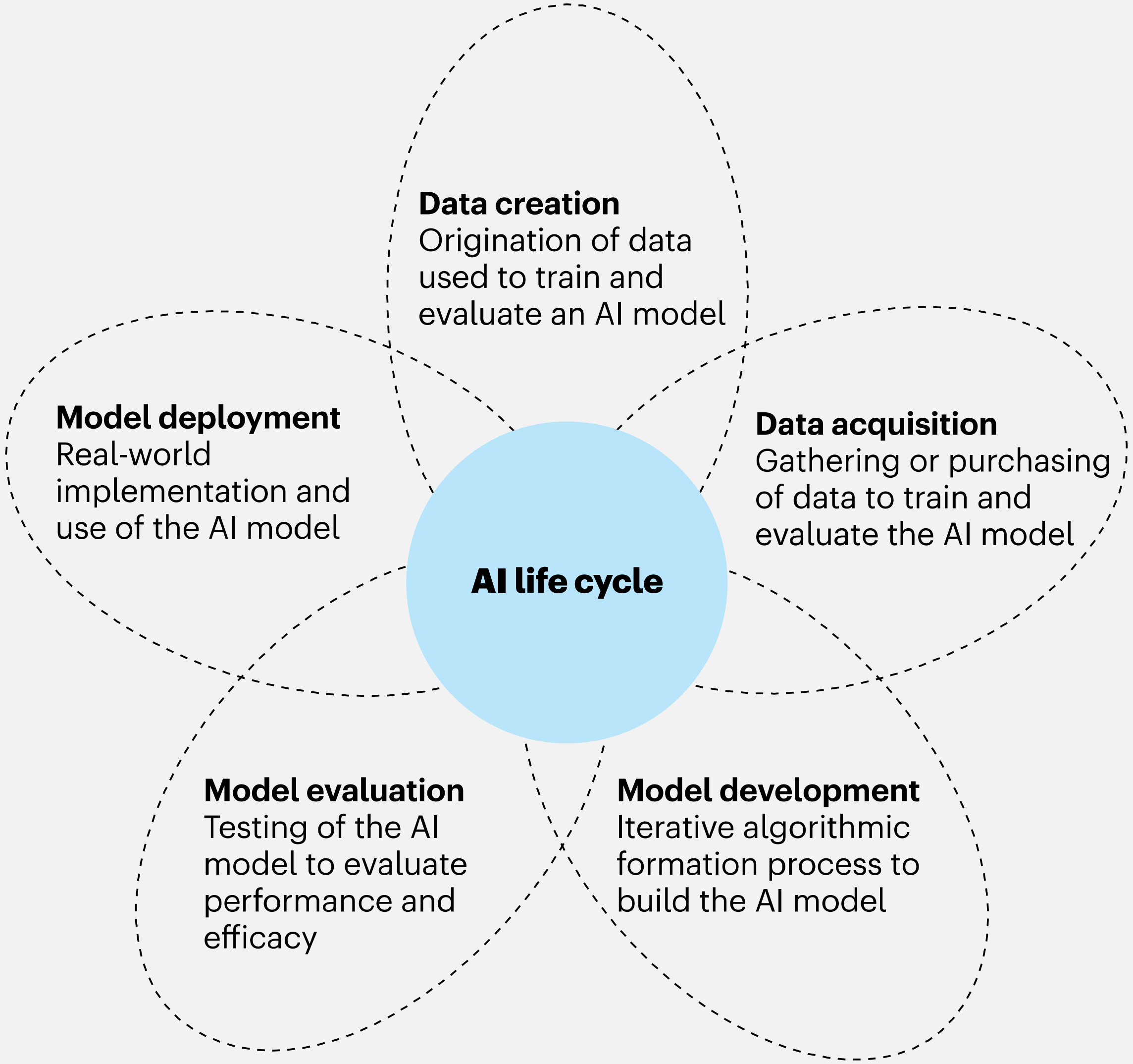
Vaginal Births after Cesarean

- Vaginal Birth after Cesarean (VBAC) Calculator estimates risk of an adverse outcome with vaginal birth after a cesarean *adjusted* for race
  - Assumed Black patients were at higher
- VBAC overestimated risk of adverse events, particularly among Black women
  - Differential recommendations
  - Limited treatment options

How do we develop Fair and Reliable Models?

Stanford MEDICINE

Life Cycle of the Artificial Intelligence



Ng MY, Kapur S, Blizinsky KD, Hernandez-Boussard T. The AI life cycle: a holistic approach to creating ethical AI for health decisions. Nat Med. 2022 Nov;28(11):2247-2249.

Stanford MEDICINE

Transparency at Every Stage

- Question formation
  - Stakeholder involvement
- Data Creation
  - How data were produced
  - Circumstances of data generation
  - Data quality
  - Data diversity
- Data Acquisition
  - Data availability
  - Legal & Regulatory mandates
  - Internal Review Boards
  - Financial agreements
- Model Development
  - Ground truth
  - Training data
  - Hyper-tuning
- Model Evaluation
  - Beyond performance
  - Parity & Calibration
  - De-biasing steps
- Deployment
  - Technical barriers
  - Integration issues
  - Organizational/technical challenges
  - Regulatory compliance

Fairness Framework to Evaluate AI


scientific data

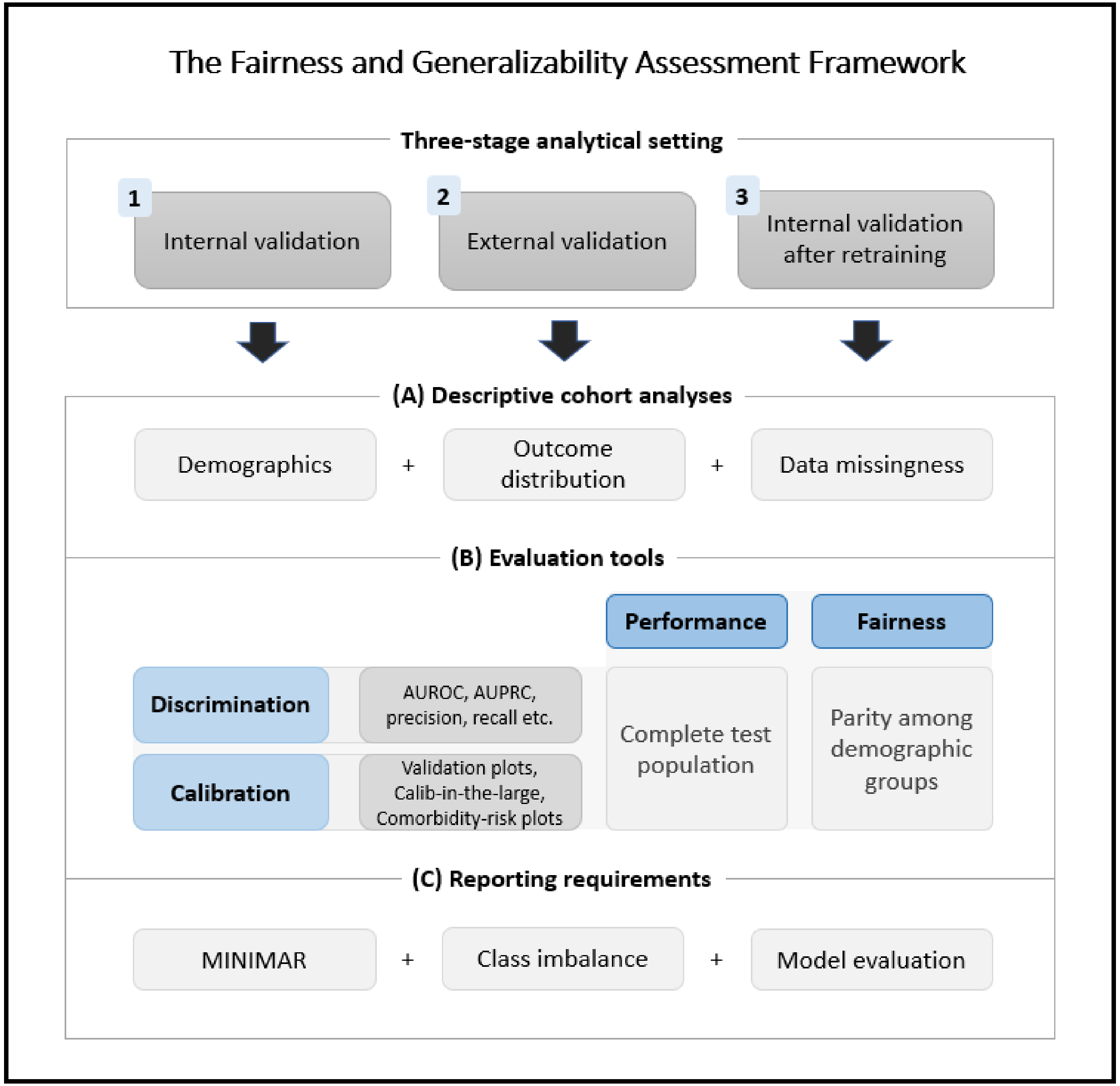
Explore content ⌄ About the journal ⌄ Publish with us ⌄

nature › scientific data › articles › article

Article | Open Access | Published: 24 January 2022

Peeking into a black box, the fairness and generalizability of a MIMIC-III benchmarking model

Eliane Röösli, Selen Bozkurt & Tina Hernandez-Boussard



Life Cycle of the Data-Driven Tools

M Chin, et al. 2023, JAMA Network Open

Guiding Principles

- Promote Health and Healthcare Equity During All Phases of the Healthcare Algorithm Lifecycle
- Ensure Healthcare Algorithms and their Use are Transparent and Explainable
- Authentically Engage Patients and Communities During All Phases of the Healthcare Algorithm Lifecycle, and Earn Trustworthiness
- Explicitly Identify Healthcare Algorithmic Fairness Issues and Tradeoffs
- Ensure Accountability for Equity and Fairness in Outcomes from Healthcare Algorithms

Stanford MEDICINE

Transparency Cycle

- Data Availability and Model Accuracy
  - Who does the model represent?
  - How does the model perform across populations
- AI bias can be prejudice and results in differential treatments and outcomes
- Impacts communities & erodes trust
  - Crucial implications related to data sharing and patient engagement
- Reduced Engagement
  - Communities reluctant to share their data, perpetuating the cycle of inaccuracy and mistrust.

Catch 22 for AI Equity

Lack of Transparency

Limited training data and suboptimal model performance

Breaks Trust

T

Erosion of trust within communities

Disparities in Outcomes

Prejudice or inaccurate model performance

No Engagement & Participation

Reluctance to share data

thank you!

boussard-lab.stanford.html

boussard@stanford.edu