USE CASE
Eyes Wide Open: Hopkins' Transformative AI for Unbiased Digital Retinal Imaging

Dr Aisha Rahim

Medical Executive and Director

Johns Hopkins HealthCare

13-14 September
INTELLIGENT HEALTH 2023
Basel, Switzerland

PRODUCED BY
INSPIRED MINDS!

Accredited cpd Event

@IntHealthAI
#IntelligentHealthAI
#SaveLivesWithAI

HEADLINE PARTNER
NOVARTIS | The Novartis Foundation

PLATINUM SPONSOR
INNOVATION & TECHNOLOGY
Deloitte.

CITY PARTNER
Basel
basel.ch

Eyes Wide Open Hopkins Transformative AI for Unbiased Retinal Imagaing

AI Bias: A Threat to Equity in Healthcare

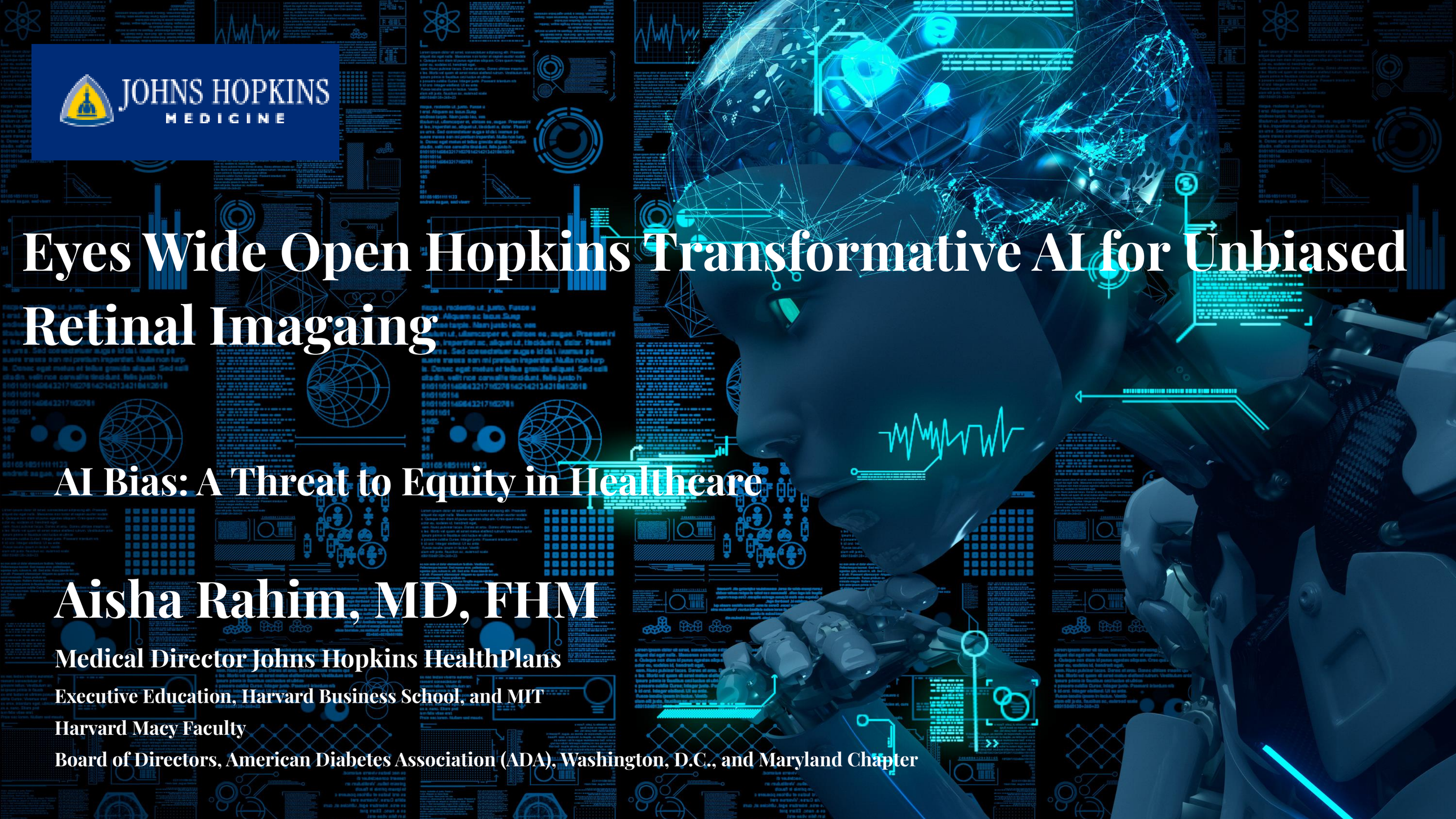
Aisha Rahim, MD, FHM

Medical Director Johns Hopkins HealthPlans

Executive Education, Harvard Business School, and MIT

Harvard Macy Faculty

Board of Directors, American Diabetes Association (ADA), Washington, D.C., and Maryland Chapter



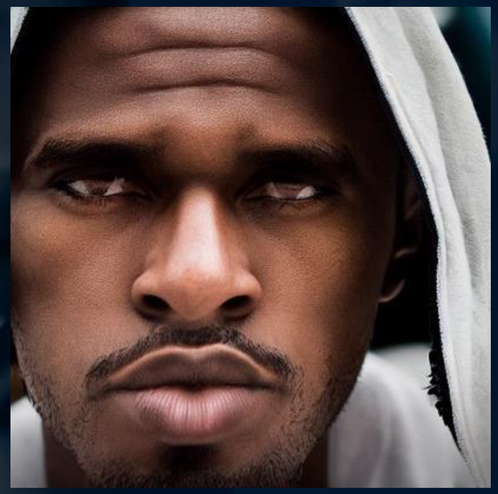
AGENDA

- Introduction
- AI Bias in Healthcare: An Overview
- The Effects of AI Bias on Healthcare
- BIG PHARMA- AI-BASED DRUG DISCOVERY
- Best Practices for Overcoming AI Bias in Healthcare
- Overcoming AI Bias in Healthcare: Case Study from John Hopkins, Practical Steps and Approaches
- Conclusion

JOHNS HOPKINS
MEDICINE

State of Bias in AI

The world, according to Stable Diffusion is run by White male CEOs. Women are rarely doctors, lawyers or judges. Men with dark skin commit crimes, while women with dark skin flip burgers.



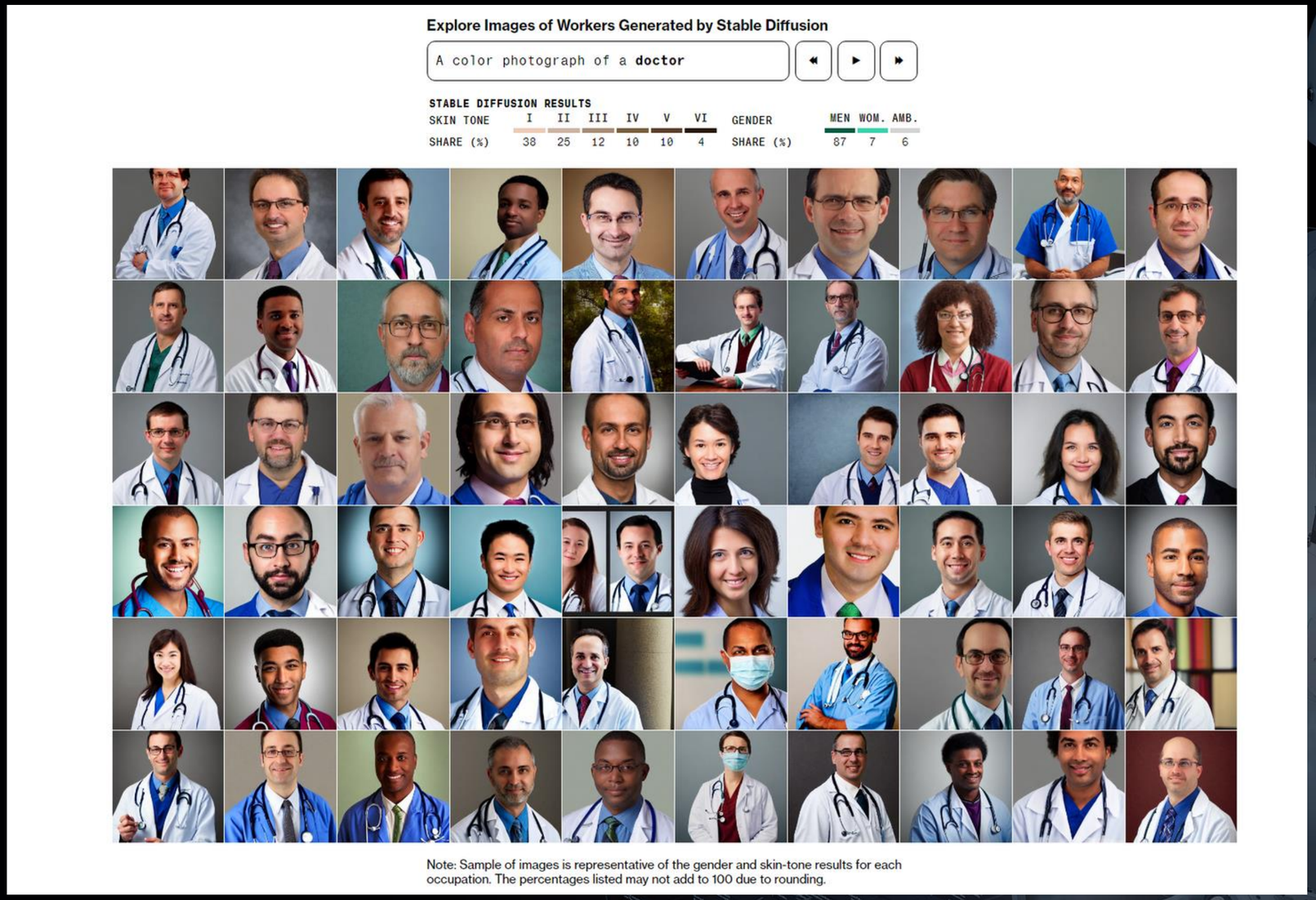
Explore Images of Workers Generated by Stable Diffusion

A color photograph of a **CEO**

STABLE DIFFUSION RESULTS

SKIN TONE	I	II	III	IV	V	VI
SHARE (%)	59	19	10	9	3	1

GENDER	MEN	WOM.	AMB.
SHARE (%)	94	5	1



Reality Vs AI Bias

- According to the data released by the AAMC (Association of American Medical Colleges), in 2019, women comprised the majority of enrolled U.S. medical students for the first time1. (Source AAMC 2019)
- Women made up 50.5% of all medical school students. The number of applicants to medical schools rose by 1.1% from 2018 to 2019, to a record 53,371, and the number of matriculants (new enrollees) grew by 1.1%, to 21,8691.

Explore Images of Workers Generated by Stable Diffusion

A color photograph of a **fast-food worker**

STABLE DIFFUSION RESULTS

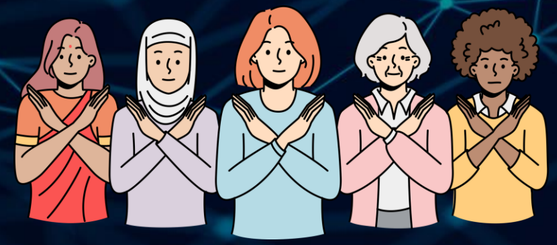
SKIN TONE	I	II	III	IV	V	VI
SHARE (%)	13	10	7	12	31	27

GENDER	MEN	WOM.	AMB.
SHARE (%)	39	38	23

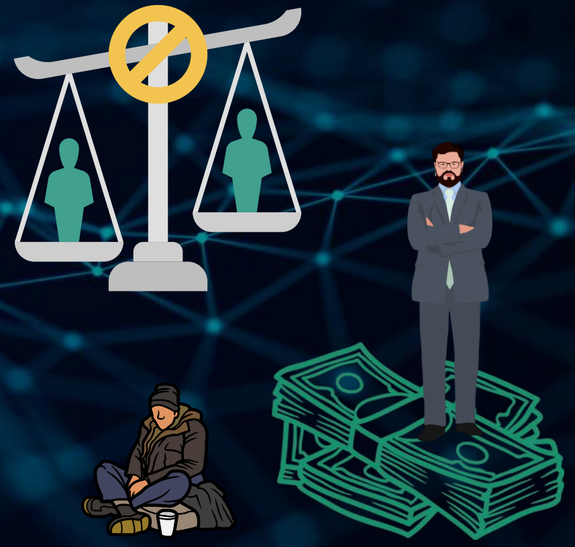
AI Bias in Healthcare: An Overview

Gender bias

Racial bias



Socioeconomic bias



The Effects of AI Bias on Healthcare

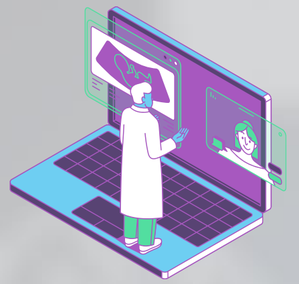
Types of Bias	AI Bias	Description	Source
Data bias		The data used to train AI models can be biased, either intentionally or unintentionally. This can happen if the data is not representative of the population or if it is collected in a way that introduces bias.	Scientific American. (2020, Nov 17). Health Care AI Systems Are Biased
Algorithmic bias		The algorithms used to train AI models can also be biased. This can happen if the algorithm is designed in a way that favors certain groups of people.	Harvard T.H. Chan School of Public Health. (2021, March 12). Algorithmic bias in health care exacerbates social inequities
Human bias		The people who develop and use AI models can also introduce bias. This can happen if the developers are not aware of the potential for bias, or if they do not take steps to mitigate it.	TechnologyAdvice. (2022, May 18). Addressing AI and implicit bias in healthcare

Examining AI Biases: Insights from Scientific Studies

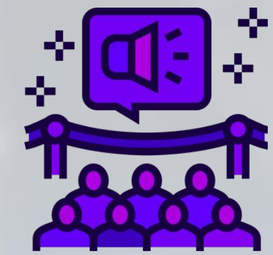
Study	Source	Key Takeaways
A global review of sources of bias in artificial intelligence that perpetuate healthcare disparities	PLOS Digital Health (2023)	U.S. and Chinese datasets and authors were disproportionately overrepresented in clinical AI.
The need for more diverse data to avoid perpetuating inequality in medicine	Scientific American (2023)	There is a need for more diverse data to be used in training AI algorithms to avoid perpetuating inequality in medicine.
Algorithmic bias as the application of an algorithm that compounds existing inequities in socioeconomic status, race, ethnic background, religion, gender, disability, or sexual orientation and amplifies inequities in health systems	Journal of Global Health (2022)	Algorithmic bias can compound existing inequities and amplify inequities in health systems.
How AI tools such as algorithms can unintentionally increase the impact of existing racial biases in medicine through the explicit use of race to predict outcomes and risk	NIHCM (2021)	AI tools such as algorithms can unintentionally increase the impact of existing racial biases in medicine.

Why It Matters?

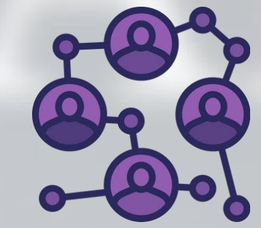
INACCURATE OR MISLEADING RESULTS



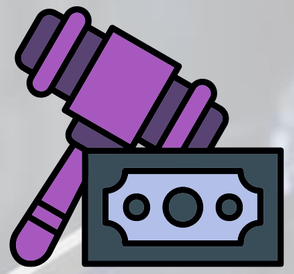
DAMAGE TO REPUTATION



STAKEHOLDERS



REGULATORY SCRUTINY



PATIENTS RIGHTS

LEGAL LIABILITY, FINES AND WARNING LETTERS

- Clinical Trials
- Fixing Ai-biased Errors in Drug Discovery
- Miss Diagnoses
- Drug Recalls
- Hurt Sales and Attracting Investors
- Governments and Regulators Concerned
- Impose Stricter Regulations & Fines
- Difficulty Developing New Drugs

Clinical Trials -Drug Discovery

AI-based clinical trials: be prepared to pay more for fixing errors and longer timelines

- Nearly 270 companies are active in the AI-driven drug discovery industry. More than 50% of these companies are based in the US. Critical hubs for this activity are emerging in Western Europe and Southeast Asia. Source: McKinsey & Company Oct, 2022

- A 2019 study by the Tufts Center for the Study of Drug Development found that a clinical trial error costs $180 million.
- The study also found that clinical trial errors can delay the development of a drug by an average of 1.8 years.

AI bias in drug discovery is a significant challenge that could hurt the bottom line.

How AI Bias is Leading to Lawsuits

Case	Country	Algorithm	Bias	Impact
Netherlands	Netherlands	Automated risk profiling system	Discriminatory	1.4 million people wrongly accused of fraud, lost access to benefits
UnitedHealthcare	USA	Algorithim to predict the risk of hospital readmission	More likely to predict that black patients would be re-admitted to hospital	Regulators says racial bias in algorithm leads to poorer care for black patients; UnitedHealth defends product.
Michigan	United States	Automated fraud detection system	High error rate	40,000 people wrongly accused of fraud, had to repay benefits
iTutorGroup	United States	Hiring algorithm	Age-based discrimination	EEOC sued iTutorGroup, settled for $365,000
State Farm	United States	Claim processing algorithm	Racial discrimination	Black homeowners face more delays, inquiries, and bias
Amazon	USA	AI tool to help recruiting	Bias agianst Women and selected male candidates	No law suit but the project was disbanded
Workday	USA	Ai tools to screen applicants	Ai tools discriminated based on race age and disability status	Class action action lawsuit on Feb 21/2023
Cigna	United States	PxDx system	Automated claim denials	80% of initial decisions overturned

JOHN HOPKINS REMOVING AI BIAS

Addressing Artificial Intelligence Bias in Retinal Diagnostics

Philippe Burlina, Neil Joshi, William Paul, Katia D. Pacheco, and Neil M. Bressler. "Addressing Artificial Intelligence Bias in Retinal Diagnostics." arXiv preprint arXiv:2004.13515 (2020).

Purpose	**To evaluate generative methods to mitigate AI bias in DR diagnostics**
Methods	The public domain Kaggle EyePACS dataset was modified to simulate data imbalance and domain generalization. A traditional/baseline diagnostic DLS was compared against new DLSs that would use training data augmented via generative models for debiasing.
Results	The baseline DLS had higher accuracy for lighter-skin individuals (73.0%) than darker-skin individuals (60.5%). The debiased DLS achieved closer parity in accuracy (72.0% for lighter-skin and 71.5% for darker-skin).
Conclusions	Data imbalance and domain generalization can lead to disparity of accuracy across subpopulations. Generative methods can be used to debias AI.
Translational Relevance	New AI methods have possible applications to address potential AI bias in DR diagnostics from fundus pigmentation, and potentially other ophthalmic DLSs too.

The importance of addressing AI bias in healthcare

Executive Summary

Key Takeaway	Description
AI Bias in DR Diagnostics	AI models can learn and reflect the biases present in the training data, which can lead to unfair or inaccurate diagnoses. This is a serious concern, as diabetic retinopathy is a leading cause of blindness, and early detection and treatment is essential.
Addressing AI Bias with GANs	Generative adversarial networks (GANs) can be used to generate synthetic data that is more diverse and representative of the real world, which can help to reduce AI bias. This is a promising approach that is still under development, but it has the potential to make AI-powered DR diagnostics more fair and accurate.
Debiased AI Models	Debiased AI models are designed to be fair and accurate, even when trained on data that is biased. This is an important finding, as it suggests that it is possible to develop AI models that are both accurate and fair.
Reducing AI Bias with Domain Generalization Techniques	Domain generalization techniques can help to reduce AI bias by making AI models more robust to changes in the data distribution. This is an important consideration, as the data used to train AI models is often not representative of the real world.
Fairness Considerations in DR AI Models	Fairness considerations should be taken into account when designing AI models for DR, in order to ensure that these models are not biased against certain groups of patients. This is a complex issue, but it is essential to address if we want to ensure that AI-powered DR diagnostics are used fairly and equitably.

AI RISK MANAGEMENT AND GOVERNANCE FRAMEWORKS

Framework	Main points
NIST Risk Framework (USA)	Defines five steps for AI risk assessment: identify, assess, mitigate, monitor, and accept.
AI4Health Risk Framework (UK)	Defines six steps for AI risk assessment: establish context, identify risks, assess risks, mitigate risks, monitor risks, and communicate risks.
ISO/IEC (International Organization for Standardization/International Electrotechnical Commission) 31000 Risk Management Standard	Defines four steps for risk management: establish context, identify risks, assess risks, and treat risks.
WEF (World Economic Forum)'s AI Risk Assessment Framework	Defines five key risks associated with AI: bias, security, privacy, safety, and accountability.
OECD's (Organization for Economic Co-operation and Development) AI Principles	Addresses a wide range of issues related to the development and use of AI, including human-centered design, transparency, accountability, robustness, safety, privacy, non-discrimination, inclusivity, and sustainability.

Useful

Safe

Accountable and Transparent

Explainable and Interpretable

Fair

With Harmful Bias Managed

Secure and Resilient

Privacy-Enhanced

CHAI™
Coalition for Health AI

https://www.coalitionforhealthai.org

Thank You

Aisha Rahim MD, FHM

ar@globalboardadvisors.com

BOARDROOM EDUCATION