Inherent Racism Induces Bias in Algorithm Development: Can Implementation of Genetic Information in Clinical Informatics Decision Be Different?

Janina M. Jeff, PhD, MS
Population Geneticist & Host of In Those Genes Podcast
Genomic Medicine XIII: Developing a Clinical Genomic Informatics Research
Email: janina@inthosegenes.com
Twitter: @inthosegenespod
Disclosures

This presentation and thoughts expressed do not reflect nor are the beliefs of Illumina. They are my own as an individual.

I am not a person with formal D&I training or expertise. I am a Black woman with a complex identity between race and science through lived experiences.
How “Race” Influences Algorithm Development

RISK OF BEING REARRESTED

- training inputs might correlate with being blue or purple
- purple individuals are more likely to be ‘false positives’: misidentified as high risk
- the purple group might have been unfairly targeted for arrest in the first place
- accurately predicting from past data that more people in the purple group will be rearrested, the algorithm could be recapitulating — and perhaps entrenching — a pre-existing societal bias

Inherently Racist Algorithms
- Criminal Justice System
- Beauty Contests
- Zoom Virtual Backgrounds
- Google Maps

How “Race” Impacts Clinical Informatics

Ziad Obermeyer et al. Science 2019;366:447-453
“How The World Inherited Blackness”
Season 2, Episode 1
Delande Justinvil (biocultural anthropologist)
Saida Grundy, PhD(Sociologist)
“Race” & Computational Genetics

Karl Pearson and The Annals of Eugenics

Galton and Regression to the Mean
Impact of “Race” in Clinical Implementation of Genetic Information

- Clinical Annotations of genetic variants are not transferable across populations

- 11.5% of genetic disorders with inflation in pathogenic variant sets and up to 92.3% for the variant set with conflicting interpretations

Can Implementation of Genetic Information in Clinical Informatics Decision Be Different?

Ziad Obermeyer et al. Science 2019;366:447-453
Can Implementation of Genetic Information in Clinical Informatics Decision Be Different?

Maybe.
Inflation increases as the level of evidence supporting the pathogenic nature of the variant decreases.

- Supporting evidence is shaped by representation of genetic studies which is a manifestation due to lack of trust in diverse communities or “racism”

- Burden of rare variants was a main contributing factor of the observed inflation, indicating collective misclassified rare variants.
  - Rare variants are population specific.

Adjusting for disease-specific population allele frequency improves inflation but does not fix it.
Can Implementation of Genetic Information in Clinical Informatics Decision Be Different?

Yes, but are we ready to be honest about how we got here and are we ready for the discomfort it will take to create it?

“today coded inequity is perpetuated precisely because those who design and adopt such tools are not thinking carefully about systemic racism”
“having this (eugenics) contemptible field at its foundations might prevent genetics ever being the anti-racist ally”
Janina M. Jeff, PhD, MS
Email: janina@inthosegenes.com
Twitter: @inthosegenespod
website: inthosegenes.com

Listen now
www.inthosegenes.com/Listen-1