The intended or unintended consequences of including race in clinical algorithms and how we can use AI to improve, not harm, health for people from all racial and ethnic backgrounds.

National Institute for Health Care Management Foundation

Fay Cobb Payton, PhD, MBA
Professor/University Faculty Scholar
North Carolina State University
September 28, 2021
Disclaimer

Any opinions, findings, conclusions or recommendations presented in these materials are only my own and do not necessarily reflect the views of the National Science Foundation, North Carolina State University or any other affiliation.
Research Findings
WE GOT DATA!!!

LET'S USE IT!!

U GET DATA, YOU GET I

DATA FOR EVERYON
Diabetes and the hospitalized patient

A cluster analytic framework for characterizing the role of sex, race and comorbidity from 2006 to 2011

Nisha Nataraj¹  •  Julie Simmons Ivy¹ • Fay Cobb Payton² • Joseph Norman³
Fig. 2  Overview of sample selection process. Records in grey boxes have been removed from the analysis.

47,911,414 Discharge Records between 2006-2011

35,721,265 Discharges over 18, not admitted for newborn delivery

28,732,666 Discharges with missing fields excluded

- 7,792,398 Discharges under 18 years old
- 4,397,751 Discharges with newborn delivery on first five procedure codes (PRCCS 1-5)

- 6,988,599 Discharges with missing values
  (missing values excluded for following fields: age, gender, race, total charges, length of stay and discharge disposition)
Fig. 1  Schematic representation of relationship between variables and outcomes of interest. The population of interest separated by sex and diabetes prevalence is represented in the inner-most circle. The primary variable of interest, comorbidity, is represented in the light gray section. Variables in the following dark gray section represent demographic variables while those in the black section represent primary outcomes of interest.
• Principal reason for the hospitalization: the disparities in outcomes for women and ethnic groups are persistent

• Ethnic groups have poorer outcomes and are less likely to have routine discharges.

• These disparities have persisted over time suggesting that without conscious effort to personalize care for women and diverse groups with diabetes, their outcomes are unlikely to improve.

• Differential impact of diabetes on physiology and treatment in women versus men.

• Homogeneity in patients and comorbidities is rare given the variability in demographic characteristics
Characterizing the impact of mental disorders on HIV patient length of stay and total charges

SHENGFAN ZHANG¹,*, FAY COBB PAYTON² and JULIE SIMMONS IVY³

¹Department of Industrial Engineering, University of Arkansas, Fayetteville, AR, USA
E-mail: shengfan@uark.edu
²Information Systems, College of Management, North Carolina State University, Raleigh, NC, USA
³Edward P. Fitts Department of Industrial and Systems Engineering, North Carolina State University, Raleigh, NC, USA
Results for principal component analysis

<table>
<thead>
<tr>
<th>ICD-9 Codes Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nondependent abuse of drugs</td>
</tr>
<tr>
<td>Disturbance of emotions specific to childhood and adolescents</td>
</tr>
<tr>
<td>Specific delays in development</td>
</tr>
<tr>
<td>Psychic factors associated with diseases classified elsewhere</td>
</tr>
<tr>
<td>Other specified mental retardation</td>
</tr>
<tr>
<td>Unspecified mental retardation</td>
</tr>
<tr>
<td>Drug-induced mental disorders</td>
</tr>
<tr>
<td>Transient mental disorders due to conditions classified els</td>
</tr>
<tr>
<td>Drug dependence</td>
</tr>
<tr>
<td>Depressive disorder not elsewhere classified</td>
</tr>
<tr>
<td>Episodic mood disorders</td>
</tr>
<tr>
<td>Personality disorders</td>
</tr>
<tr>
<td>Acute reaction to stress</td>
</tr>
<tr>
<td>Adjustment reaction</td>
</tr>
<tr>
<td>Specific nonpsychotic mental disorders due to brain dam;</td>
</tr>
<tr>
<td>Hyperkinetic syndrome of childhood</td>
</tr>
<tr>
<td>Mild mental retardation</td>
</tr>
<tr>
<td>Dementias</td>
</tr>
<tr>
<td>Persistent mental disorders due to conditions classified els</td>
</tr>
<tr>
<td>Special symptoms or syndromes not elsewhere classified</td>
</tr>
<tr>
<td>Delusional disorders</td>
</tr>
<tr>
<td>Pervasive developmental disorders</td>
</tr>
<tr>
<td>Anxiety, dissociative and somatoform disorders</td>
</tr>
<tr>
<td>Sexual and gender identity disorders</td>
</tr>
<tr>
<td>Physiological malfunction arising from mental factors</td>
</tr>
<tr>
<td>Disturbance of conduct not elsewhere classified</td>
</tr>
<tr>
<td>Alcohol-induced mental disorders</td>
</tr>
<tr>
<td>Alcohol dependence syndrome</td>
</tr>
<tr>
<td>Schizophrenic disorders</td>
</tr>
<tr>
<td>Drug-induced mental disorders</td>
</tr>
<tr>
<td>Other nonorganic psychoses</td>
</tr>
</tbody>
</table>

In each cluster are highlighted in bold.
• Longer stay lengths, comorbid conditions affected HIV patient outcomes.

• Mental disorders generally results in a decrease in both LOS and total charges.

• Patients with mental illness are more likely to be transferred to other facilities so that their true LOS will not be observed.

• The most important conditions were drug related mental disorders (304, 305), mood disorders (296), depression (311) and anxiety (300).

• Health services delivery approach to adherence and treatment to better address chronic diseases and their severity with comorbidity.
Cluster 1: Age

Age factors in mental health experiences

Document Frequency: 7

Percentage of Sample: 4%

Cluster 2: Race

Racial factors in mental health experiences

Document Frequency: 30

Percentage of Sample: 18%

Cluster 3: Crime

Factors related to crime that lead to mental health experiences

Document Frequency: 3

Percentage of Sample: 2%

Cluster 4: Services

What institutions are doing to assist with mental health experiences

Document Frequency: 113

Percentage of Sample: 68%

Cluster 5: Aftermath

What happens after a mental health experience

Document Frequency: 4

Percentage of Sample: 2%

Cluster 6: Victim

Factors related to victimization that lead to mental health experiences

Document Frequency: 8

Percentage of Sample: 5%
When Race (Inclusion) is Excluded

- Race-blinded algorithms may be well-intended
- Race-blinded \(\rightarrow\) elevate disparate conditions/experiences/health outcomes
- Race-blinded \(\rightarrow\) elevate structural inequities
- Race-blinded \(\rightarrow\) health costs as a proxy for health needs
  - Fewer resources spent on Black patients who have the same level of need
  - Algorithm falsely concludes that Black patients are healthier than equally sick White patients (Obermeyer, et al 2019 in *Science*)
So, What Was/Is Missing?
The Healthy People 2020 Social Determinants of Health topic area is organized into 5 place-based domains:

1. Economic Stability
2. Education
3. Health and Health Care
4. Neighborhood and Built Environment
5. Social and Community Context

Discrimination is a key issue in the Social and Community Context domain.
‘Big data’ was supposed to fix education. It didn’t. It’s time for ‘small data.’ - The Washington Post
Here Comes the #Engagement: A serious health initiative made trendy

Creating a user experience to communicate the seriousness of HIV prevention and awareness can be both educational while entertaining. This combination along with a sense of cultural influence helps to both attract and engage millennials.

By Fay Cobb Payton and KinMar Galloway
DOI: 10.1143/2691362
Suicide is the 3rd leading cause of death on college campuses.

Mood disorders such as depression are the 3rd most common cause of hospitalization in the U.S. for both youth and adults ages 18 to 44.

The death rate from suicide for African American men was almost 4x that for African American women, in 2009.

African Americans are 20% more likely to report having serious psychological distress than Non-Hispanic Whites.

75% of lifetime cases of mental health conditions begin at age 24.

#Activist UX
MyHealthImpactNetwork
Historical Context

• Think about the time period of the story and things like technology, events, or issues related to this particular time frame.

• How could the author be using the historical context to impact the plot?

• How is he using characters, events to make a point to the reader?
Parting Thoughts on It’s Not Just A Data Issue
Algorithmic Equity in the Hiring of Underrepresented IT Job Candidates

Authors
Lynette Yarger, Penn State
Fay Cobb Payton, NC State
Bikalpa Neupane, Penn State

- Dataset bias: data used to train machine learning models does not represent the diversity of the customer base (e.g., Voice recognition technologies that only work well for male users because the initial training data excluded women.)
- Association bias: data used to train a model reinforces and multiplies a cultural bias (e.g., Language translation tools that make gender assumptions.)
- Automation bias: automated decisions override social and cultural considerations (e.g., Beautification photo filters reinforcing a European notion of beauty on facial images.)
- Interaction bias: humans interact with AI and create biased results (e.g., Humans deliberately input sexist language into a chatbot to train it to say offensive things.)
- Confirmation bias: overly simplified personalization makes biased assumptions for a group or an individual (e.g., Job advertisements for executive positions are displayed only to male users.)
‘Health equity tourists’: How white scholars are colonizing research on health disparities

By Usha Lee McFarling  Sept. 23, 2021 | Reprints

A glaring example occurred in August when the Journal of the American Medical Association — a leading medical journal already under fire for how it handles issues of race — published a special themed issue on racial and ethnic health disparities in medicine. Meant to highlight JAMA’s new commitment to health equity, it served up an illustration of the structural racism embedded in academic publishing: Not one of the five research papers published in the issue included a Black lead or corresponding author, and just one lead author was Hispanic.
Data Science Inclusion

Health Disparities (De)Colonization

Race/Ethnicity of General Assembly Students by Course

<table>
<thead>
<tr>
<th>Course</th>
<th>White</th>
<th>Asian (S/E/SE)</th>
<th>Hispanic or Latino</th>
<th>African-American</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVERAGE</td>
<td>46.9%</td>
<td>22%</td>
<td>9.8%</td>
<td>7.7%</td>
<td></td>
</tr>
<tr>
<td>Product Management</td>
<td>50.5%</td>
<td>21.4%</td>
<td>10.1%</td>
<td>5.7%</td>
<td></td>
</tr>
<tr>
<td>Digital Marketing</td>
<td>49.8%</td>
<td>15.9%</td>
<td>12.2%</td>
<td>8.1%</td>
<td></td>
</tr>
<tr>
<td>Data Science</td>
<td>46.1%</td>
<td>28%</td>
<td>7.8%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>JavaScript Development</td>
<td>46.1%</td>
<td>19.7%</td>
<td>9.5%</td>
<td>7.7%</td>
<td></td>
</tr>
<tr>
<td>Front-End Web Development</td>
<td>45.5%</td>
<td>20.8%</td>
<td>11.0%</td>
<td>10.9%</td>
<td></td>
</tr>
<tr>
<td>User Experience Design</td>
<td>44.4%</td>
<td>24.5%</td>
<td>9.0%</td>
<td>6.9%</td>
<td></td>
</tr>
<tr>
<td>Data Analytics</td>
<td>43.8%</td>
<td>24.6%</td>
<td>9.3%</td>
<td>9.8%</td>
<td></td>
</tr>
</tbody>
</table>

Source: General Assembly part-time student data (09/2016-01/2017)
"Asian" represents South, Southeast, and East Asia
*Average = the courses listed above

Course enrollment by demographic. GENERAL ASSEMBLY

Data source: General Assembly
The Lived Experience & Algorithmic Biases - Fairness

Small Data → Social Determinants of Health → Place and Space

Design Justice: Centering Race in Design

Ecosystem Thinking → Data Sanitation → Context Matters
- Augmentation of AI Findings/Results Interpretation

Overfitting vs Underfitting Has Implications.
Fay Cobb Payton

fcpayton@ncsu.edu