# FINGER LAKES HEALTH DATA **CONFERENCE** Presented by the ROCHESTER RHIO



#### Keynote Speaker: Joshua Vest, Ph. D., M.P.H.



Associate Professor of Health Policy & Management and director of the Center for Health Policy, Indiana University

Affiliated Scientist, Regenstrief Institute, Inc.



# Applications of data in health care

Joshua R Vest, PhD, MPH Indiana University Richard M. Fairbanks School of Public Health Regenstrief Institute

## HEALTH CARE WILL GENERATE 2,134,000,000 TERABYTES OF DATA IN 2020



#### CLINICAL



Progress notes Vital signs Medical histories Diagnoses Medications Immunizations Allergies Radiology images Lab and test results

# CLINICALLEGAL & IDENTITYImage: constant of the second sec

Medical histories Diagnoses Medications Immunizations Allergies Radiology images Lab and test results Name Contact information DNR Advance Directives Gender Race/ethnicity





Lab and test results



## **HOW CAN ALL THAT DATA BE APPLIED?**

#### PATIENT CARE & POPULATION HEALTH

MANAGING ORGANIZATIONS



#### **PUBLIC HEALTH & POLICY**

**RESEARCH & DEVELOPMENT** 



https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5051581/pdf/egems1244.pdf

## THE APPLICATIONS OF HEALTH CARE DATA

#### PATIENT CARE & POPULATION HEALTH

MANAGING ORGANIZATIONS



#### PUBLIC HEALTH & POLICY

**RESEARCH & DEVELOPMENT** 



#### OBVIOUS - HEALTH DATA SHOULD SUPPORT MEDICAL CARE & DECISION MAKING. BUT DATA ARE OFTEN MISSING.





## AT A POPULATION LEVEL QUALITY DATA ARE NECESSARY TO RISK STRATIFY

The concept that patients' posses different underlying severity of disease or different probabilities of negative outcomes





Of the total population

Of all out-of-pocket health care costs

- 1% of the population accounted for nearly a fifth of all out-of-pocket spending on health services in 2016
- Top 5% of spenders accounted for 46%.
- the 50% of the population with the lowest out-ofpocket spending accounted for 2% of all out-ofpocket health spending.



https://www.ahrq.gov/sites/default/files/wysiwyg/professionals/prevention-chronic-care/decision/mcc/mccchartbook.pdf

#### lezzoni 1997

https://www.healthsystemtracker.org/chart-collection/health-expenditures-vary-across-population/#item-out-of-pocket-spending-on-health-services-is-almost-as-concentrated-as-overall-health-spending\_2016



## **RISK STRATIFICATION**

The goal of risk stratification is to segment patients into distinct groups of similar complexity and care needs.



Also referred to as market segmentation.

Fundamentally an identification & allocation strategy.

Systematically find those patients in the most need of services.

Allows organization to focus resources.

## **RISK STRATIFICATION FAILS WITH POOR QUALITY DATA**

#### **RESEARCH ARTICLE**

# Dissecting racial bias in an algorithm used to manage the health of populations

**Ziad Obermeyer**<sup>1,2,\*</sup>, **Brian Powers**<sup>3</sup>, **Christine Vogeli**<sup>4</sup>, **Sendhil Mullainathan**<sup>5,\*,†</sup> + See all authors and affiliations

Science 25 Oct 2019: Vol. 366, Issue 6464, pp. 447-453 DOI: 10.1126/science.aax2342

#### Abstract

Health systems rely on commercial prediction algorithms to identify and help patients with complex health needs. We show that a widely used algorithm, typical of this industry-wide approach and affecting millions of patients, exhibits significant racial bias: At a given risk score, Black patients are considerably sicker than White patients, as evidenced by signs of uncontrolled illnesses. Remedying this disparity would increase the percentage of Black patients receiving additional help from 17.7 to 46.5%. The bias arises because the algorithm predicts health care costs rather than illness, but unequal access to care means that we spend less money caring for Black patients than for White patients. Thus, despite health care cost appearing to be an effective proxy for health by some measures of predictive accuracy, large racial biases arise. We suggest that the choice of convenient, seemingly effective proxies for ground truth can be an important source of algorithmic bias in many contexts.

## **PRISK STRATIFICATION FAILS WITH POOR QUALITY DATA**

#### **RESEARCH ARTICLE**

# Dissecting racial bias in an algorithm used to manage the health of populations

**Ziad Obermeyer**<sup>1,2,\*</sup>, **Brian Powers**<sup>3</sup>, **Christine Vogeli**<sup>4</sup>, **Sendhil Mullainathan**<sup>5,\*,†</sup> + See all authors and affiliations

Science 25 Oct 2019: Vol. 366, Issue 6464, pp. 447-453 DOI: 10.1126/science.aax2342

#### Abstract

Health systems rely on commercial prediction algorithms to identify and help patients with complex health needs. We show that a widely used algorithm, typical of this industry-wide approach and affecting millions of patients, exhibits significant racial bias: At a given risk score, Black patients are considerably sicker than White patients, as evidenced by signs of uncontrolled illnesses. Remedying this disparity would increase the percentage of Black patients receiving additional help from 17.7 to 46.5%. The bias arises because the algorithm predicts health care costs rather than illness, but unequal access to care means that we spend less money caring for Black patients than for White patients. Thus, despite health care cost appearing to be an effective proxy for health by some measures of predictive accuracy, large racial biases arise. We suggest that the choice of convenient, seemingly effective proxies for ground truth can be an important source of algorithmic bias in many contexts.

## THE APPLICATIONS OF HEALTH CARE DATA

#### PATIENT CARE & POPULATION HEALTH

MANAGING ORGANIZATIONS



#### **PUBLIC HEALTH & POLICY**

**RESEARCH & DEVELOPMENT** 



## MANAGING ORGANIZATIONS

Using EHR data to draw inferences about your organization's performance





Performance improvement

Workflows & t

Workflows & time allocations



## **MEASURES FROM EHR LOGS**



https://academic.oup.com/jamia/article/27/4/639/5728718

## **CURRENTLY EHRs HAVE CHALLENGES**



#### EHR Data Reports Fall Short for Reporting Performance Indicators

Using MRA and EHR reports can be useful for reporting some performance indicators, but the tools are vastly inaccurate for reporting aspirin-related performance measures.



The EHR had 72 different ways to document "aspirin".

Only 59% of practices were able to generate rolling 12-month measures on preventative cardiovascular care.



The Joint Commission Journal on Quality and Patient Safety 2020; 46:11–17

#### Using Electronic Health Records to Measure Quality Improvement Efforts: Findings from a Large Practice Facilitation Initiative

David T. Liss, PhD; Yaw A. Peprah, MPH; Tiffany Brown, MPH; Jody D. Ciolino, PhD; Kathryn Jackson, MS; Abel N. Kho, MD, MS; Linda Murakami, RN MSHA; Theresa L. Walunas, PhD; Stephen D. Persell, MD, MPH

ehrintelligence.com/news/ehr-data-reports-fall-short-for-reporting-performance-indicators https://www.jointcommissionjournal.com/action/showPdf?pii=S1553-7250%2819%2930394-0

## THE APPLICATIONS OF HEALTH CARE DATA



#### THE VALUE OF QUALITY DATA TO PUBLIC POLICY & PUBLIC HEALTH



stats.oecd.org/Index.aspx?DataSetCode=SHA#

Chetty et al JAMA 2016



#### THE VALUE OF QUALITY DATA TO **PUBLIC POLICY & PUBLIC HEALTH**

Unique

County

© Mapbox © OSM

Patients

#### AGGREGATED HEALTH CARE DATA



Age Group

0-4

5-19

20-29

20.20

Male

Female

397

106

Total

49

126

606

002

26

48

209

207

23

78





AGGREGATING EHR DATA TO GENERATE POPULATION LEVEL ESTIMATES OF HEALTH INDICATORS

Table 1. Prevalence of Obesity, Smoking, Depression and Influenza Vaccination among Adults in Care, New York City, 2013

OUTCOME	2013 NYC MACROSCOPEª % (95% CI)	2013–2014 NYC HANES⁵ % (95% CI)	2013 NYC CHS <sup>c</sup> % (95% CI)
Obesity	27.8 (27.7-27.9)	31.3 (28.5-34.2)	24.7 (23.2-26.3)
Smoking	15.2 (15.1-15.3)	17.7 <b>(</b> 15.1–20.8)	14.9 (13.6-16.3)
Depression Self-Report (SR)⁴	8.2 (8.1-8.2) n/a	19.0 (16.4-21.9) 15.2 (13.0-17.7)	n/a 16.4 (15.1-7.9)
Influenza Vaccination	20.9 (20.8-21.0)	47.6 (44.0-51.3)	47.3 (45.5-49.0)

Notes: <sup>a</sup> Weighted to the NYC HANES distribution of the population in care.

<sup>b</sup> New York City Health and Nutrition Examination Survey.

<sup>c</sup> New York City Community Health Survey.

<sup>d</sup> Alternate definition: Self-reported diagnosis.

www.ncbi.nlm.nih.gov/pmc/articles/PMC5226379/pdf/egems1267.pdf



## WHY SHOULD A HEALTH CARE ORGANIZATION CARE ABOUT USING "THEIR" DATA FOR PUBLIC POLICY & PUBLIC HEALTH?

# NEEDED





## THE ABSENCE OF QUALITY DATA FOR PUBLIC POLICY & PUBLIC HEALTH IS DAMAGING.



## **Right? Georgia, Florida And The Deadly Trend Of Science Suppression**

## POLICY AND PUBLIC HEALTH USE CASES REQUIRE ATTENTION TO THE QUALITY OF NON-CLINICAL DATA



www.kff.org/disparities-policy/issue-brief/disparities-in-health-and-health-care-five-key-questions-and-answers/









#### **RACE & ETHNICITY ARE FREQUENTLY NOT RECORDED**



## THE APPLICATIONS OF HEALTH CARE DATA

#### PATIENT CARE & POPULATION HEALTH

MANAGING ORGANIZATIONS



#### PUBLIC HEALTH & POLICY

**RESEARCH & DEVELOPMENT** 

R&D a process intended to create new or improved technology that can provide a competitive advantage at the business, industry, or national level.

www.inc.com/encyclopedia/research-and-development.html

O o





#### BASIC

APPLIED

DEVELOPMENT











#### **APPLICATION OF EHR DATA IN ACADEMIC RESEARCH**

Academic papers each year on electronic health record data (in 100,000s)

#### Results per 100,000 citations in PubMed

proportion for each search by year, 2010 to 2020



#### Establish Research Cohorts

#### **Reevaluating Prior Findings**

Rare Outcomes

#### Research on Stigmatized Conditions

esperr.github.io/pubmed-by-year/?q1="electronic health record data"&startyear=2010 https://www.annualreviews.org/doi/10.1146/annurev-publhealth-032315-021353

#### VENTURE CAPITAL FUNDING FOR DIGITAL HEALTH INCREASED BY 858%



www.healthaffairs.org/do/10.1377/hblog20181218.956406/full/



## WHAT TECHNICAL INNOVATIONS PROMISE





## WHAT TECHNICAL INNOVATIONS PROMISE

THE RISK WITH POOR DATA QUALITY



No amount of magic can fixt this problem.



xkcd.com/1838/



https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5051581/pdf/egems1244.pdf

## SUMMARY: WHY DO CARE ABOUT QUALITY DATA?

ို

#### PATIENT CARE & POPULATION HEALTH

ATIONS

#### **PUBLIC HEALTH & POLICY**

**RESEARCH & DEVELOPMENT** 

MANAGING ORGANIZATIONS

# Applications of data in health care

Joshua R Vest, PhD, MPH Indiana University Richard M. Fairbanks School of Public Health Regenstrief Institute