ML FOR PUBLIC HEALTH ANALYSIS

Agenda

- Introduction
  - What is public health?
  - What affect public health?
  - What are challenges?
  - How can we do to promote public health?
- An empirical case
  - Feng and Jiao (2021). Predicting and mapping neighborhood-scale health outcomes: A machine learning approach
- Lab time

ML4HEALTH

Figure source: https://www.broadinstitute.org/ml4h

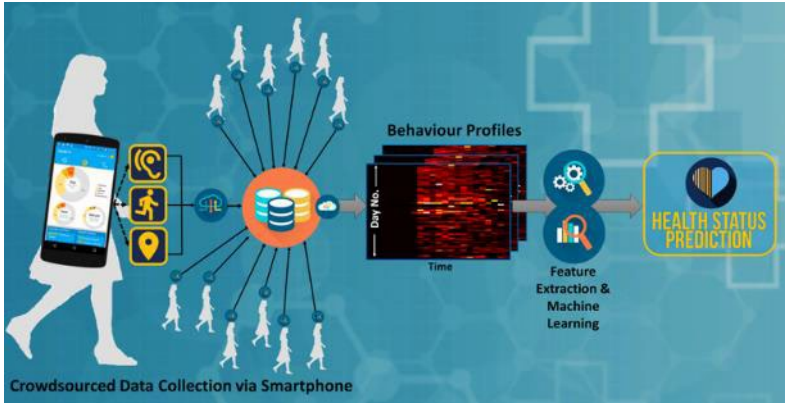


Figure source: https://www.embs.org/jbhi/articles/public-health-informatics-vol21-issue6/

Introduction

- Public Health is defined as **“the art and science of preventing disease, prolonging life and promoting health through the organized efforts of society”** (Acheson, 1988; WHO).

Source: https://medium.com/@faizan81/what-are-the-implications-of-each-of-the-four-components-of-public-health-26a1e6c198cc

Introduction

- How can we define health?

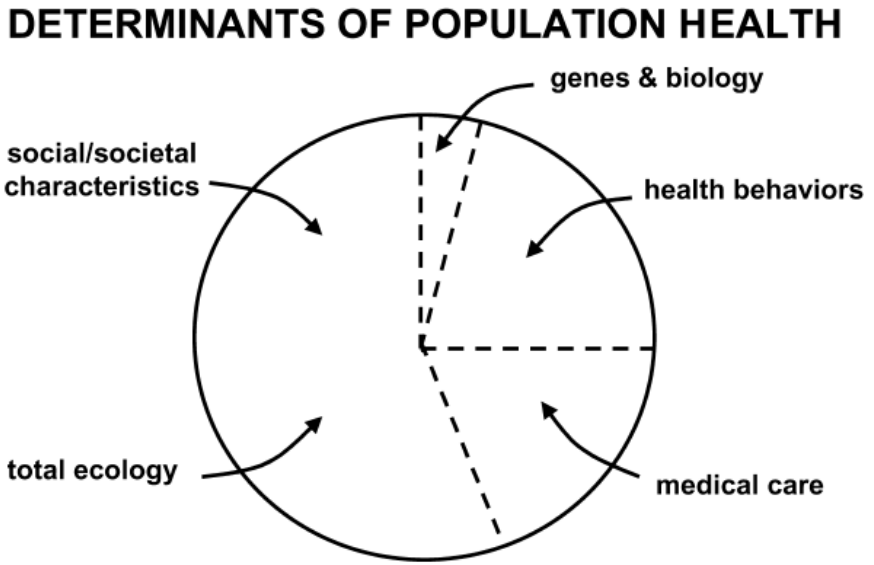


FIGURE 1. Relative influence of the five major determinant categories of population health: rough approximations.

(Tarlov, 2006)

Introduction

- About Social Determinants of Health (SDOH)

About Social Determinants of Health (SDOH)

What are social determinants of health?

Social determinants of health (SDOH) are conditions in the places where people live, learn, work, and play that affect a wide range of health risks and outcomes.

Healthy People 2030 uses a place-based framework that outlines five key areas of SDOH:

Healthcare Access and Quality

The connection between people's access to and understanding of health services and their own health. This domain includes key issues such as access to healthcare, access to primary care, health insurance coverage, and health literacy.

Education Access and Quality

The connection of education to health and wellbeing. This domain includes key issues such as graduating from high school, enrollment in higher education, educational attainment in general, language and literacy, and early childhood education and development.

Social and Community Context

The connection between characteristics of the contexts within which people live, learn, work, and play, and their health and wellbeing. This includes topics like cohesion within a community, civic participation, discrimination, conditions in the workplace, and incarceration.

Economic Stability

The connection between the financial resources people have – income, cost of living, and socioeconomic status – and their health. This area includes key issues such as poverty, employment, food security, and housing stability.

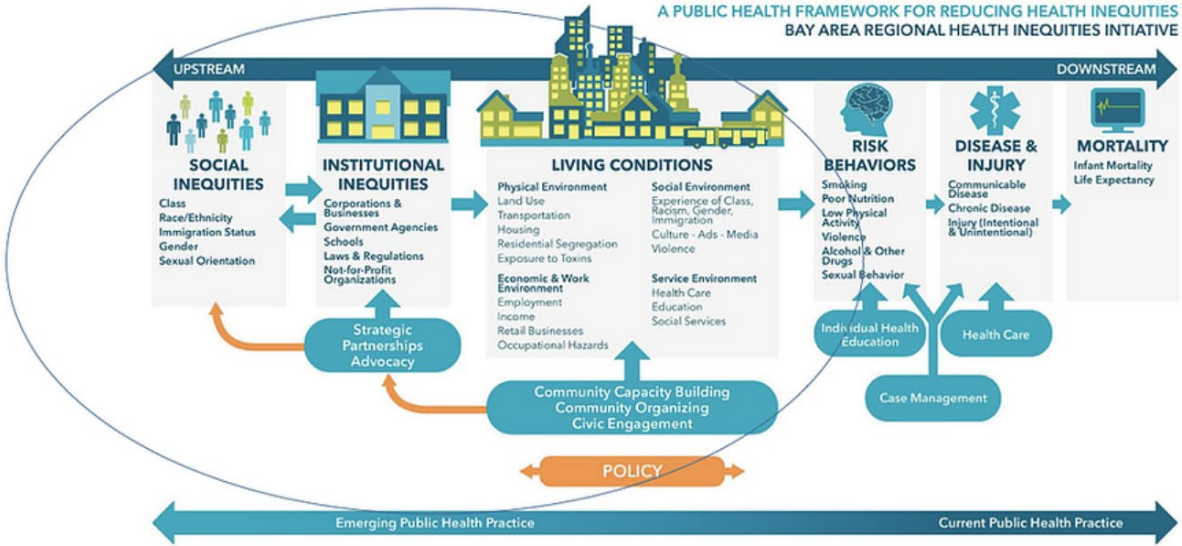
Neighborhood and Built Environment

The connection between where a person lives – housing, neighborhood, and environment – and their health and wellbeing. This includes topics like quality of housing, access to transportation, availability of healthy foods, air and water quality, and neighborhood crime and violence.

Source: https://www.cdc.gov/socialdeterminants/about.html

Introduction

- How can we do to promote health?



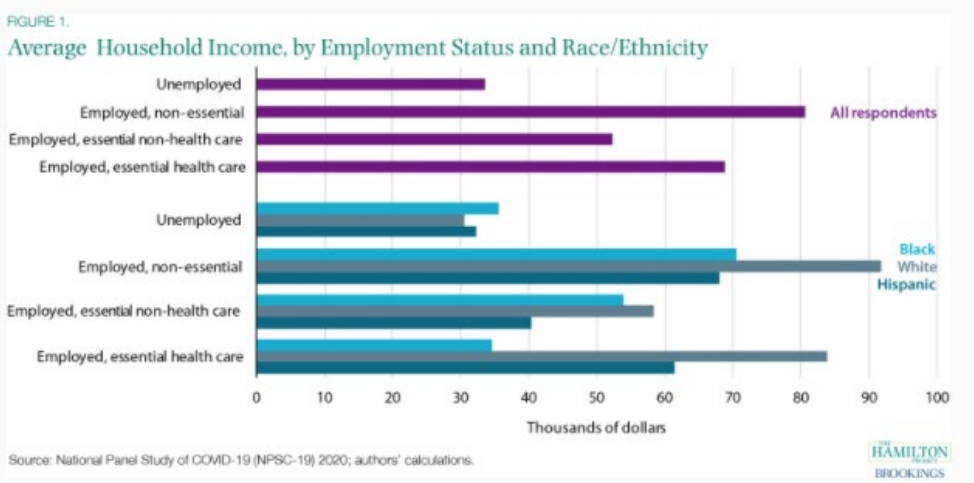
Source: https://www.barhii.org/enlarged-framework

Introduction

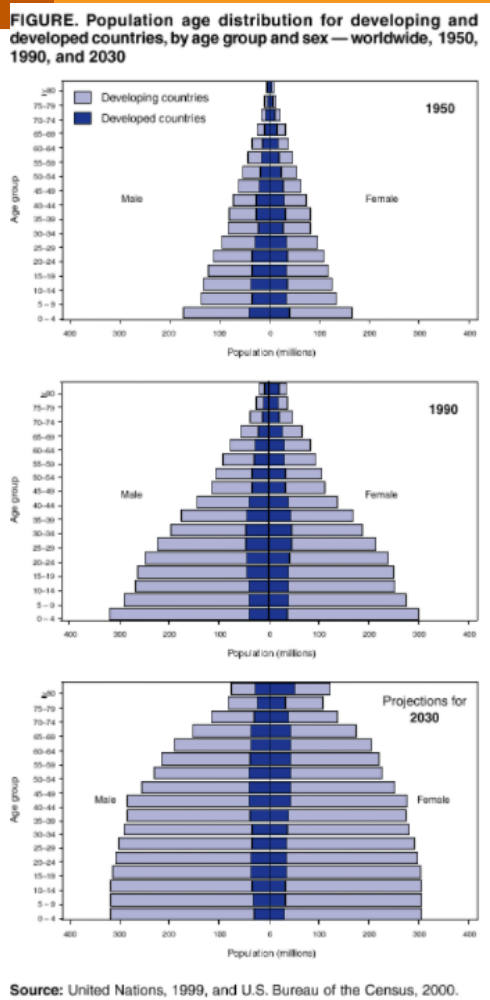
- Challenges of public health:
  - economic crisis;
  - ageing population;
  - increasing levels of chronic disease;
  - migration and urbanization; and
  - environmental damage and climate change

Introduction

- Challenges of public health: economic crisis



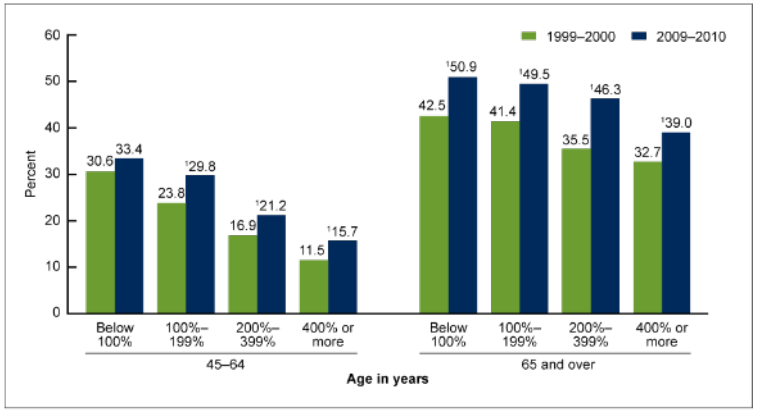
- Challenges of public health: ageing population



Introduction

- Challenges of public health: increasing levels of chronic disease

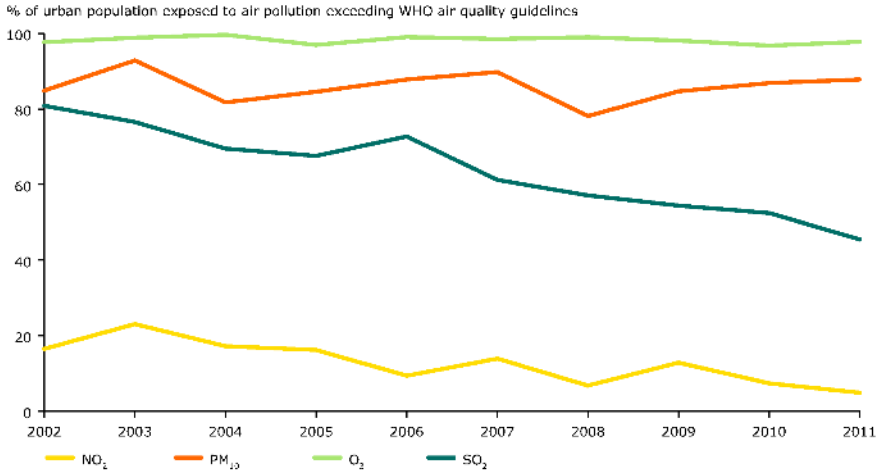
Figure 3. Prevalence of two or more of nine selected chronic conditions among adults aged 45 and over, by age and percentage of poverty level: United States, 1999–2000 and 2009–2010



†Significantly different from 1999–2000, $p < 0.05$.
NOTE: Access data table for Figure 3 at: http://www.cdc.gov/nchs/data/databriefs/db100_tables.pdf#3.
SOURCE: CDC/NCHS, National Health Interview Survey.

Source: https://www.cdc.gov/nchs/products/databriefs/db100.htm

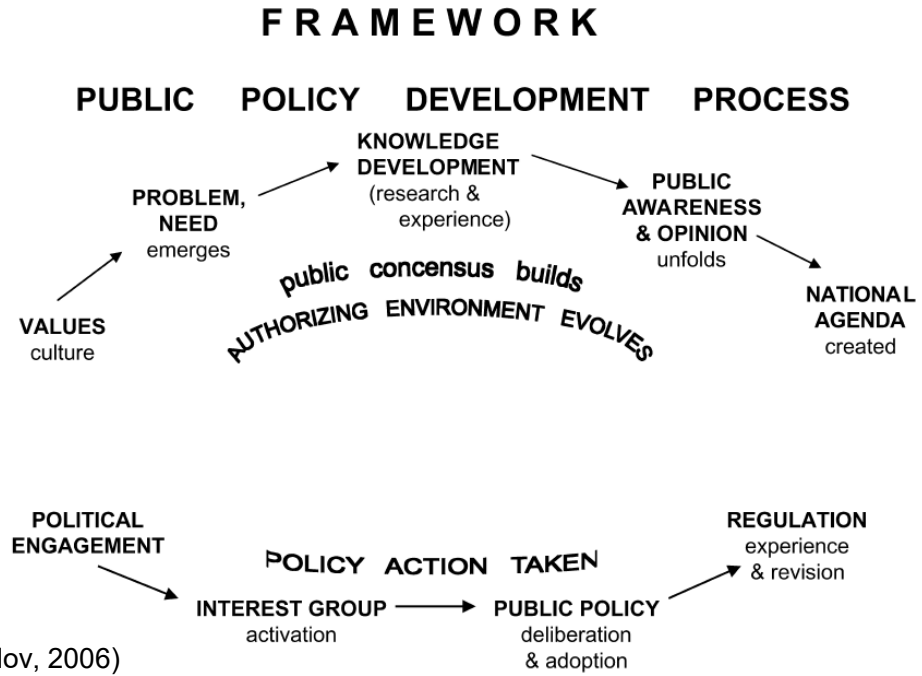
- Challenges of public health: migration and urbanization and environmental damage and climate change



Source: https://www.eea.europa.eu/data-and-maps/figures/percentage-of-the-eu-urban

Introduction

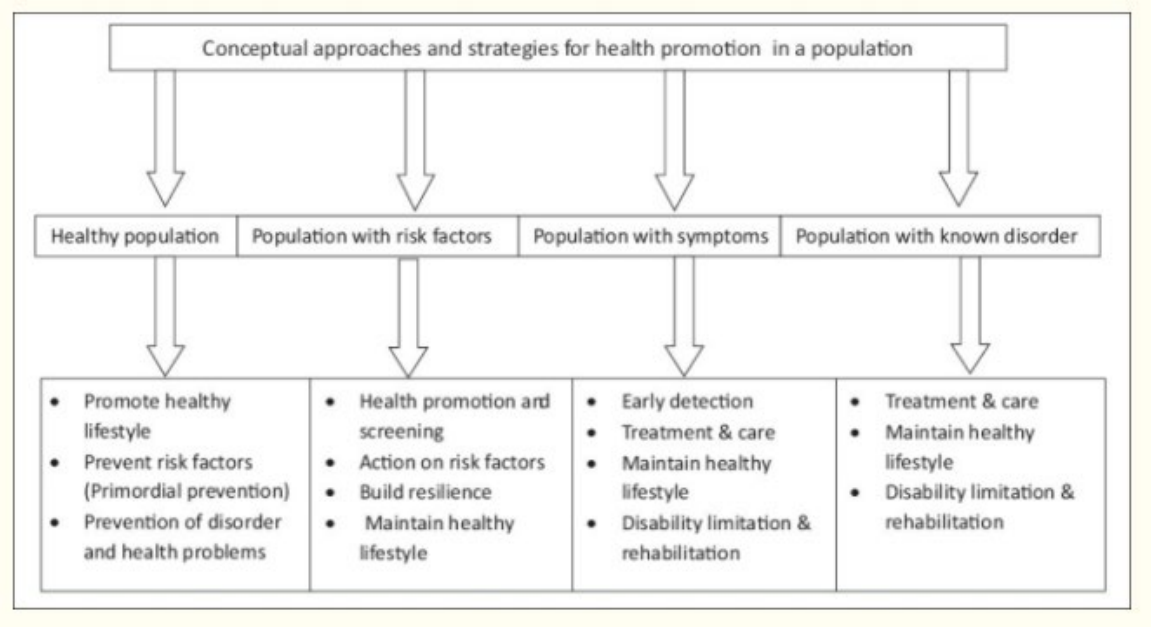
- How can we do?



(Tarlov, 2006)

Introduction

- How can we do?



(Kumar and Preetha, 2012)

Introduction

- How can we do?

Interventions in BE?

About Social Determinants of Health (SDOH)

What are social determinants of health?

Social determinants of health (SDOH) are conditions in the places where people live, learn, work, and play that affect a wide range of health risks and outcomes.

Healthy People 2030 uses a place-based framework that outlines five key areas of SDOH:

Healthcare Access and Quality

The connection between people's access to and understanding of health services and their own health. This domain includes key issues such as access to healthcare, access to primary care, health insurance coverage, and health literacy.

Education Access and Quality

The connection of education to health and wellbeing. This domain includes key issues such as graduating from high school, enrollment in higher education, educational attainment in general, language and literacy, and early childhood education and development.

Social and Community Context

The connection between characteristics of the contexts within which people live, learn, work, and play, and their health and wellbeing. This includes topics like cohesion within a community, civic participation, discrimination, conditions in the workplace, and incarceration.

Economic Stability

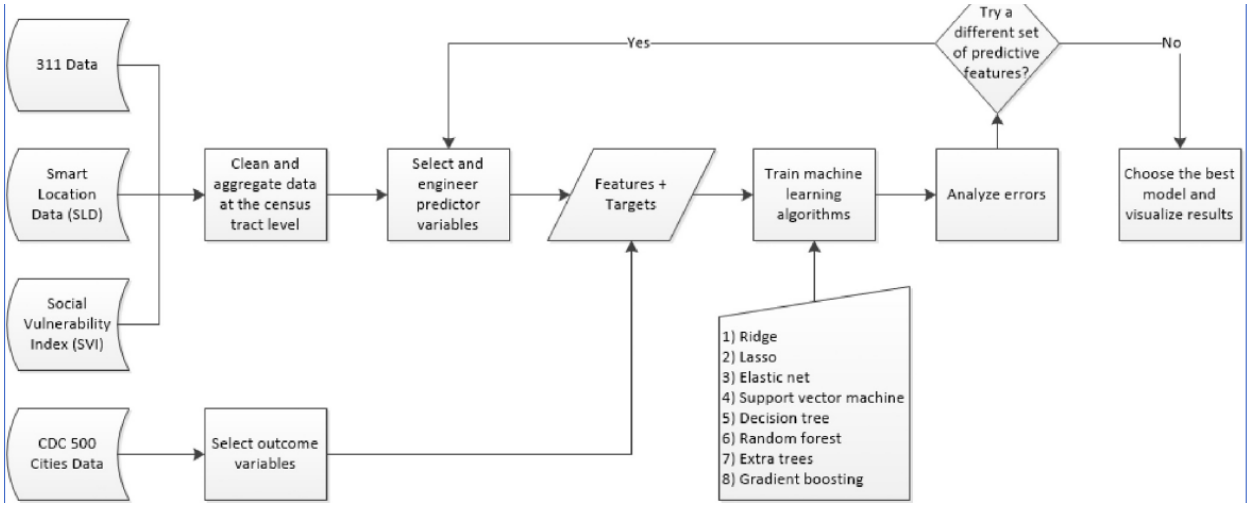
The connection between the financial resources people have – income, cost of living, and socioeconomic status – and their health. This area includes key issues such as poverty, employment, food security, and housing stability.

Neighborhood and Built Environment

The connection between where a person lives – housing, neighborhood, and environment – and their health and wellbeing. This includes topics like quality of housing, access to transportation, availability of healthy foods, air and water quality, and neighborhood crime and violence.

An empirical study

- Feng and Jiao (2021). Predicting and mapping neighborhood-scale health outcomes: A machine learning approach



An empirical study

Research gaps:

- A variety of machine learning approaches have recently been applied to predict health outcomes, albeit most have focused on the individual- level health outcomes

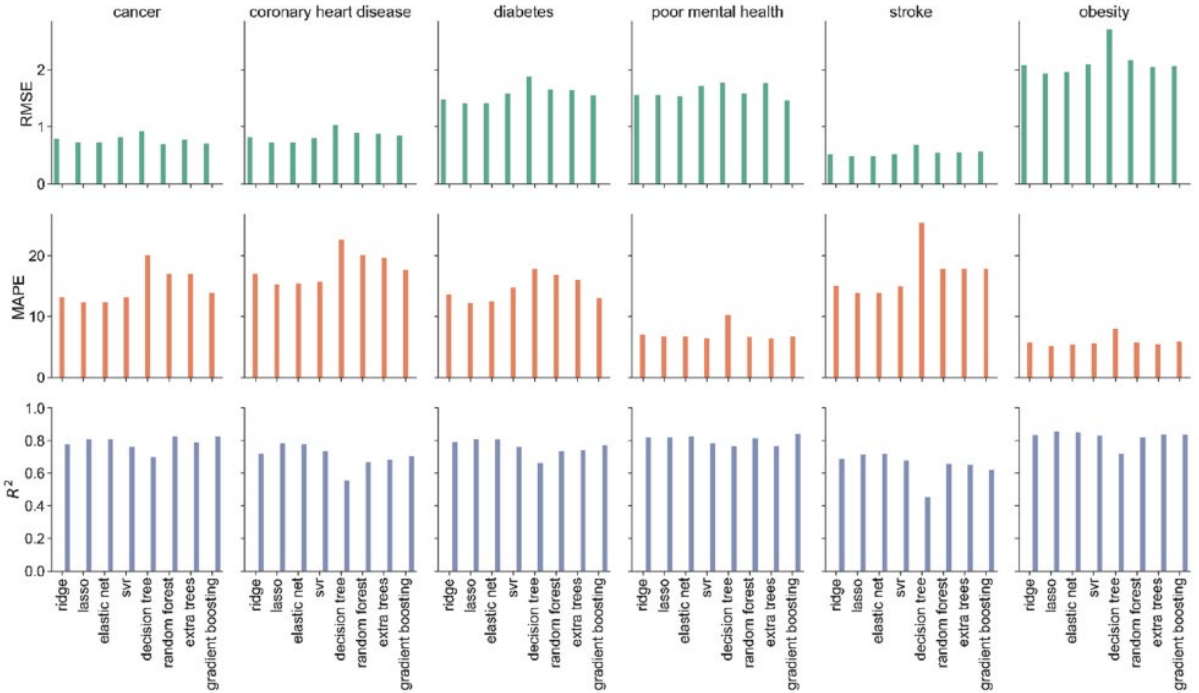
Data sources:

- 500 Cities Project data from the Centers for Disease Control and Prevention
- the Smart Location Database (SLD)
- the Social Vulnerability Index (SVI) data
- 311 service request data accessed via the City of Austin's open data portal

Methods: Ridge, Lasso, Elastic net, SVM, Decision tree, Random forest, Extra trees, Gradient boosting

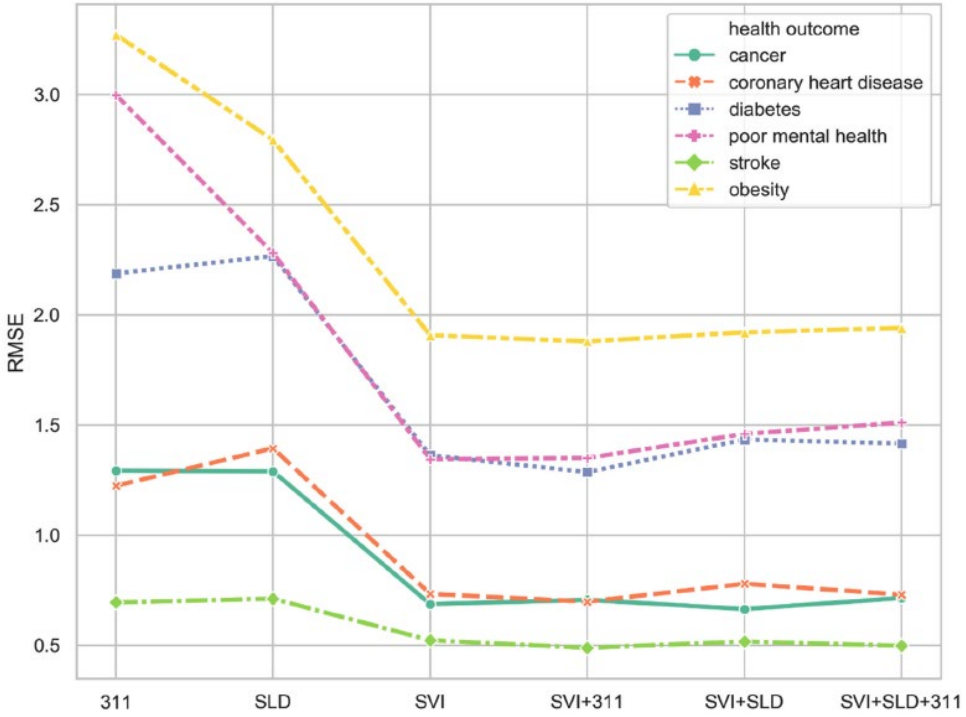
Outcome variables: cancer, coronary heart disease, diabetes, poor mental health, stroke, obesity

An empirical study



Determine the method

An empirical study

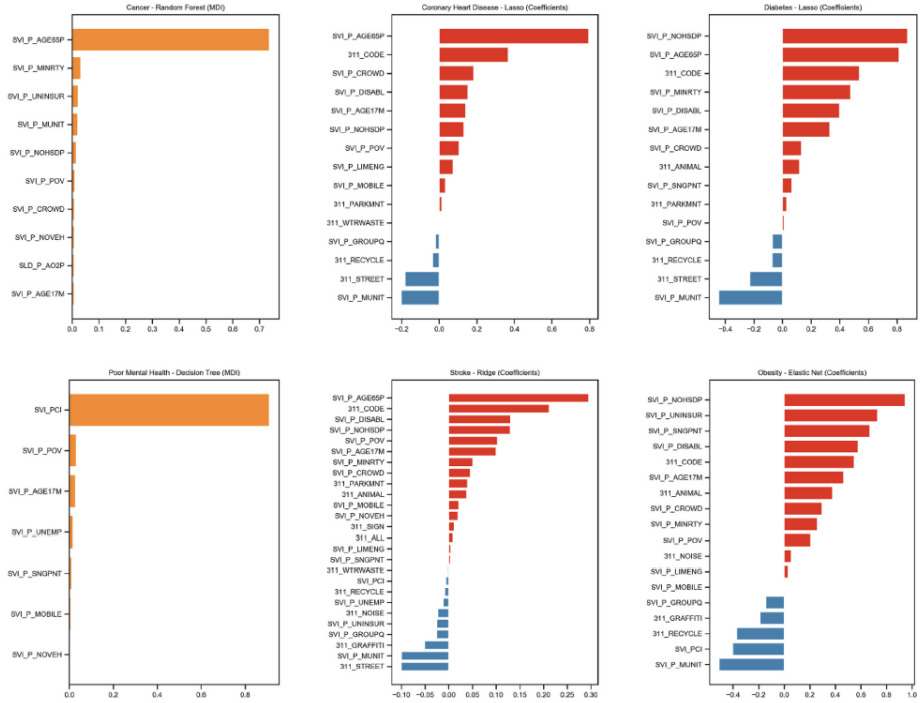


Determine the dataset

An empirical study

Algorithms and datasets that delivered the best test RMSEs in predicting the prevalence of health outcomes.

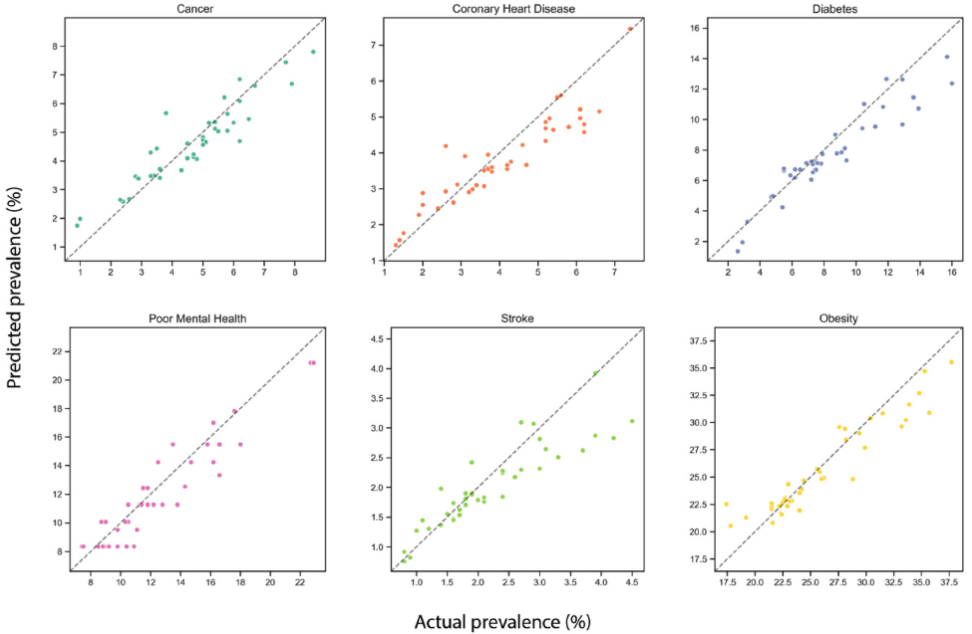
Health outcome	Algorithm	Datasets
Cancer	Random Forest	SVI + SLD
Coronary heart disease	Lasso Regression	SVI + 311
Diabetes	Support Vector Machine	SVI + 311
Poor mental health	Decision Tree	SVI
Stroke	Ridge Regression	SVI + 311
Obesity	Elastic Net	SVI + 311



An empirical study

Findings:

- Even at the census tract level, the sociodemographic and socioeconomic factors were the most powerful predictors of the prevalence of the health outcomes.
- In several cases, the additional predictive power contributed by the 311 data is particularly encouraging.



An empirical study

Two limitations:

- The 311 data may be biased and cannot be easily aggregated based on a meaningful set of categories
- The machine learning models are developed with only their predictive performance, thus providing very few insights on the potential causal pathways.

What do you think?

LAB TIME

HW: using the technologies learned today to explore the situations in Baltimore