Metabolic Syndrome

Analysis and Predictive Modeling

Cameron Peace

What is a syndrome? How does it differ from a disease?

"A group of signs and symptoms that occur together and characterize a particular abnormality or condition"

-Merriam Webster

"A syndrome is a set of medical signs and symptoms which are correlated with each other and often associated with a particular disease or disorder.[1] The word derives from the Greek σύνδρομον, meaning "concurrence".[2]:1818 When a syndrome is paired with a definite cause this becomes a disease."

-Wikipedia

"A syndrome is a recognizable complex of symptoms and physical findings which indicate a specific condition for which a direct cause is not necessarily understood. "

-Diagnoses, Syndromes, and Diseases: A Knowledge Representation Problem

Franz Calvo, MD,1 Bryant T Karras, MD,1,2 Richard Phillips, MD,1 Ann Marie Kimball, MD,1,2 and Fred Wolf, PhD1

What defines Metabolic Syndrome?

Metabolic syndrome is a group of conditions that together raise your risk of coronary heart disease, diabetes, stroke, and other serious health problems.

The National Institutes of Health guidelines define metabolic syndrome as having **three or more of the following traits**, including traits for which you may be taking medication to control:

- **Large waist —** A waistline that measures at least 35 inches (89 centimeters) for women and 40 inches (102 centimeters) for men
- **High triglyceride level —** 150 milligrams per deciliter (mg/dL), or 1.7 millimoles per liter (mmol/L), or higher of this type of fat found in blood
- **Reduced "good" or HDL cholesterol —** Less than 40 mg/dL (1.04 mmol/L) in men or less than 50 mg/dL (1.3 mmol/L) in women of high-density lipoprotein (HDL) cholesterol
- **Increased blood pressure —** 130/85 millimeters of mercury (mm Hg) or higher
- **Elevated fasting blood sugar —** 100 mg/dL (5.6 mmol/L) or higher

-NIH

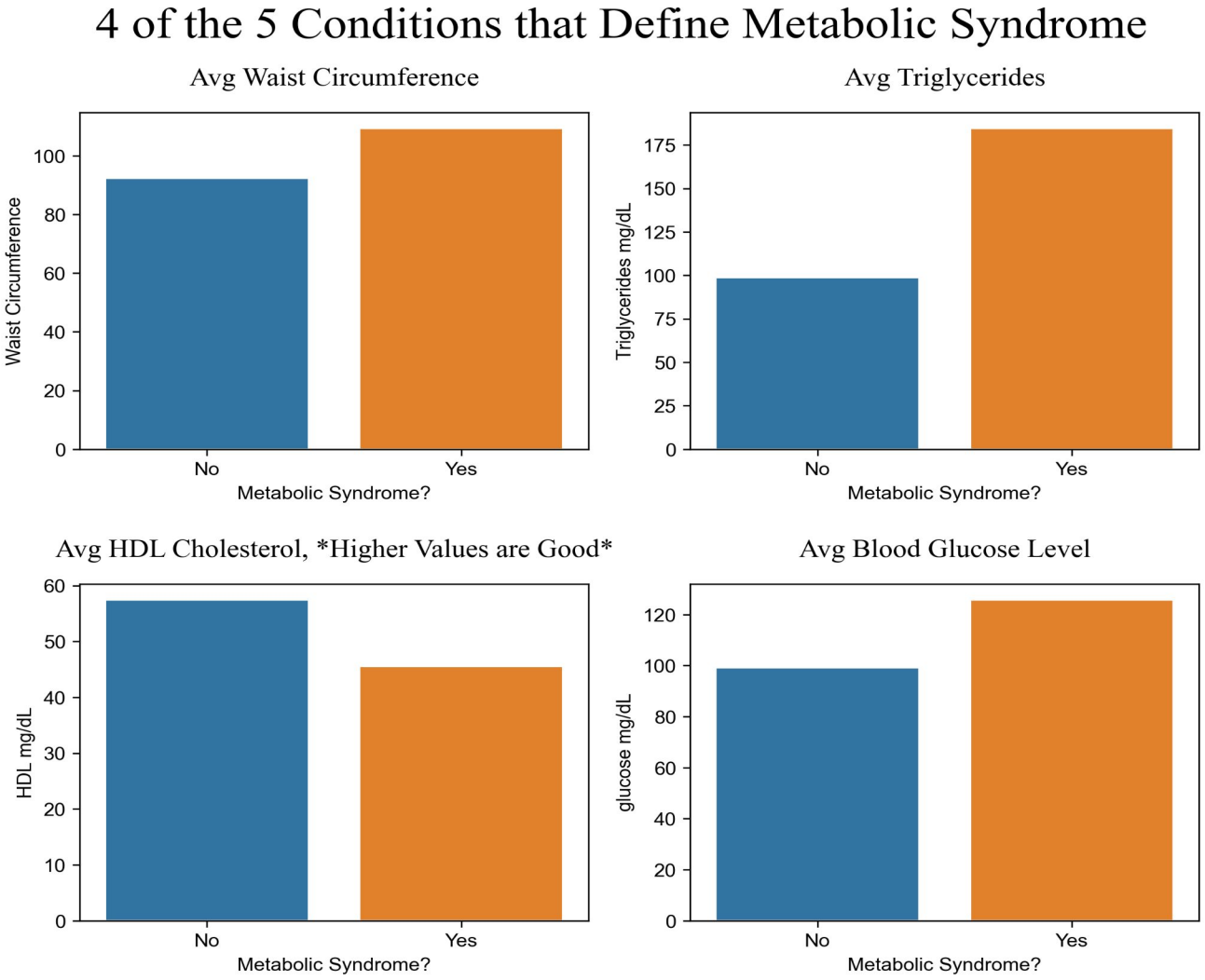
What data were used?

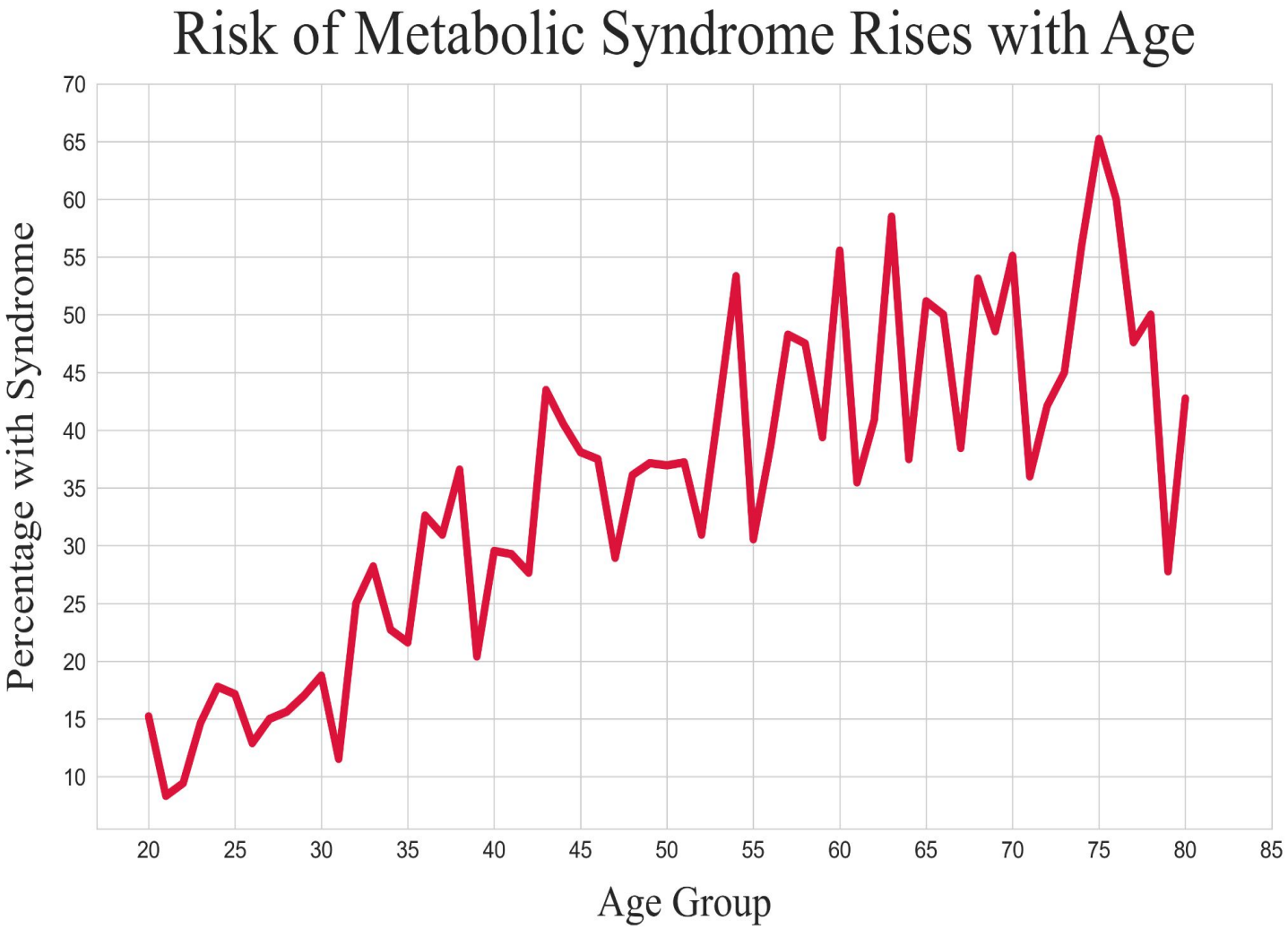
The dataset for analysis came from the NHANES initiative where the following variables were combined from multiple tables with SQL: abnormal waist circumference, triglycerides above 150, HDL cholesterol below 50 in women or 40 in men, history of hypertension and mildly elevated fasting blood sugar (100-125). Numerous other variables were added, such as uric acid, race, income, etc.

- From CDC website: More information on NHANES
  The National Health and Nutrition Examination Survey (NHANES) is a program of studies designed to assess the health and nutritional status of adults and children in the United States. The survey is unique in that it combines interviews and physical examinations. NHANES is a major program of the National Center for Health Statistics (NCHS). NCHS is part of the Centers for Disease Control and Prevention (CDC) and has the responsibility for producing vital and health statistics for the Nation.
- It is not clear from the description on data.world when the data were collected and what regions or demographic groups the individuals in the study represent.
- It was uploaded to data.world by Robert Hoyt MD on July 22nd, 2019.

Which defining features of metabolic syndrome are present in the data?

We are missing data on blood pressure





The data may exhibit "survivorship" bias

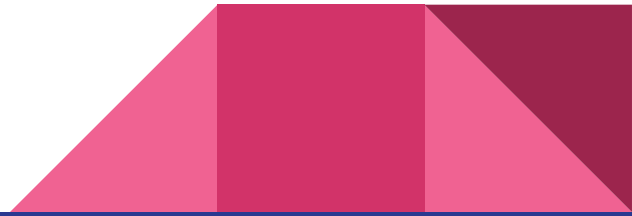
Who might benefit from a predictive model?

- Hospitals and other healthcare organizations could deploy the model directly to diagnose patients with metabolic syndrome outright or simply flag their risk level.
- Companies that develop healthcare software could incorporate the model into a suite of diagnostic tools

How does the model perform?

- The model was 89% accurate at predicting a diagnosis for metabolic syndrome given the data available.
- However, it was only able to correctly label positive diagnoses 82% of the time.

In other words, 89% of the time the model correctly predicted a positive or negative diagnosis of metabolic syndrome, but 18% of the time it missed a positive diagnosis and labeled it as negative.



Model Limitations

- Because a diagnosis of metabolic syndrome involves the presence of 3 or more well defined quantitative measurements above certain thresholds, a predictive model is really only useful in cases where data is missing or unavailable.
- Most of the metrics that define metabolic syndrome are routinely collected in the average visit for preventative care, so scenarios where deploying and maintaining a predictive model is preferable over data collection may be unlikely.
- The accuracy of the current model is currently too low to be effective. It would likely lead to lack of trust by users.
- The dataset used for modeling was very small and also unbalanced, the model would need much more testing on additional data to validate its efficacy.

Recommendations

- In my opinion this is a case where predictive modeling is unnecessary.
- Simpler, filter based screening software could be deployed that would be cheaper, faster, and much easier to maintain. It could even provide warning “levels”, based on what ranges the appropriate patient metrics fall within.
- If there is a use case where predictive modeling is needed, I would recommend collecting much more data, better quality data, and broader data (i.e. more patient features/attributes). This may allow for developing a more accurate model.