Predicting the Risk of Recurrent MI

SIBS Hackathon Presentation

Kevin Krupa, Lucy Liu, Kendall McClellan, Connor McNeill

NCSU-Duke Summer Institute in Biostatistics

July 21, 2022

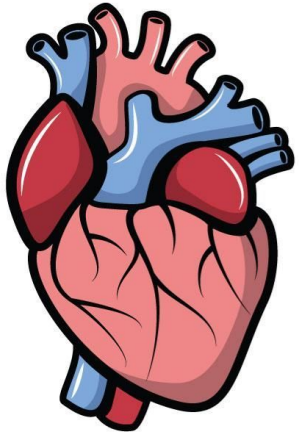
Outline

1. Background
2. Research Questions
3. Descriptive Analysis
4. Inferential Analysis
5. Discussion

Outline

1. Background
2. Research Questions
3. Descriptive Analysis
4. Inferential Analysis
5. Discussion

Cardiovascular Disease



Disorders of the heart and blood vessels

Leading cause of death globally

Most prevalent in wealthy, urban populations

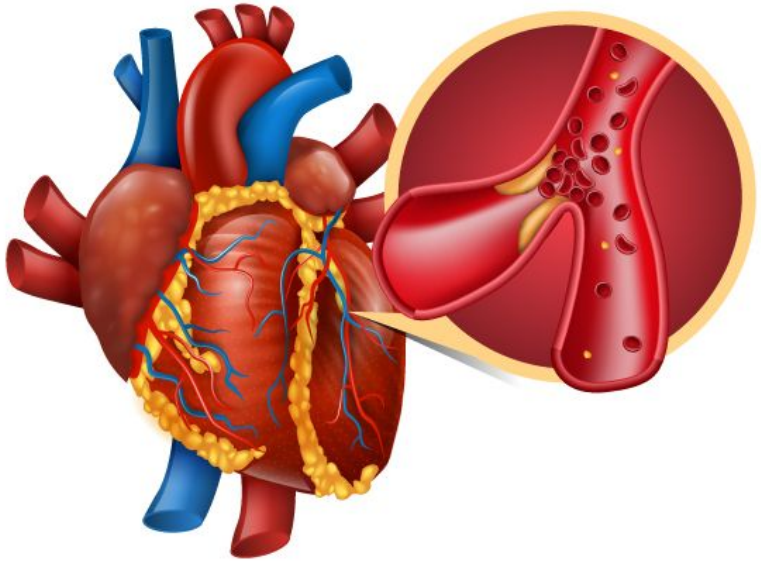
Source: World Health Organization and Golovenkin et al.

Myocardial Infarction

aka **Heart Attack**

Blockage of blood flow to heart

Heart lacks oxygen

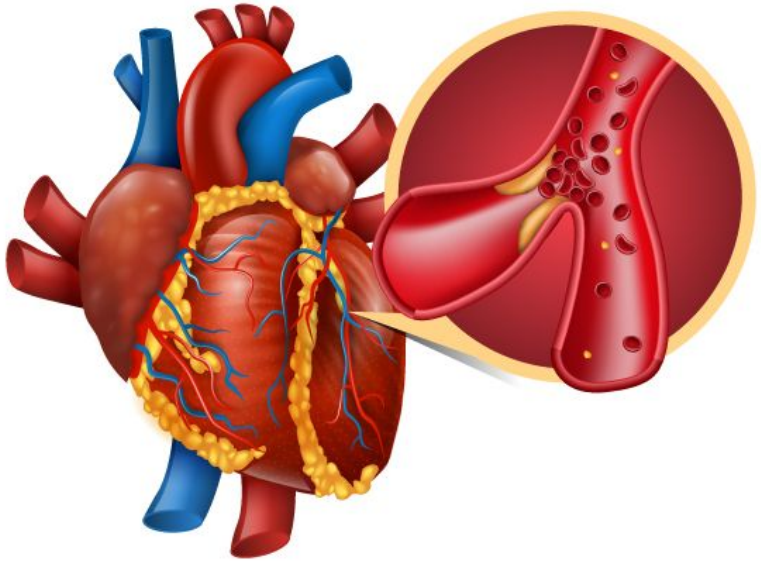


Source: Johns Hopkins Medicine

Myocardial Infarction

Risk factors: hypertension, age, lifestyle, comorbidities

Complications: pulmonary edema, A-Fib, **relapse**



Source: Johns Hopkins Medicine

Outline

1. Background
2. Research Questions
3. Descriptive Analysis
4. Inferential Analysis
5. Discussion

Research Questions

Question 1

What can we look at either during admission or the hospital stay in patients post-heart attack (myocardial infarction) to predict the risk of having a relapse?

Question 2

Which risk prediction model is best at predicting a recurrent heart attack?

Outline

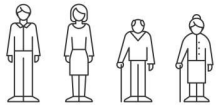
1. Background
2. Research Questions
3. Descriptive Analysis
4. Inferential Analysis
5. Discussion

The Data

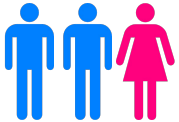
Myocardial Infarction Complications Database

Patients admitted to Russian hospital for MI from **1992-1995**

1700 patients ages **26-92**



50% older than **63**



2x as many males as females

Question 1

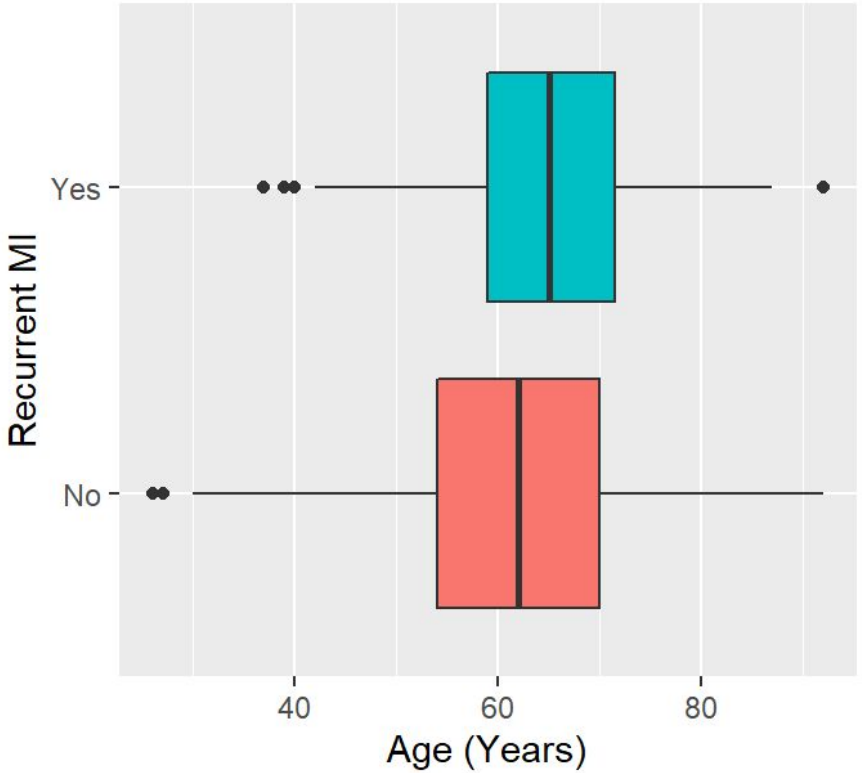
What can we look at either during admission or the hospital stay in patients post-heart attack (myocardial infarction) to predict the risk of having a relapse?

Background research suggests the following (among others):

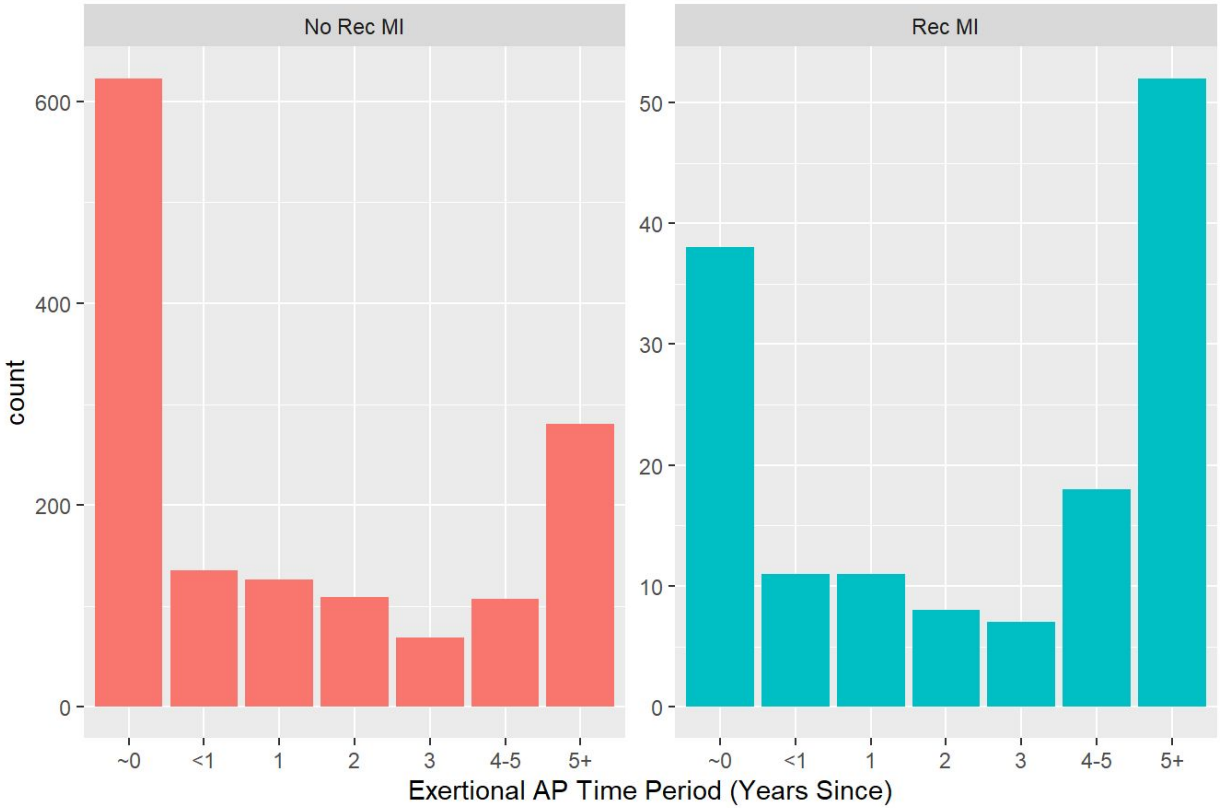
- Age
- Diabetes mellitus
- Medically managed intervention
- Smoking history and/or COPD

We will determine prospective covariates using model selection

Age



History of Chest Pain



Question 2

Which risk prediction model is best at predicting a recurrent heart attack?

Basis for models is **logistic regression**

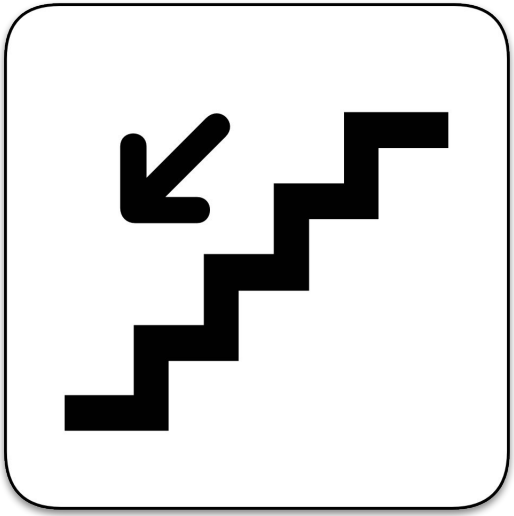
- Response variable is binary (0 or 1)
- Multiple predictors
  - Includes Binary, Ordinal, and Continuous variables
  - **> 100 variables** to begin
  - Variables removed if significant amount of missingness & insignificant
- Two methods used for model building & selection:
  1. **Stepwise Regression**
  2. **Elastic Net Regression**

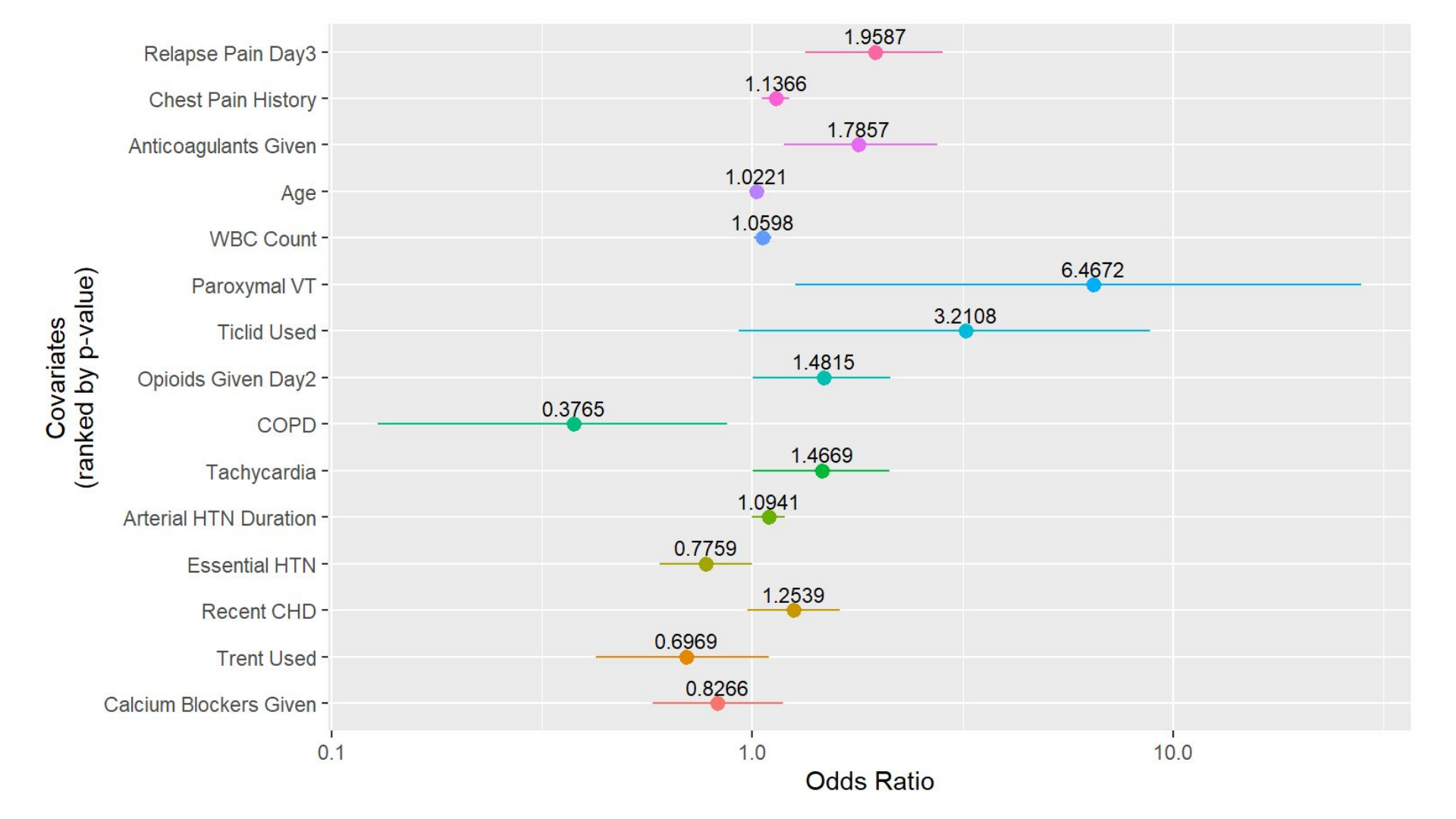
Outline

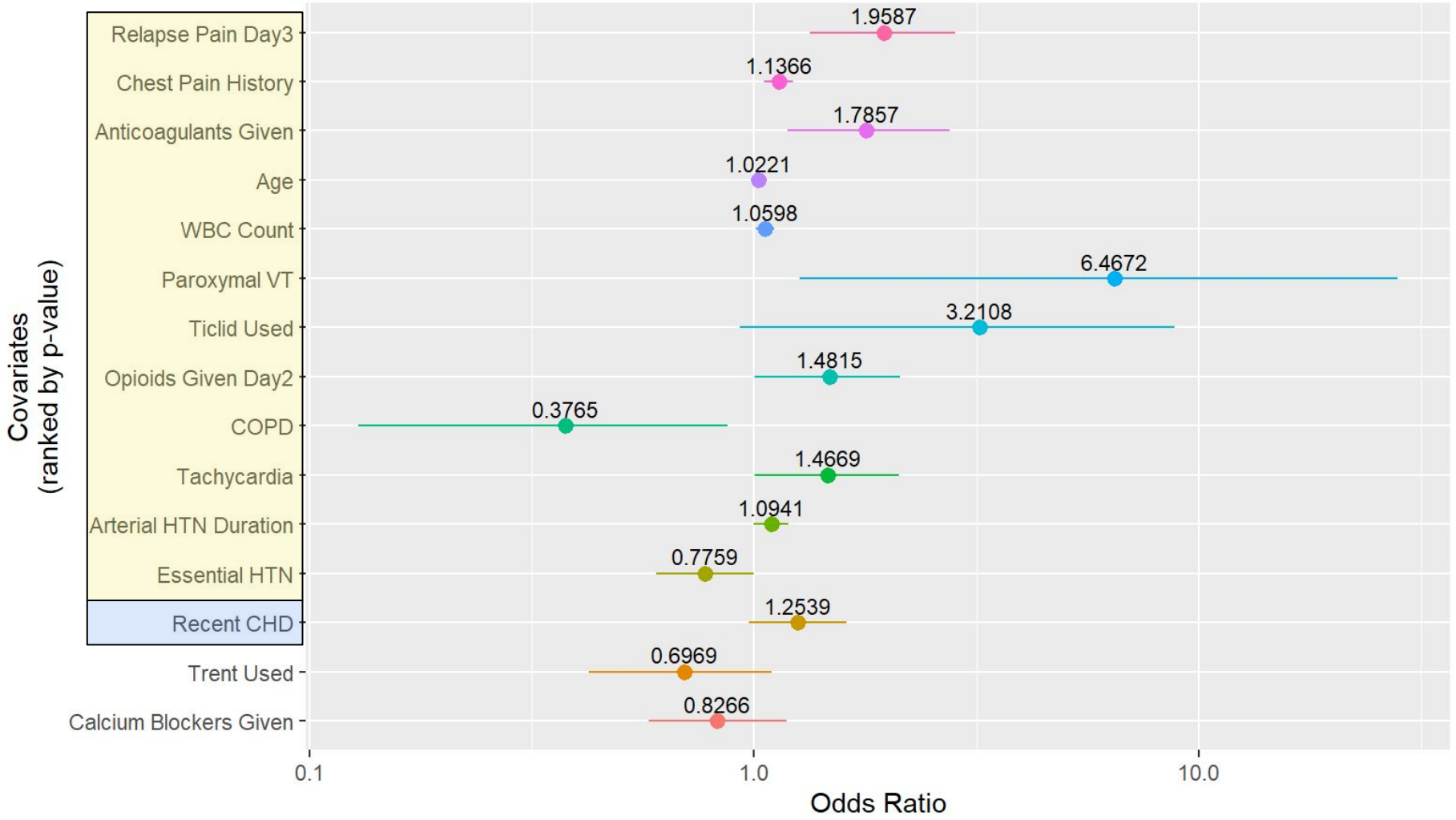
1. Background
2. Research Questions
3. Descriptive Analysis
4. Inferential Analysis
5. Discussion

Model 1: Stepwise Regression Model

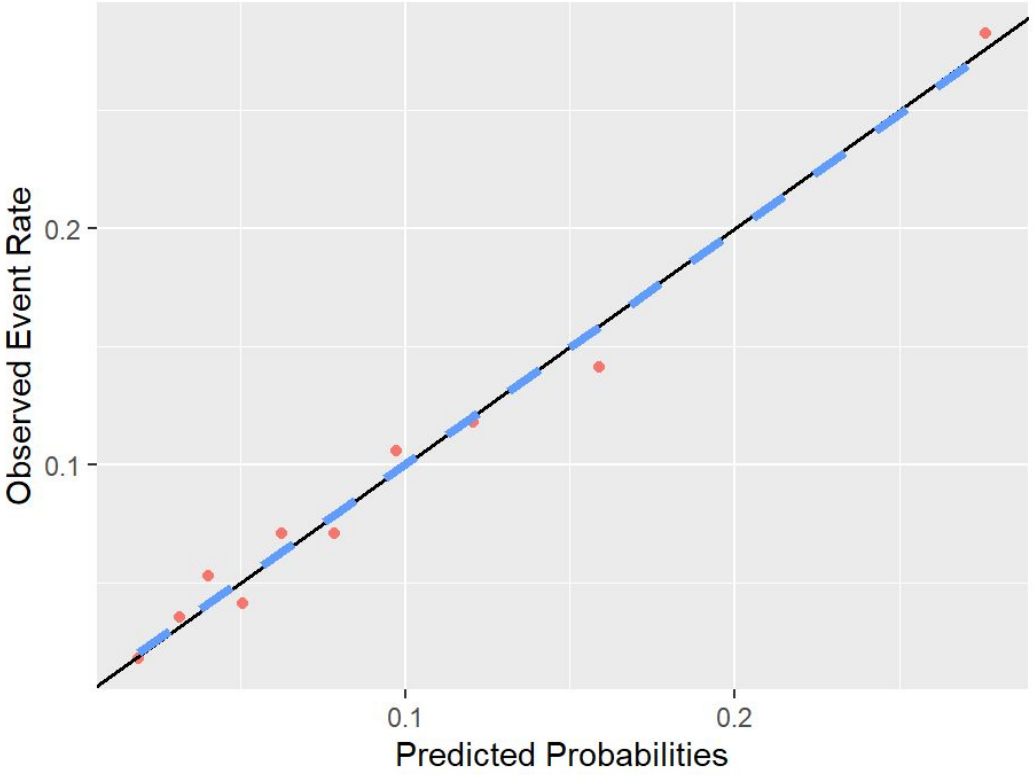
- Used `stepAIC` from the `MASS` library
- Started with **48 candidate** covariates
- Started with full model and removed/added variables
- **AIC** used to determine best model
- Removed one more covariate due to **high p-value**




Relapse Pain Day3
Chest Pain History
Anticoagulants Given
Age
WBC Count
Paroxymal VT
Ticlid Used
Opioids Given Day2
COPD
Tachycardia
Arterial HTN Duration
Essential HTN
Recent CHD
Trent Used
Calcium Blockers Given
1.9587
1.1366
1.7857
1.0221
1.0598
6.4672
3.2108
1.4815
0.3765
1.4669
1.0941
0.7759
1.2539
0.6969
0.8266
0.1
1.0
10.0
Odds Ratio
Covariates
(ranked by p-value)


Covariates
(ranked by p-value)
Relapse Pain Day3
Chest Pain History
Anticoagulants Given
Age
WBC Count
Paroxymal VT
Ticlid Used
Opioids Given Day2
COPD
Tachycardia
Arterial HTN Duration
Essential HTN
Recent CHD
Trent Used
Calcium Blockers Given
1.9587
1.1366
1.7857
1.0221
1.0598
6.4672
3.2108
1.4815
0.3765
1.4669
1.0941
0.7759
1.2539
0.6969
0.8266
0.1
1.0
10.0
Odds Ratio

Model 1: Calibration Plot

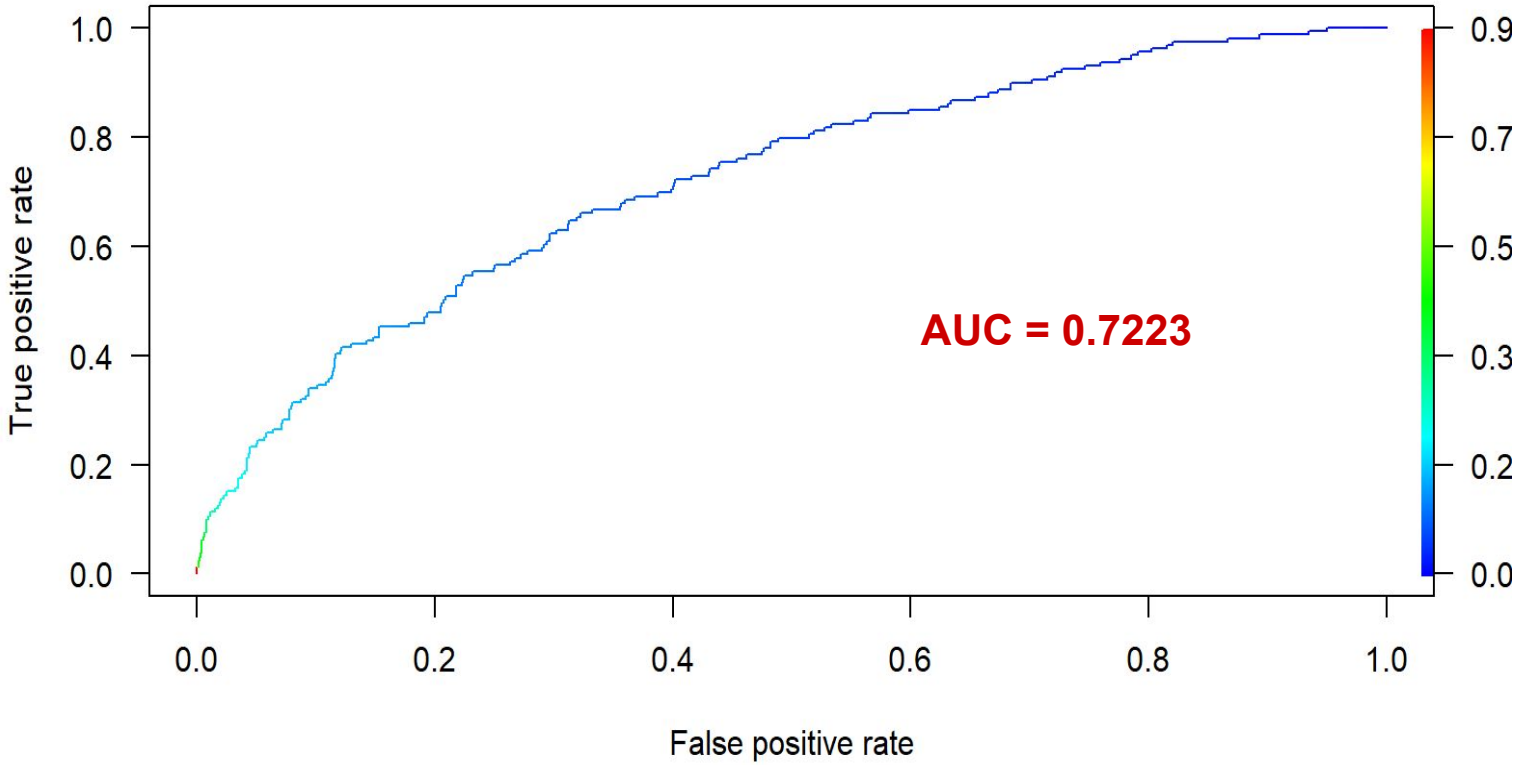


Calibration intercept: 0.001149

Calibration slope: 0.987710

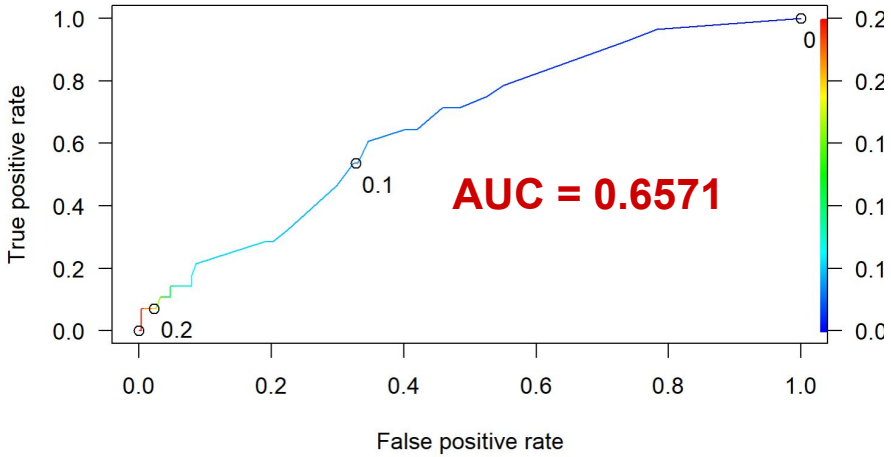
Values are close to the ideal - model is well calibrated

Model 1: ROC Curve



Model 2: Elastic Net Regression

- LASSO/RIDGE/Elastic Net Investigated
- Created Test/Train Sets
  - Multiple Imputation
- Enforced Strength of Some Model 1 Predictors
  - Relapse of Pain
  - History of Chest Pain



Model 2: Elastic Net Regression

Predictor	**Odds Ratio**
Chest Pain History (STENOK_AN)	1.060
1st Degree AV Block (np_01)	6.207
Ventricular Tachycardia (GT_POST)	2.357
Incomplete RBBB (n_p_ecg_p_11)	1.379
Relapse Pain Day 3 (R_AB_3_n)	1.610
Use of Opioid Drugs (NA_KB)	1.116

Outline

1. Background
2. Research Questions
3. Descriptive Analysis
4. Inferential Analysis
5. Discussion

Question 1

What can we look at either during admission or the hospital stay in patients post-heart attack (myocardial infarction) to predict the risk of having a relapse?

- Both models selected:
  - Relapse of Pain during Hospital Stay (on Day 3)
  - Past History of Chest Pain due to CHD
- Shows strength of those two predictors
- Were also most significant by p-value for stepwise model

Question 1

What can we look at either during admission or the hospital stay in patients post-heart attack (myocardial infarction) to predict the risk of having a relapse?

- Other strong covariates:
  - Age
  - White Blood Cell Count
  - Certain medications reduced risk (Trent, Calcium Blockers)
  - Others are associated with higher risk (Opioids, Ticlid, Anticoagulants)
    (Risk-benefit analysis with medications and sicker → more meds)

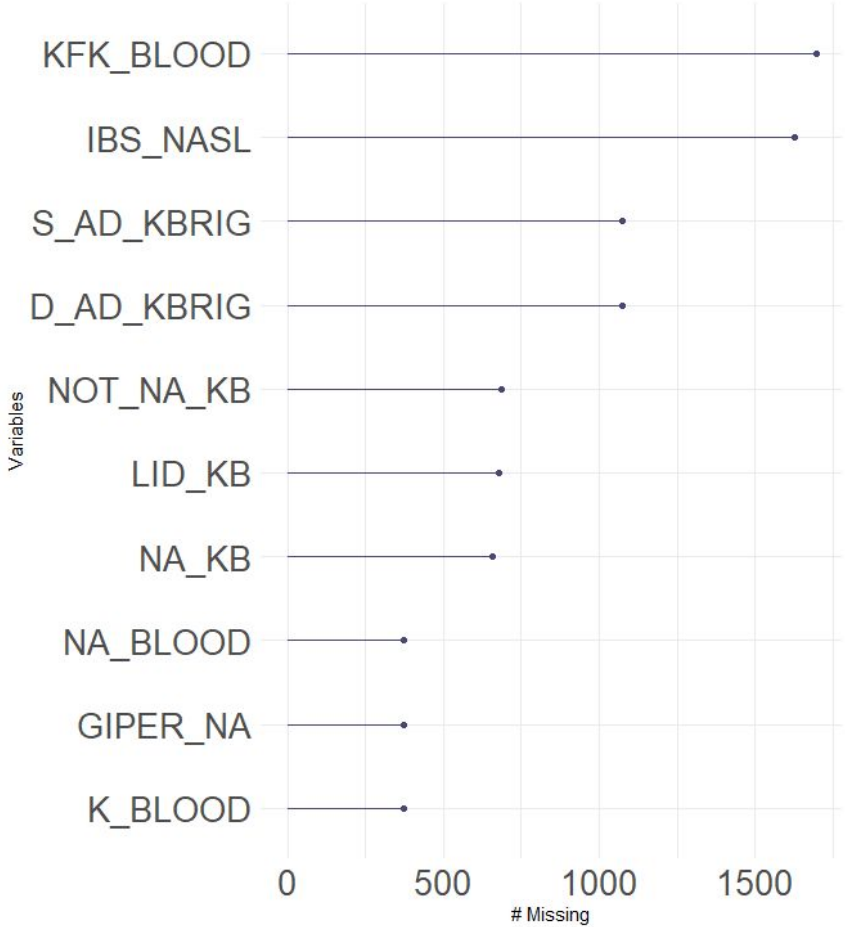
Question 2

Which risk prediction model is best at predicting a recurrent heart attack?

- Stepwise Regression Model
  - Best result for predicting between the two
  - Still not amazing though (AUC < 0.8)
- Elastic Net Regression Model
  - Not as accurate but still helpful
    - Penalized Regression helps determine significant covariates

Missing Data

- 10 variables with most missingness
- 7.6% missing data
- Solution: **`mice`** (Multivariate Imputation by Chained Equations)
- Method: Predictive Mean Matching (`pmm`)



Limitations

Of the Data…

- Missing data
- Lack of diversity
- No data on socioeconomic factors
- Binary/Ordinal less accurate

Of the Analysis…

- MICE accuracy unknown

Moving Forward

- Examine social determinants of health, incl. demographics
- Role of genetics
- More numerical variables
- Analyze more recent data
- Model comparing hospital interventions

References

“Cardiovascular Diseases (CVDs).” World Health Organization. World Health Organization, June 11, 2021. https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds).

Golovenkin, S.E.; Gorban, Alexander; Mirkes, Evgeny; Shulman, V.A.; Rossiev, D.A.; Shesternya, P.A, Nikulina S Y, Orlova Y V and Dorrer M G. 2020: Myocardial infarction complications Database. University of Leicester. Dataset. https://doi.org/10.25392/leicester.data.12045261.v3

“Heart Attack.” Johns Hopkins Medicine. The Johns Hopkins University, 2022. https://www.hopkinsmedicine.org/health/conditions-and-diseases/heart-attack.

Nair, Raunak, Michael Johnson, Kathleen Kravitz, Chetan Huded, Jeevanantham Rajeswaran, Moses Anabila, Eugene Blackstone, et al. 2021. “Characteristics and Outcomes of Early Recurrent Myocardial Infarction After Acute Myocardial Infarction.” Journal of the American Heart Association 10 (16): e019270. https://doi.org/10.1161/JAHA.120.019270.

Song, Jiali, Karthik Murugiah, Shuang Hu, Yan Gao, Xi Li, Harlan M. Krumholz, and Xin Zheng. 2021. “Incidence, Predictors, and Prognostic Impact of Recurrent Acute Myocardial Infarction in China.” Heart 107 (4): 313–18. https://doi.org/10.1136/heartjnl-2020-317165.

Questions??