Examining the Presence & Severity of Coronary Artery Disease based on a Cardiac Catheterization

By: Thu Vu

SCHOOL OF PUBLIC HEALTH
UNIVERSITY of WASHINGTON

INTRODUCTION

- Heart Disease is the leading cause of death in the U.S.
- Coronary Artery Disease (CAD): most common type of heart disease
  - Plaque build-up in the coronary arteries
  - Can lead to a heart attack
  - Risk Factors: Age, Sex, & Cholesterol Levels
- Cardiac Catheterization (or Heart Cath): catheter used to examine condition of the coronary arteries
  - Diagnose and/or treat CAD

DATASET BACKGROUND

- Observational cross-sectional study with 3504 participants
- Contains 6 variables with 1249 missing observations:
  - sex (0 = Male, 1 = Female)
  - age (years)
  - cad_dur (days): duration of chest pain symptoms
  - choleste (milligrams per deciliter) --- serum cholesterol level
  - sigdz (0 = No, 1 = Yes): Significant Coronary Disease
  - tvdlm (0 = No, 1 = Yes): Severe Coronary Disease (Either Three Vessel Disease or Left Main Disease)
- Significant Coronary Disease: a blockage of at least 75% in at least one of the major coronary arteries
- Three Vessel Disease & Left Main Disease: severe forms of CAD & are both associated with high risk of mortality

SCIENTIFIC QUESTIONS

- Target Population: adults aged 20 years and older who are experiencing chest pain symptoms
- Primary Objective: Is there an association between having Significant or Severe Coronary Disease and the duration of chest pain symptoms adjusting for age, sex and cholesterol levels?
- Secondary Objective: Based on the primary objective, is cholesterol level a possible effect modifier?
  - Continuous variable, cholesterol level, will be dichotomized:
    - High cholesterol: greater than or equal to 239 mg/dl (0=Otherwise, 1=Yes)

STATISTICAL METHODS

- Fit 4 Multiple Logistic Regression Models
- Reference Group for Sex are Males
- Baselines: Duration of Chest Pain (0 Days), Age (20 Years old), & Cholesterol Level (0 mg/dl)
- Confidence Intervals using robust standard errors
- Likelihood Ratio Tests to derive p-values (significance level = 0.05)

	Model 1	Model 2	Model 3	Model 4
Response	$sigdz$	$tvdlm \mid sigdz = 1$	$sigdz$	$tvdlm \mid sigdz = 1$
Covariates	$cad.dur$ $sex{=}Female$ $age\text{-}20$ $choleste$	$cad.dur$ $sex{=}Female$ $age\text{-}20$ $choleste$	$cad.dur$ $sex{=}Female$ $age\text{-}20$ $choleste{=}high$ $cad.dur \times choleste = high$	$cad.dur$ $sex{=}Female$ $age\text{-}20$ $choleste{=}high$ $cad.dur \times choleste = high$

DESCRIPTIVE ANALYSIS

- Table 1 (Model 1):
  - ⅔ has Significant CAD; ⅓ does not
  - Duration of Chest Pain & High Cholesterol: Patients who have Significant CAD experience more days and larger % has High Cholesterol
- Table 2 (Model 2):
  - 48% has Severe CAD; 52% does not
  - Duration of Chest Pain & High Cholesterol: Larger mean difference and not as large % difference than Table 1

Table 1: Summary Statistics for observed data grouped by indicator of Significant Coronary Disease

	Doesn't Have Significant Coronary Disease	Has Significant Coronary Disease	Overall
Proportion (N)	0.333 (1167)	0.667 (2334)	1 (3501)
Sex (Female)			
Proportion (N)	0.546 (637)	0.198 (462)	0.314 (1099)
Age (years)			
Mean (SD)	48.9 (9.72)	54 (9.51)	52.3 (9.89)
Median [Min, Max]	49 [20.0, 80.0]	54 [25.0, 82.0]	52 [20.0, 82.0]
Duration of Chest Pain (days)			
Mean (SD)	36.5 (53.9)	46.3 (59.9)	43.0 (58.2)
Median [Min, Max]	13 [0, 372]	22 [0, 416]	18 [0, 416]
Cholesterol Level (mg/dl)			
Mean (SD)	220 (44.3)	235 (52.8)	230 (50.6)
Median [Min, Max]	216 [89, 500]	230 [29, 576]	225 [29, 576]
Missing (%)	401 (34.4%)	844 (36.2%)	1245 (35.6%)
Has High Cholesterol (>238 mg/dl)			
Proportion (N)	0.292 (224)	0.432 (643)	0.384 (867)
Missing (%)	401 (34.4%)	844 (36.2%)	1245 (35.6%)

Table 2: Summary Statistics for observed data grouped by indicator of Severe Coronary Disease given individual has Significant Coronary Disease

	Doesn't Have Severe Coronary Disease	Has Severe Coronary Disease	Overall
Proportion (N)	0.515 (1203)	0.484 (1129)	1 (2334)
Sex (Female)			
Proportion (N)	0.227 (273)	0.167 (188)	0.198 (462)
Age (years)			
Mean (SD)	52.4 (9.45)	55.8 (9.25)	54.0 (9.51)
Median [Min, Max]	52 [25, 82]	56 [25, 81]	54 [25, 82]
Duration of Chest Pain (days)			
Mean (SD)	34.6 (48.8)	58.7 (67.7)	46.3 (59.9)
Median [Min, Max]	12 [0, 364]	35 [0, 416]	22 [0, 416]
Cholesterol Level (mg/dl)			
Mean (SD)	233 (51)	237 (54.6)	235 (52.8)
Median [Min, Max]	230 [105, 576]	230 [29, 500]	230 [29, 576]
Missing (%)	436 (36.2%)	406 (36.0%)	844 (36.2%)
Has High Cholesterol (>238 mg/dl)			
Proportion (N)	0.425 (326)	0.438 (317)	0.432 (643)
Missing (%)	436 (36.2%)	406 (36.0%)	844 (36.2%)

DESCRIPTIVE ANALYSIS

- Distribution of Duration of Chest Pain Values is right-skewed
- Log-transform(+1 for 0 values) covariate of interest for future analyses

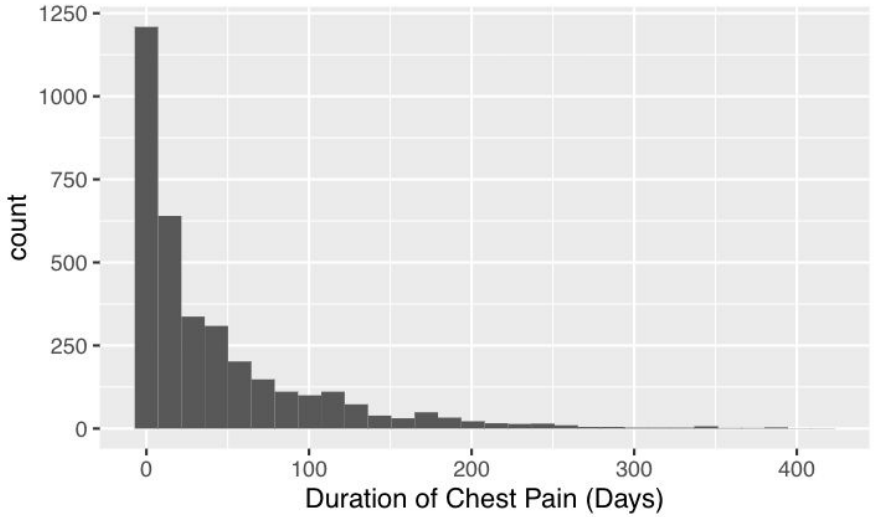

Figure 1: Histogram of Duration of Chest Pain

DESCRIPTIVE ANALYSIS

- Both have a higher median compared to those who don't have Significant/Severe CAD
- Larger Median difference in duration of chest pain in those who have Severe CAD as compared to those who have Significant CAD

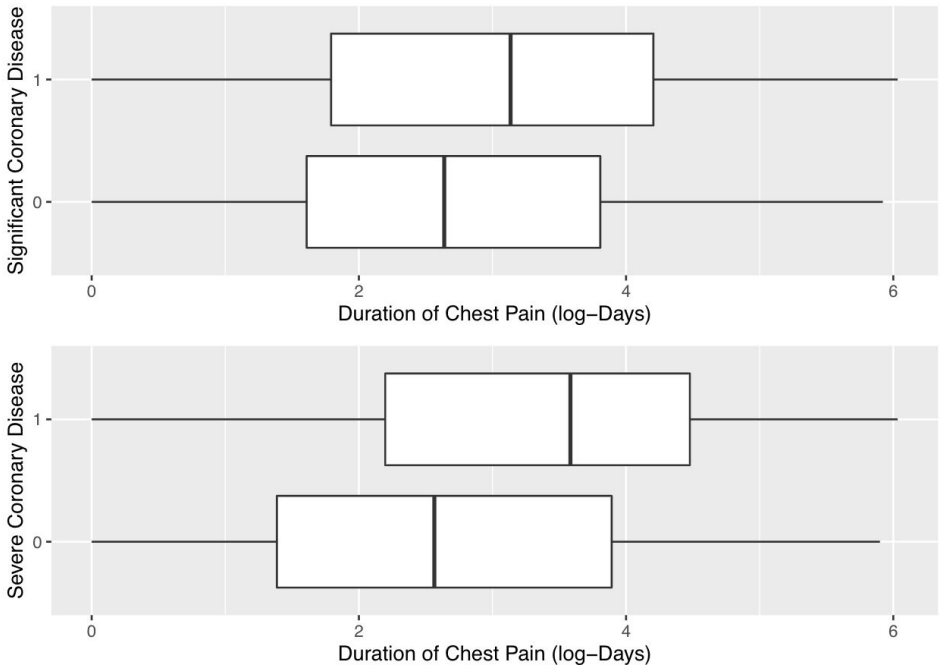

Figure 2: Boxplot of Significant/Severe Coronary Disease vs log-Duration of Chest Pain

DESCRIPTIVE ANALYSIS

- All have a higher median compared to those who don't have Significant/Severe CAD
- All have a noticeable median difference except for those who have Significant CAD and high cholesterol
  - The median difference is much smaller compared to those without high cholesterol

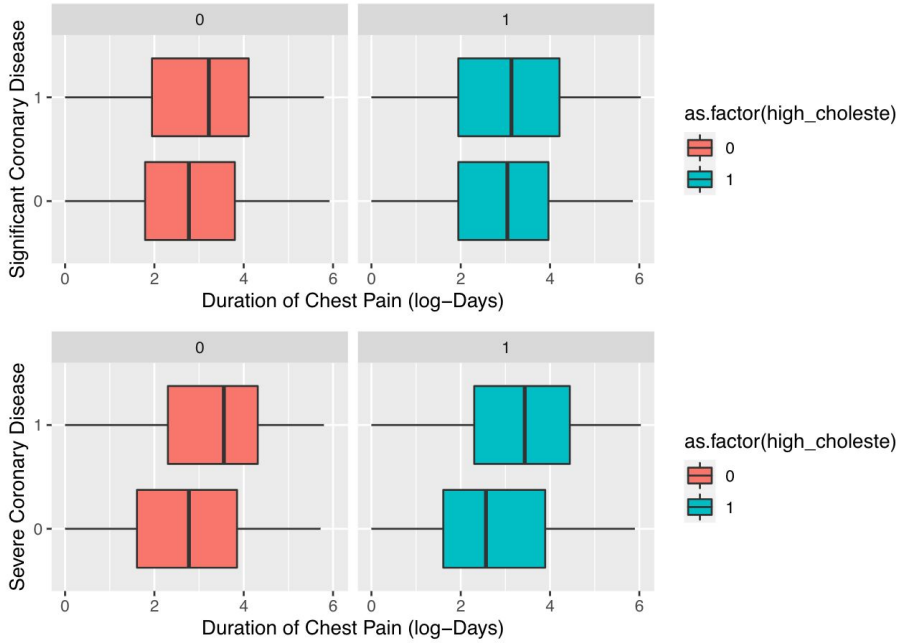

Figure 3: Boxplots of Significant/Severe Coronary Disease vs log-Duration of Chest Pain stratified by High Cholesterol Level

MISSING DATA ANALYSIS

- 6% Missing compared to 94% Observed
- Majority missing in cholesterol
- Some missing in indicator for Severe CAD
- Compare cholesterol to each covariate

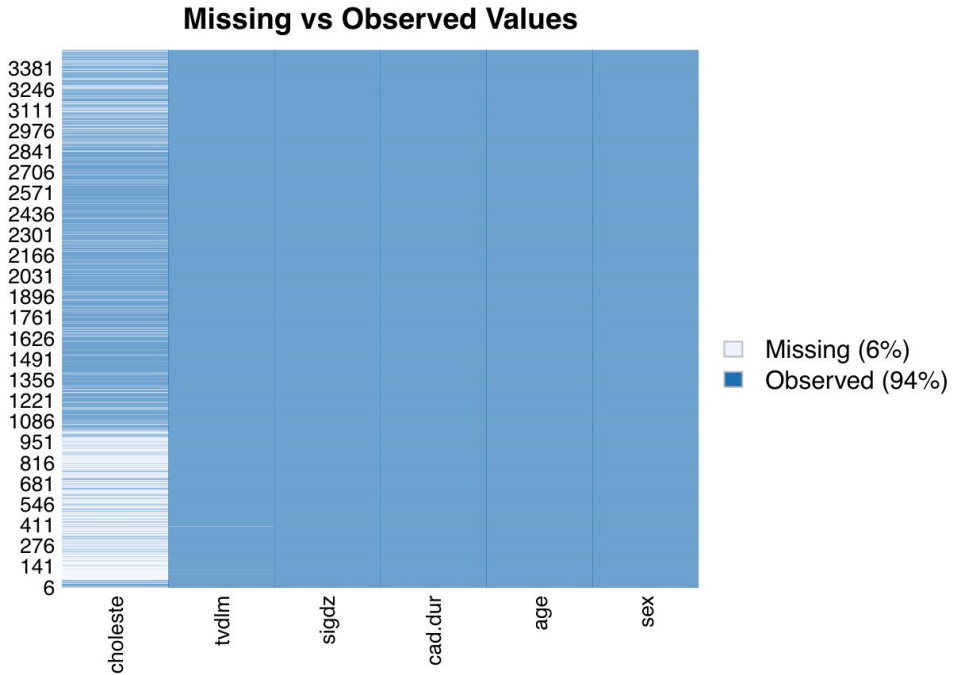

Figure 4: Plot of Missing vs Observed Values

MISSING DATA ANALYSIS

- 33% of the patients with missing cholesterol values are Females (~31% of the patients in the observed dataset are Females)
- Appears to have no missingness pattern (assume MAR)
- Multiple Regression Imputation to handle missing data

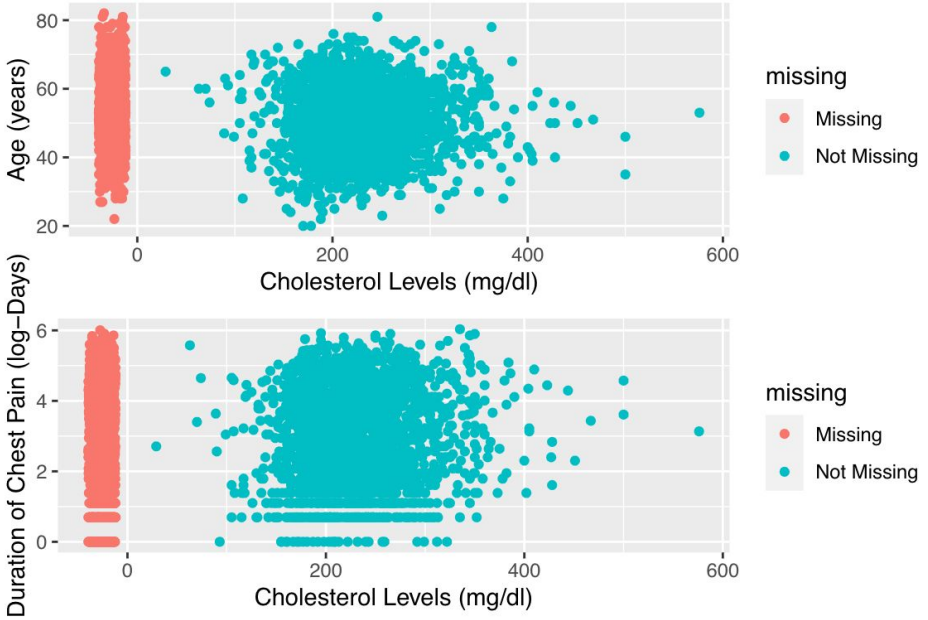

Figure 5: Plots of Cholesterol vs Age & Duration of Chest Pain stratified by missingness

RESULTS (MODEL 1)

- Same cholesterol level, sex, and age, but differ in duration of chest pain by 1 log-day, the group who experiences more log-days of chest pain has estimated odds of having Significant CAD 1.018 times greater than those who experiences less (95% CI: 0.961-1.079)
- Not statistically significant (p-value = 0.536)
- Fail to reject null hypothesis (no association between having Significant CAD and duration of chest pain)

Table 3: Estimated Odds Ratios of Having Significant Coronary Disease

	Estimate	Lower 95% CI	Upper 95% CI	LRT P-Value
Duration of Chest Pain (log-days)	1.018	0.961	1.079	0.536
Cholesterol Level (mg/dl)	1.009	1.007	1.010	–
Sex (Female)	0.123	0.102	0.148	–
Age (years over 20)	1.081	1.071	1.091	–

RESULTS (MODEL 2)

- Same cholesterol level, sex, and age, but differ in duration of chest pain by 1 log-day, the group who experiences more log-days of chest pain has estimated odds of having Severe CAD 1.304 times greater than those who experiences less (95% CI: 1.23-1.382)
- Is statistically significant (p-value < 0.05)
- Reject null hypothesis (Significant association between having Severe CAD and duration of chest pain)

Table 4: Estimated Odds Ratios of Having Severe Coronary Disease given the individual has Significant Coronary Disease

	Estimate	Lower 95% CI	Upper 95% CI	LRT P-Value
Duration of Chest Pain (log-days)	1.304	1.230	1.382	2.01e-19
Cholesterol Level (mg/dl)	1.002	1.001	1.004	–
Sex (Female)	0.529	0.421	0.665	–
Age (years over 20)	1.038	1.028	1.048	–

RESULTS (MODEL 3)

- Same cholesterol level, sex, and age, but differ in duration of chest pain by 1 log-day, the estimated difference in odds of having Significant CAD between two groups that have high cholesterol is 0.96 times lower than the difference between two groups that don't have high cholesterol (95% CI: 0.855-1.079)
- Not statistically significant (p-value = 0.495)
- High Cholesterol is likely not an effect modifier

Table 5: Estimated Odds Ratios of Having Significant Coronary Disease (with interaction)

	Estimate	Lower 95% CI	Upper 95% CI	LRT P-Value
Duration of Chest Pain (log-days)	1.049	0.978	1.125	–
Cholesterol Level (High)	2.471	1.691	3.611	–
Sex (Female)	0.127	0.106	0.152	–
Age (years over 20)	1.079	1.069	1.089	–
Duration of Chest Pain (log-days):Cholesterol Level (High)	0.960	0.855	1.079	0.495

RESULTS (MODEL 4)

- Same cholesterol level, sex, and age, but differ in duration of chest pain by 1 log-day, the estimated difference in odds of having Significant CAD between two groups that have high cholesterol is 1.058 times greater than the difference between two groups that don't have high cholesterol (95% CI: 0.943-1.187)
- Not statistically significant (p-value = 0.341)
- High Cholesterol is likely not an effect modifier

Table 6: Estimated Odds Ratios of Having Severe Coronary Disease given the individual has Significant Coronary Disease (with interaction)

	Estimate	Lower 95% CI	Upper 95% CI	LRT P-Value
Duration of Chest Pain (log-days)	1.278	1.183	1.380	–
Cholesterol Level (High)	0.951	0.648	1.396	–
Sex (Female)	0.548	0.437	0.686	–
Age (years over 20)	1.037	1.027	1.047	–
Duration of Chest Pain (log-days):Cholesterol Level (High)	1.058	0.943	1.187	0.341

DISCUSSION & LIMITATIONS

- Significant association between having Severe CAD and duration of chest pain
  - Symptoms are more severe, so likely to last longer
  - For Significant CAD, symptoms might be mild/moderate, thus, duration might fluctuate; Might be difficult to keep track of
- Observational study, so results are exploratory in nature and does not imply a causal relationship
- Did not record severity of symptoms
- Was duration of symptoms recorded based on consecutive days or days in a certain timeframe?
- Breaking down the chest pain symptoms into more specific symptoms
- Adjust for other possible confounders (smoking status, diet, exercise, etc.)

REFERENCES

- Clinic, Cleveland. 2018. “Understanding Your Cholesterol Numbers | Cleveland Clinic.” Cleveland Clinic.https://my.clevelandclinic.org/health/articles/11920-cholesterol-numbers-what-do-they-mean.
- Clinic, Mayo. 2019. “Cardiac Catheterization - Mayo Clinic.” Mayoclinic.org. https://www.mayoclinic.org/t ests-procedures/cardiac-catheterization/about/pac-20384695.
- ———. 2020. “Coronary Artery Disease.” Mayo Clinic. https://www.mayoclinic.org/diseases-conditions/coronary-artery-disease/symptoms-causes/syc-20350613.
- Digirad. 2020. “Triple Vessel Disease: Pathology, Diagnosis and Treatment.” Digirad. https://www.digirad. com/triple-vessel-disease-diagnosis-treatment/.
- Harrell Jr, Frank E. 2002. “Datasets.” hbiostat.org. http://hbiostat.org/data/#vanderbilt-biostatistics-datasets.
- Higgins, Peter. 2021. “Cath Dataset Description Document.” htmlpreview.github.io. https://htmlpreview.github.io/?https://github.com/higgi13425/medicaldata/blob/master/man/description_docs/cath_desc.html.