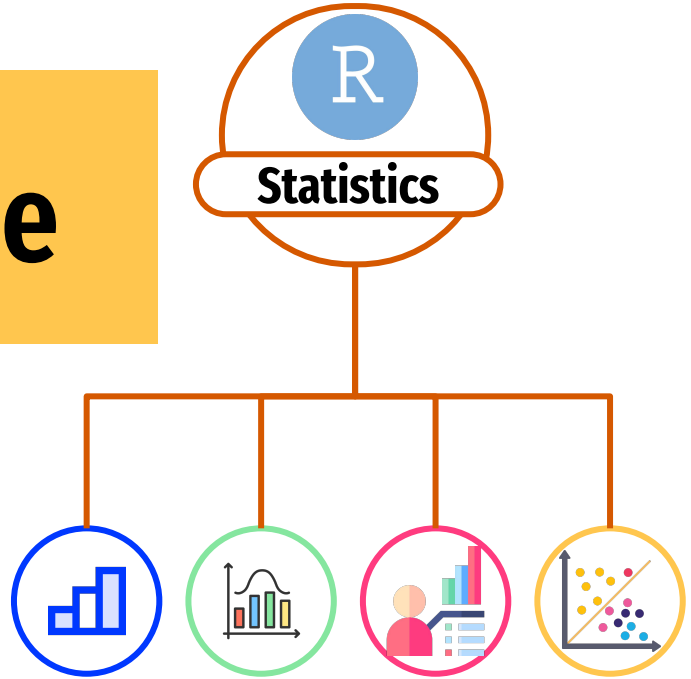
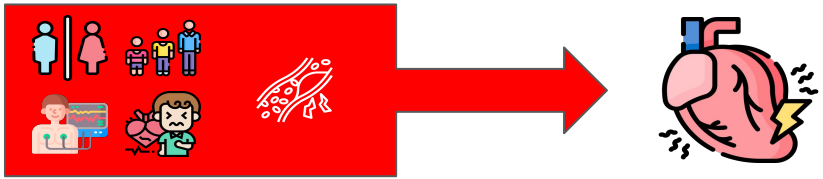
Predicting Heart Disease

Health Science Data





Group 4: Eisha Segaran, Seazer Wu, Seline May, Megan Soo, Melissa van Loenen

** all icons in the presentation are from Flaticon (cited at the end)*

Our Data

Data Repository →

Kaggle

- *Hospital admissions* reporting chest pain
- Sample is **NOT** a subset of the **general population** or **general hospital admissions**
- ☆ 28 years is the minimum age, biased towards males

Sources →

Combined 5 Datasets

- Cleveland
- Hungary
- Switzerland
- VA Long Beach
- Stalog Data Set (UCI)

Rows →

Observations

- **1190** total observations
- 272 duplications
- 172 counts with 0 cholesterol → biologically impossible, so values were removed
- Final dataset: **746** observations

Variables ♺

Total Variables: **12**
Analyzed: **6**

1. *Age*
2. *Sex*
3. *Serum cholesterol, i.e. the total amount of cholesterol that is in a person's blood*
4. *Resting Electrocardiogram (ECG) Results, i.e. a non-invasive test that is used for heart arrhythmia detection*
5. *Exercise-induced Angina, i.e. heart pain brought on by exercise*
6. *Heart disease, i.e. includes various heart conditions*

Variables of Interest

Independent Variables

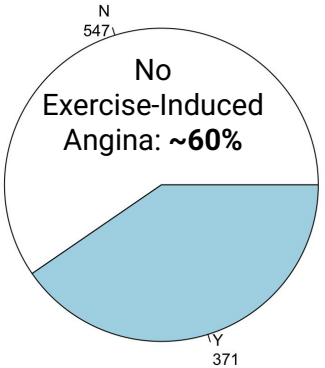
- **Sex** — Binary → (F) Female/Male (M)
- **Age** — Continuous → Ranges from 28-77 years old
- **Serum Cholesterol** — Continuous → Ranges from 85-603; Normal levels are < 200 milligrams per deciliter
- **Resting Electrocardiogram** — Categorical → Normal ECG Results, ST-T Abnormality, Left Ventricular Hypertrophy (LVH)

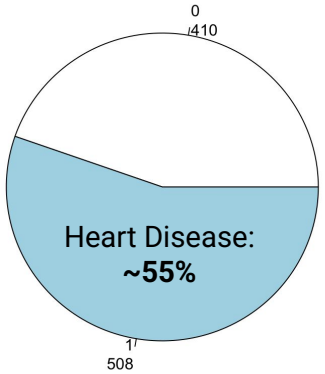
Used as Both Independent and Dependent

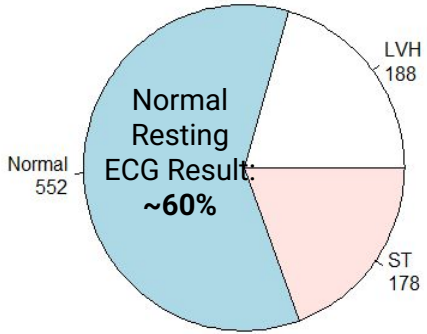
- **Exercise-Induced Angina** — Binary → Yes/No

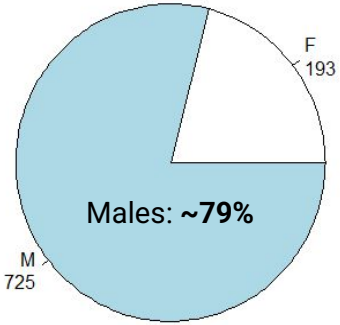
Dependent Variable

- **Heart Disease** — Binary → Yes/No









Research Question: Predicting Heart Disease

ADDRESSING KNOWLEDGE GAPS

Question 1: Can ***age and sex*** predict an **exercise-induced angina**?

WHY?

Age & sex are known predictors of heart disease & exercise-induced angina
(Rodgers et al., 2019)

- Requires assessment of strength of these predictors in multinational data

Question 2: Can ***cholesterol levels and resting ECG results*** be predictive of an **exercise-induced angina**?

WHY?

Cholesterol & abnormal ECG results are predictors of heart disease & exercise-induced angina
(Ekelund et al., 1989)

- Investigating a sample of patients already reporting with various types chest pain
- Requires assessment of strength of these predictors in multinational data

Question 3: Does the occurrence of an ***exercise-induced angina*** predict **heart disease**?

WHY?

Exercise-Induced Angina linked to coronary artery disease
(Harvard Health Publishing, 2021)

- Investigating a sample of patients already reporting with various types chest pain
- Requires assessment of strength of these predictors in multinational data

MODELS & METHODOLOGY

Principal Component Analysis: Allows observation of trends and relationships

Question 1: Given that age and sex are both predictors of heart disease, can ***age and sex*** predict an **exercise-induced angina**?

Question 2: Can ***cholesterol levels and resting ECG results*** be predictive of an **exercise-induced angina**?

Question 3: Does the occurrence of an ***exercise-induced angina*** predict **heart disease**?

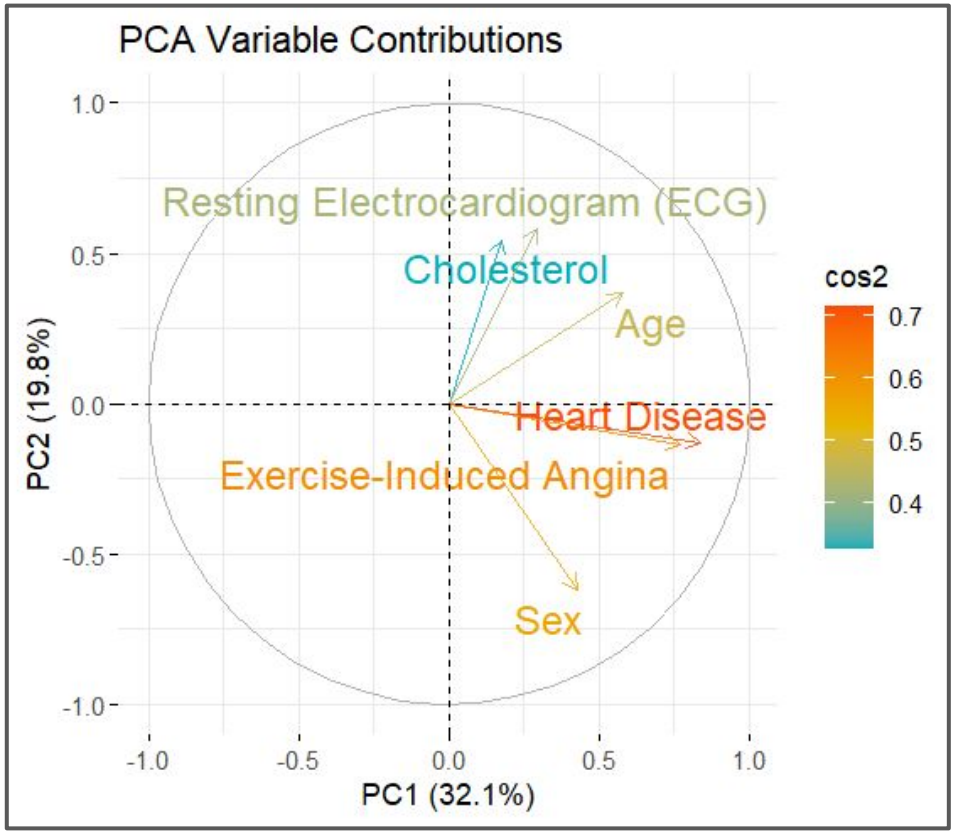
Logistic ANCOVA & Logistic Regression

- Binary Dependent Variables: Heart Disease & Exercise-Induced Angina
- Independent variables are both continuous and categorical

Test Nested Model Pairs

- Use AIC to predict best model
- Identify interactions between selected independent variables

Risk Factors of Heart Disease



- PC1 and PC2 together explain 51.9% of the total variance.

Relatively Positive Correlation

- Heart Disease, Exercise-induced angina, Sex
- Explains majority of variance in PC1

Moderate Positive Correlation

- Age moderately contributes to both PC1 and PC2.

Relatively Weak Correlation

- Resting ECG and Cholesterol
- Weaker contribution to PC1 and PC2

Question 1: Sex & Age

Given that age and sex are both predictors of heart disease, can age and sex predict an exercise-induced angina?

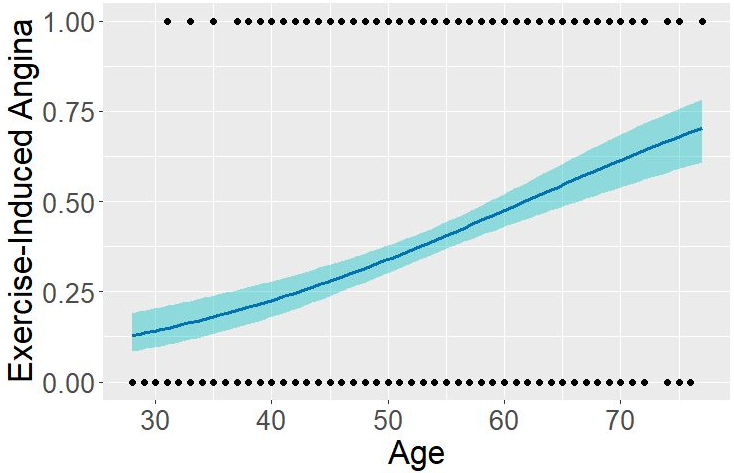
Sex × Age Interaction

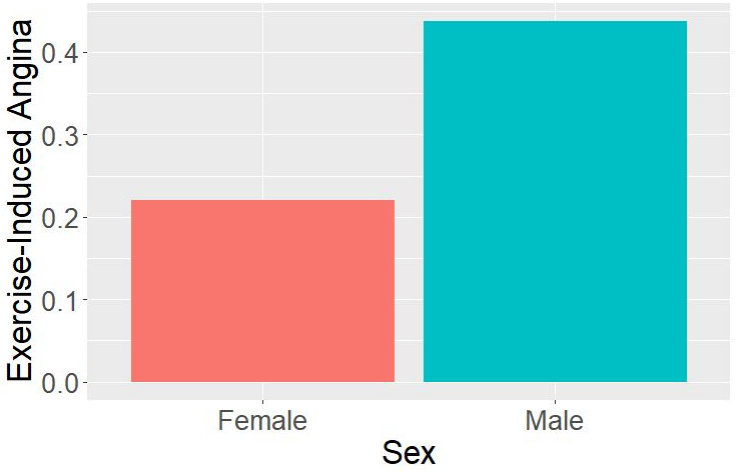
P-value → **Not significant**

Can use simpler model

Exercise-Induced Angina

- Chi-square test
  - P-value for sex < 0.001
  - P-value for age < 0.001



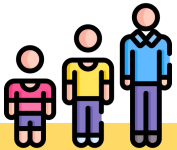


Risk of Exercise-Induced Angina

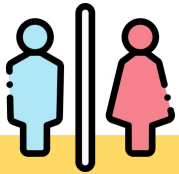
Men are **2.77x** more likely

1.06x more likely per added year of age

P-value < 0.001
Significant



Question 1: Sex & Age



Given that age and sex are both predictors of heart disease, can age and sex predict an exercise-induced angina?

Interaction Model

- Age explains the most variability
- **Age and sex do not have significant effects**

No Interaction

- Sex explains the most variability
- **Age and sex have significant effects**

AIC and ANCOVA

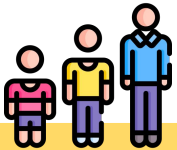
- No interaction → AIC value = 925.67
- Interaction → AIC value = 926.37

difference <2: two models are not significantly different

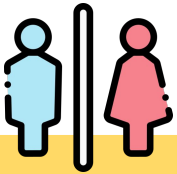
ANCOVA p-value = 0.2506

Model with* no interaction *is taken:

- ***Age and sex are insignificant in interaction model***
- ***ANCOVA tells us they are significant***

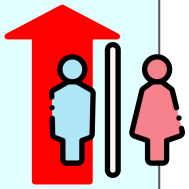


Question 1: Sex & Age

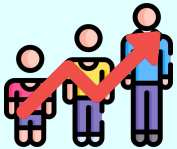


Given that age and sex are both predictors of heart disease, can age and sex predict an exercise-induced angina?

Past Literature



- Coronary heart disease incidence was **300x higher in men** in a random sample
  - ~14,000 people aged 25-64
  - Roughly equal sex distribution
  - Excluded those with previous history of myocardial infarction *(Jousilahti et al. 1999)*
- Increased age increases heart disease risk *(Jousilahti et al., 1999)*
  - More profoundly in women
    - Synergistic non-additivity



- Most at risk of cardiac events: **older males** *(Rodgers et al., 2019)*

Results

- No significant interactions between sex and age in this dataset
  - Does not corroborate past literature
    - Likely due to the study sample have already reported chest pain
      - **Not representative** of general population

- **Sex** is a **stronger** predictor than age
  - May be due to ~79% of the sample being male
  - *Jousilahti et al.* had a roughly even proportion of sexes

Question 2: Cholesterol and Resting ECG

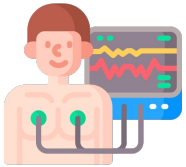
Can cholesterol levels and resting ECG results be predictive of an exercise-induced angina?

Interaction

P-value < 0.05 (*)

- Significant
- Suggests complex model is better

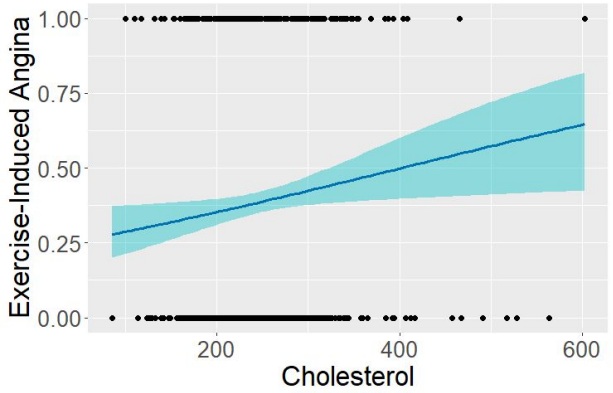
But AIC says otherwise!

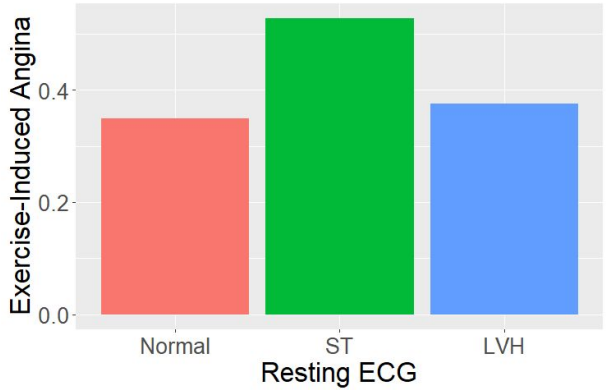


Exercise-Induced Angina

Chi-Squared Test

- **Cholesterol** (p = 0.03978) and **its interaction with resting ECG** (p = 0.03313) are significant predictors of exercise-induced angina.
- **Resting ECG** alone (p = 0.12206) is **not** a significant predictor.





Question 2: Cholesterol and Resting ECG

Can cholesterol levels and resting ECG results be predictive of an exercise-induced angina?

Interaction Model

- Resting ECG has a significant effect
- Cholesterol has a significant effect
- **Interaction is significant**

No interaction

- Resting ECG has a significant effect
- Cholesterol has a significant effect

AIC and ANCOVA

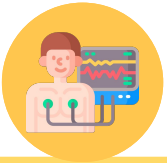
- No interaction → AIC value = 983.12
- Interaction → AIC value = 980.36

ANCOVA p-value < 0.05

Interaction model is taken:

- ***Significantly smaller AIC value, so it should be used.***
- ***Aligns with the previous ANCOVA results which say the interaction is significant.***

Question 2: Cholesterol and Resting ECG



Can cholesterol levels and resting ECG results be predictive of an exercise-induced angina?

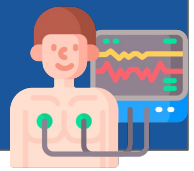
Past Literature

- **Abnormal ECG results** are a considerable predictor of heart disease in men, especially those with hypercholesterolemia, a condition where one has too much bad cholesterol/low-density lipoprotein/LDL *(Ekelund et al., 1989)*.

- Relationship between exercise-induced angina and cholesterol levels/ECG results
    - Exercise-induced angina is often acknowledged as a predictor of heart disease and serum cholesterol is a measurement of total cholesterol levels (good/HDL and bad/LDL).

Results

- Interaction between cholesterol levels and resting ECG **is significant.**
    - Corroborates past studies by **Ekelund et al**.
- While cholesterol alone **is** a significant predictor, resting ECG independently **is not** (according to ANCOVA).



Question 3: Exercise-Induced Angina

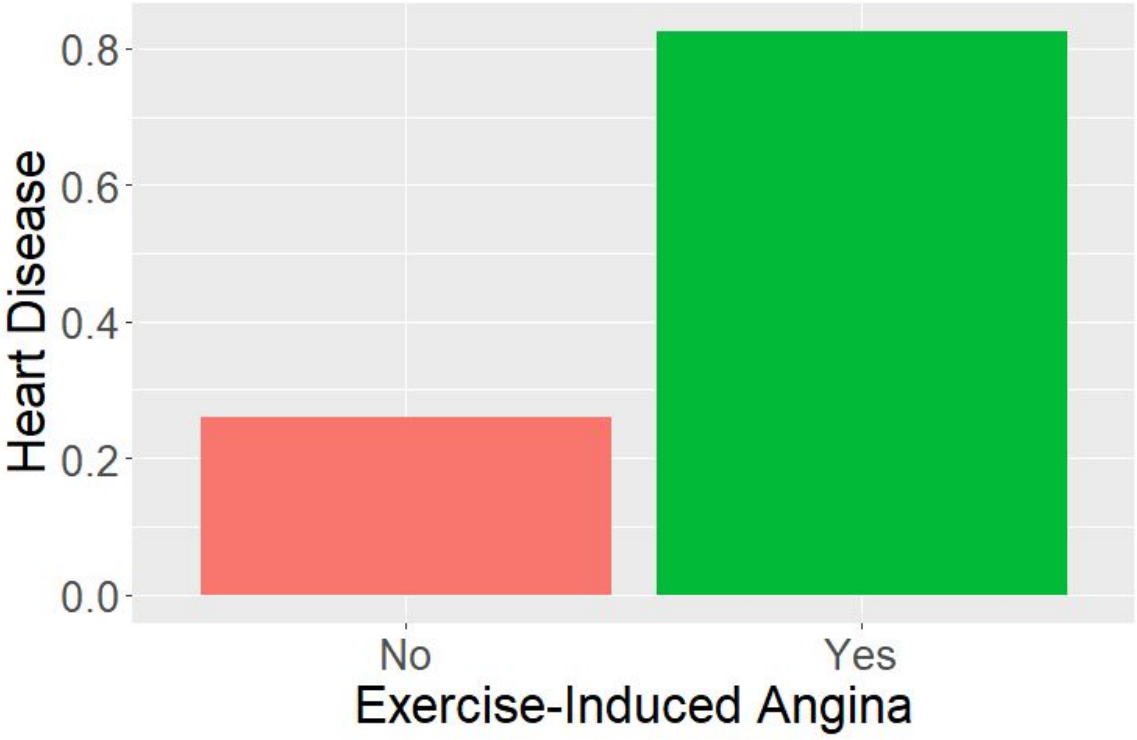
Does the occurrence of an exercise-induced angina predict heart disease?

Risk of Heart Disease

People with exercise-induced angina are **13.5x** more likely to get heart disease

P-value <0.001

Statistically significant



Question 3: Exercise-Induced Angina

Does the occurrence of an exercise-induced angina predict heart disease?

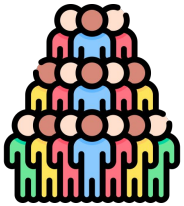
Past Literature

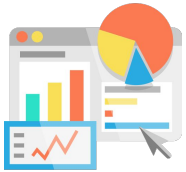
- Among typical risk factors spanning both the psychological and physiological, the **strongest individual predictor** of coronary heart disease was a history of exercise-induced angina *(Bekkouche et al., 2014)*.
- Therefore, we want to see if exercise-induced angina alone is able to predict the occurrence of heart disease, separate from other interacting variables.

Results

- Those with **history of exercise-induced** angina are **significantly** more likely to have **heart disease**
  - This corroborates past literature
  - Suggests that exercise-induced angina is a **considerable predictor** of heart disease

Possible Improvements to Experimental Design





Better Representation of General Population

- Roughly equal distribution of males and females to observe a true effect of sex
- Use stratified random sampling to ensure accurate and **proportionate** representation of sexes within general population

More Lifestyle Variables

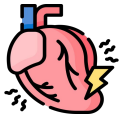
Possibility of confounding variables in multinational data

- Diet is known to influence heart disease
- Certain ethnicities are at greater risk of heart disease
- Timing of the exercise-induced angina should have been recorded
- Useful for determining temporal links

Interactions

- Determine if interactions between serum cholesterol levels and resting ECG results are significant
- AIC and ANCOVA provided conflicting outcomes

Conclusion



Predicting Heart Disease

01 Given that age and sex are both predictors of heart disease, can age and sex predict an exercise-induced angina?

Yes, individually as separate predictors

02 Can cholesterol levels and resting ECG results be predictive of an exercise-induced angina?

Yes to cholesterol and its interactions with resting ECG, but not resting ECG by itself

03 Does the occurrence of an exercise-induced angina predict heart disease?

Yes

References

Angina: Symptoms, diagnosis and treatments. (2021, September 21). Harvard Health Publishing: Harvard Medical School. https://www.health.harvard.edu/heart-health/angina-symptoms-diagnosis-and-treatments#:~:text=September%2021%2C%202021,angina%20can%20be%20very%20reassuring

Bekkouche N., Wawrzyniak AJ., Whittaker KS., Ketterer, KW., & Krantz, DS. (2014). PSYCHOLOGICAL AND PHYSIOLOGICAL PREDICTORS OF ANGINA DURING EXERCISE-INDUCED ISCHEMIA IN PATIENTS WITH CORONARY ARTERY DISEASE. *Psychosomatic medicine, 75(4)*, 413-421, https://doi.org/10.1097/PSY.0b013e31828c4cb4.

Buchner, S., Debl, K., Haimerl, J., Djavidani, B., Poschenrieder, F., Feuerbach, S., Riegger, G. A., & Luchner, A. (2009). Electrocardiographic diagnosis of left ventricular hypertrophy in aortic valve disease: Evaluation of ECG criteria by Cardiovascular Magnetic Resonance. *Journal of Cardiovascular Magnetic Resonance*, *11*(1), 18. https://doi.org/10.1186/1532-429x-11-18

Casas, R., Castro-Barquero, S., Estruch, R., & Sacanella, E. (2018). Nutrition and Cardiovascular Health. *International Journal of Molecular Sciences, 19(12),* 3988, doi: 10.3390/ijms19123988.

Cleveland Clinic. (2024). *Cholesterol numbers in the heart-healthy range can help lower your risk of heart disease.* Cleveland Clinic. Retrieved from https://my.clevelandclinic.org/health/articles/11920-cholesterol-numbers-what-do-they-mean.

DeVon, H. A., Mirzaei, S., & Zègre‐Hemsey, J. (2020). Typical and atypical symptoms of acute coronary syndrome: time to retire the terms?. *Journal of the American Heart Association*, *9*(7), e015539, doi: 10.1161/JAHA.119.015539.

Ekelund, L., Suchindran, CM., McMahon, RP., Heiss, G., Leon, AS., Romhilt, DW., Rubenstein, CL., Probstfield, JL., & Ruwitch, JF. Coronary heart disease morbidity and mortality in hypercholesterolemic men predicted from an exercise test: The lipid research clinics coronary primary prevention trial. *Journal of the American College of Cardiology, 14(3)*, 556-563, https://doi.org/10.1016/0735-1097(89)90092-2

fedesoriano. (September 2021). Heart Failure Prediction Dataset. https://www.kaggle.com/fedesoriano/heart-failure-prediction.

Flaticon. (2025). flaticon.com

Heart disease. (n.d.). Mayo Clinic. https://www.mayoclinic.org/diseases-conditions/heart-disease/symptoms-causes/syc-20353118

References

How Race and Ethnicity Impact Heart Disease. (n.d.). Cleveland Clinic. https://my.clevelandclinic.org/health/articles/23051-ethnicity-and-heart-disease.

Huizen, Jennifer. (2021, September 22). *What is serum cholesterol?* Medical News Today. https://www.medicalnewstoday.com/articles/321519#:~:text=Cholesterol%20is%20a%20type%20of,and%20triglycerides%20in%20the%20blood

Jousilahti, P., Vartiainen, E., Tuomilehto, J., & Puska, P. (1999). Sex, age, cardiovascular risk factors, and coronary heart disease: a prospective follow-up study of 14 786 middle-aged men and women in Finland. *Circulation, 99(9),* 1165-1172, doi: 10.1161/01.cir.99.9.1165.

Nonanginal chest pain: How it may feel, causes, and more. (2023, March 9). Medical News Today. https://www.medicalnewstoday.com/articles/nonanginal-chest-pain

Rodgers, JL., Jones, J., Bolleddu, SI., Vanthenapalli, S., Rodgers, LE., Shah, K., Karia, K., & Panguluri, SK. (2019). Cardiovascular Risks Associated with Gender and Aging. *Journal of Cardiovascular Development and Disease, 6(2),* 19, https://doi.org/10.3390/jcdd6020019

RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL http://www.rstudio.com/.

The T-Wave: physiology, variants, and ECG features. (2025). Retrieved from https://ecgwaves.com/the-t-wave-physiology-variants-and-ecg-features/.

U.S. National Library of Medicine. (2016, April). *Preoperative Tests (Update): Routine Preoperative Tests for Elective Surgery.* National Library of Medicine | National Center for Biotechnology Information . https://www.ncbi.nlm.nih.gov/books/NBK367910/