MM958: Statistical Modelling and Analysis – Research Abstract

Patrick Montgomery

University of Strathclyde

# Introduction

Cardiovascular diseases (CVDs) are the leading cause of death globally, taking an estimated 17.9 million lives each year (World Health Organisation, 2024). Coronary heart disease is the most diagnosed heart disease worldwide and is one of the leading causes of cardiovascular disease-related morbidity and mortality. A build-up of plaque and fats narrow the arteries, decreasing the blood flow causing chest pain, shortness of breath and in certain cases heart attacks. (Mayo Clinic, 2024) This study aimed to evaluate the predictive capability of multiple risk factors for CHD and examine the underlying patterns of variation in these risk factors using both traditional statistical modelling and dimensional reduction techniques.

# Methods

A retrospective analysis was conducted on a dataset of male subjects (n=422) from Western Cape, South Africa. Eight potential risk factors were analysed: systolic blood pressure, tobacco use, LDL cholesterol, adiposity, Type-A behaviour, obesity, alcohol consumption, and age. The analysis employed logistic regression with bidirectional stepwise selection using AIC criterion for model optimisation. Principal Component Analysis (PCA) was performed to explore the underlying structure of risk factors. The data was split into training (80%) and test (20%) sets for model validation.

# Results

The optimised logistic regression model identified four significant predictors of CHD: tobacco use (OR=1.09, 95% CI: 2.81-3.19), LDL cholesterol (OR=1.19, 95% CI: 2.85-3.86), Type-A behaviour (OR=1.04, 95% CI: 2.76-2.92), and age (OR=1.06, 95% CI: 2.81-2.95). At the optimal classification threshold of 0.35, the model achieved 74.2% accuracy on test data, with sensitivity of 70.4% and specificity of 75.8%. Principal Component Analysis revealed three major components explaining 64.05% of total variance: a general "health status" component (35.72%), a "lifestyle habits" component (15.04%), and a "behavioural" component (13.29%).

# Conclusion

The study demonstrates that CHD risk can be effectively predicted using a reduced set of risk factors, with LDL cholesterol and age emerging as particularly strong predictors. The complementary PCA findings suggest that while the relationship between risk factors and CHD is complex, it can be largely captured by three underlying dimensions. The model's balanced predictive performance on test data suggests good generalizability, though the presence of unexplained variance indicates that additional factors may influence CHD risk. These findings are particularly relevant in the South African context, where CHD represents a significant and growing public health concern, with an estimated 38% increase in CHD cases from 2000 to 2016 (Abdelatif, Peer, & Manda, 2021). The study's results align with the urgent need for improved risk assessment strategies in clinical settings, especially given that risk factors for CVDs are largely modifiable and up to 80% of premature heart attacks could be preventable. While our model shows promise for risk stratification, its implementation should be considered within the broader context of South Africa's healthcare infrastructure, where there is a recognised need for well-structured CHD registries and improved diagnostic capabilities. Future research should focus on validating these predictive models across different population groups and healthcare settings to ensure their effectiveness in this high-risk population.

References

Abdelatif, N., Peer, N., & Manda, S. (2021). National prevalence of coronary heart disease and stroke in South Africa from 1990–2017: a systematic review and meta-analysis. *Cardiovascular Journal of Africa*, 156-160.

Mayo Clinic. (2024, November 26). *Coronary artery disease*. Retrieved from Mayo Clinic: https://www.mayoclinic.org/diseases-conditions/coronary-artery-disease/symptoms-causes/syc-20350613#:~:text=CAD%20happens%20when%20coronary%20arteries,or%20even%20a%20heart%20attack.

World Health Organisation. (2024, November 26). *World Health Organisation*. Retrieved from Cardiovascular Diseases: https://www.who.int/health-topics/cardiovascular-diseases#tab=tab\_1