Derivation and Validation of a Clinical Tool to Predict Obstructive Coronary Artery Disease Among Patients with Coronary Calcium Score of Zero

By:

Ali Alshahrani

Thesis submitted to faculty of graduate and post-doctoral studies in partial fulfillment of requirements for MSc in Epidemiology

School of Epidemiology, Public Health and Preventive Medicine, Faculty of Medicine, University of Ottawa

Abstract

Coronary artery disease (CAD) is associated with significant morbidity and mortality. Coronary artery calcification (CAC) indicates presence of CAD. Absence of CAC is associated with very low risk of having CAD but not equal to zero. In this study, we aim at developing a clinical prediction tool to predict presence of obstructive CAD among patients with zero calcium score. We developed two models. A full prespecified model with 7 variables based on input from clinical experts, and a reduced model with 4 variables based on univariate screening. Both models showed an acceptable performance (c-statistics of 0.68 for both). Both models performed well when validated, externally for the full model and internally for the reduced one. We derived a clinical risk score of 20 points from the full model. We found that a score threshold of ≥ 14 is associated with presence of obstructive CAD with positive likelihood ratio of 5.5.

Executive Summary

Background

Coronary artery calcification (CAC) is a marker of underlying atherosclerotic coronary artery disease (CAD) and commonly quantified using Agatston score. Presence of CAC is associated with increased risk of cardiovascular events and overall mortality. Absence of CAC is associated with a lower prevalence of cardiovascular events, mortality and obstructive CAD. However, there is a residual risk of obstructive CAD among patients with zero calcium score.

Objectives

Our main objective was to derive a clinical tool that can be added to Agatston score =0 to predict the presence of obstructive CAD with high accuracy using two distinct approaches for development of clinical prediction tools. In addition, we sought to compare the developed models methodologically and assess the impact of that on the clinical utility of these models.

Methods

We developed two clinical models to predict obstructive CAD using two distinct approaches retrospectively from a cohort of 4,903 consecutive patients with an Agatston score =0 from the University of Ottawa Heart Institute Cardiac CT Registry. In the first approach, the model pre-specification approach, we included a pre-specified list of predictors to be included in a multivariable logistic clinical prediction model (the full model). In the second approach, the univariate screening approach, we included predictors in the multivariable model based on the statistical significance of the association between the predictor and the outcome in the univariable analysis (the reduced model). Our outcome was defined as luminal stenosis of any coronary artery segment of $\geq$ 50% on coronary computed tomographic angiography (CTA). The score derived from the full model was validated using a cohort of 4,290 patients with an Agatston score =0 from a multi-center international registry of coronary CTA. We validated the score derived from the reduced model internally using the bootstrap technique. We assessed the discriminative performance for both models using the Area Under the Curve (AUC) of the Receiver Operating Characteristics (ROC) curve and the Hosmer-Lemeshow statistic was used for assessment of goodness of fit.

Results

We derived a score of 7 variables: age, sex, typical chest pain, hyperlipidemia, hypertension, family history and diabetes mellitus from the full model. The score was robust with an area under the curve of 0.68 in the derivation cohort and 0.69 in the validation cohort. The Hosmer-Lemeshow statistic showed that the model fit well both derivation and validation datasets (p value >0.05). Patients were divided into three risk groups based on the score, low (≤6), intermediate (7-13) and high (≥ 14). Patients who score ≤6 have a low negative likelihood ratio of 0.42 and a negative predictive value of 99% for obstructive CAD, while those who score ≥14 have a positive likelihood ratio of >5.5 for obstructive CAD. The proportion of patients in the low risk group with the outcome in the validation cohort was < 3%.

The score derived from the reduced model consisted of 4 variables: sex, typical chest pain, hyperlipidemia and baseline systolic blood pressure. The model had an area under the curve of 0.68 and the Hosmer-Lemeshow statistic showed that the model fit well the dataset. The model had an acceptable performance up on internal validation. A score of 3 on a scale of 5 points was chosen as a cut point for ruling out the outcome of obstructive CAD with a specificity of 90% and negative predictive value of 98%. Both models seem to perform better to rule out the outcome.

Conclusion

We developed two clinical scores using two distinct approaches to predict obstructive CAD in patients with zero calcium score with good performance when validated. Both scores can be used to rule out obstructive CAD in patients with an Agatston score =0 with high confidence and comparable performance. However, the score derived from the full model seem to be more clinically relevant.

Acknowledgement

By completing this work, I must acknowledge the great support and help provided by my supervisor, Dr. George Wells and by Dr. Benjamin Chow, the director of Advanced Cardiac Imaging Fellowship program who facilitated my work with the database and was my clinical mentor for this project.

This work would not be possible without the tremendous help of my wife, Nora; who has been encouraging and supportive all the time during my journey.

Table of Contents

Abstract ii
Executive Summary iii
Acknowledgement v
Table of Contents vi
List of Tables xii
List of Figures xiv
CHAPTER 1: INTRODUCTION 1
1.1: Importance of Prediction of Obstructive CAD 1
1.2 Objectives of The Thesis 2
1.3: Description of the Thesis Chapters 3
CHAPTER 2: BACKGROUND 5
2.1: Coronary Artery Disease 5
2.1.1: Risk Factors for CAD 5
2.1.2: Diagnosis of CAD 6
2.1.3: Coronary Artery Calcification 7
2.1.3.1: Clinical Significance of Coronary Artery Calcification 7
2.1.3.2: Clinical Significance of Calcium Score of Zero 8
2.1.3.3: Coronary Calcification and Cardiovascular Risk Modification 10
2.2: Clinical Prediction Models 10
2.2.1: Overview 10
2.2.2: Fundamentals of Clinical Prediction Models' Development 11
2.2.2.1: Role of Study Design in Clinical Prediction Models 11
2.2.2.3: Outcomes in Clinical Prediction Rules 13
2.2.2.4: Analytical Methods for Clinical Prediction Rules 14
2.2.3: Methodological Considerations for Clinical Prediction Models 15
2.2.3.1: Over-fitting 15
2.2.3.2: Mechanisms and Handling of Missing Data 16
2.2.3.3: Handling Continuous Predictors 19
2.2.4: Measurement of Model Performance 20
2.2.5: Validation of Developed Models 20
2.2.5.1: External Validation 20

2.2.5.2: Internal Validation 21
2.2.6: Model Presentation 23
2.3: Clinical Prediction Models for Coronary Artery Disease 24
2.3.1: Overview of Models Predicting CAD 25
2.3.1.1: Derivation Cohorts, Setting and Design 25
2.3.1.2: Predictors Selection and Model Specification 26
2.3.1.3: Definition of Predicted Outcome 31
2.3.1.4: Over-fitting 33
2.3.1.5: Missing data 34
2.3.1.6: Model performance 34
2.3.1.7: Validation 35
2.3.1.8: Summary of Published Models 35
2.3.2: Clinical Utility of Calcium Score of Zero in Prediction of CAD 36
CHAPTER3: METHODS 38
3.1: Overview of Standard Methodology for Developing Clinical Prediction Models 38
3.2: Overview of Methods Used for This Study 39
3.3: Data Source for Model Derivation 40
3.3.1: Definition of Variables 42
3.3.2: Image Acquisition and Interpretation 42
3.3.3: Definition of the Outcome 43
3.3.3.1: Diagnostic Accuracy of CTA for Diagnosis of Coronary Artery Disease 43
3.4: Data Preparation and Explanatory Analysis 46
3.4.1: Data Exploration and Variables Recoding 47
3.4.2: Statistical Summarization 50
3.4.3: Missing Data 51
3.5: Overview of Model Derivation Approaches 51
3.5.1: Model Pre-specification (Full Model) Approach 51
3.5.2: Univariate Screening (Reduced Model) Approach 52
3.5.2.1: Univariate Predictors Screening 52
3.5.2.2: Step-wise Selection of Predictors 53
3.5.2.3: Issues of Statistically-Driven Predictors Selection 53
3.6: Model Development - Pre-Specification Approach (Full Model) 54
3.6.1: Model Pre-Specification 54
3.6.2: Multivariable Modelling 54

3.6.3: Assessment of Model Performance 55
3.7: Model Presentation 55
3.7.1: Development of the Scoring System 56
3.7.1.1: Dealing with Continuous Predictors in The Score 56
3.7.2: Assessment of Score Performance 56
3.7.3: Diagnostic Characteristics of Score Threshold 57
3.8: External Validation of the Score Derived from the Full Model 57
3.8.1: Overview of Validation Cohort 58
3.8.2: Data Preparation of Validation Dataset 58
3.8.3: Exploratory Analysis and Variables Recoding 58
3.8.4: Dealing with Missing Values in Validation Dataset 59
3.8.4.1: Multiple Imputations of Missing Values in Final Validation Dataset 60
3.8.5: Score Validation 60
3.8.5.1: Relatedness of Derivation and Validation Cohorts 61
3.8.5.2: Assessment of Score Performance in Validation Dataset 61
3.8.5.3: Diagnostic Characteristics of the Score Thresholds in Validation Cohort 61
3.9: Model development - Reduced Model 62
3.9.1: Multiple Imputations 62
3.9.2: Univariate Screening 62
3.9.3: Multivariable Modelling- Reduced Model 63
3.9.4: Assessment of Model Performance 63
3.10: Internal Validation- Reduced Model 63
3.10.1: Bootstrapping Technique 63
3.11: Model Presentation 64
3.11.1: Development of the Scoring System 64
3.11.1.1: Dealing with Continuous Predictors in The Score 64
3.11.2: Assessment of Score Performance 65
3.11.3: Diagnostic Characteristics of Score Threshold 65
CHAPTER 4: RESULTS 66
4.1: Derivation Dataset Preparation 66
4.1.1: Characteristics of Participants in Derivation Dataset 66
4.1.3: Summary Statistics of Predictors and Outcome 68
4.2: Results of Full Model (Prespecified Model Approach) 69

4.2.1: Model Specification 69
4.2.2: Multivariable Modelling 74
4.2.2.1: Initial Multivariable Logistic Regression Model 74
4.2.2.2: Final Multivariable Logistic Regression Model 77
4.2.2.3: Interaction 77
4.2.3: Model Performance 77
4.3: Model Presentation- Score Development 80
4.3.1: Scoring System Points 80
4.3.2: Score Performance 83
4.3.3: Operating Characteristics of Risk Score Levels 83
4.3.4: Risk Classification of Score Thresholds 83
4.3.4.1: Risk Re-classification of Intermediate Risk Group 84
4.3.5: Accuracy of Risk Classification 85
4.3.6: Proposed Clinical Algorithm Based on the Risk Score 85
4.4: External Validation of the Risk Score 85
4.4.1: Results of Validation Dataset Preparation 85
4.4.2: Exploratory Analysis of Preliminary Validation Dataset 89
4.4.3: Predictors and Outcome in Preliminary Validation Dataset 91
4.4.3.1: Missing Data in Preliminary Validation Cohort 91
4.4.4: Final Validation Dataset 91
4.4.4.1: Final Validation Dataset vs Data from Excluded Centers 94
4.4.5: Score Validation 94
4.4.5.1: Relatedness of Derivation and Validation Datasets 94
4.4.5.2: Imputation of Missing Values in the Final Validation Dataset 97
4.4.5.3: Score Performance in the Final Validation Dataset 99
4.4.5.3: Diagnostic Characteristics of Risk Thresholds in Validation Dataset 101
4.4.5.4: Score Performance in the Derivation vs Validation Datasets 103
4.4.5.6: Score Calibration in Derivation vs Validation Datasets 103
4.5: Results of Reduced Model Development (Univariate Screening Approach) 105
4.5.1: Univariate Screening 105
4.5.2: Multivariable Modelling (Reduced model) 107
4.5.2.1: Multiple Imputation 107
4.5.2.2: Variables Selection (Reduced-model approach) 109

4.5.2.3: Final Multivariable Prediction Model (Reduced model) 109
4.5.3: Model Performance 109
4.6: Internal Validation of Reduced Model 111
4.6.1: Results of Variable Selection in Bootstraps 111
4.6.2: Optimism and Overfitting 111
4.6.3: Pooled Estimates of Final Model in Bootstraps 112
4.7: Model Presentation (Reduced Model) 113
4.7.1: Risk Score Development 113
4.7.2: Operating Characteristics of Score Thresholds 114
CHAPTER 5: DISCUSSION 119
5.1: Summary of Findings 119
5.2: Interpretation of the Results in Context of Available Studies 123
5.2.1: Available Tools to Predict Obstructive CAD 123
5.2.2: Diagnostic Utility of Zero Calcium Score 126
5.3: Methodological Aspects of Models Development 127
5.3.1: Methods of Models' Specification 127
5.3.2: Modelling Continuous Predictors 128
5.3.3: Dealing with Missing Data 129
5.3.4: Validation of the Developed Models 129
5.3.5: Classification Thresholds of the Developed Scores 130
5.4: Limitations 131
5.5: Strengths 131
5.6: Future Research 132
5.7: Conclusion 132
References 134
Appendices 151
Appendix 1: TRIPOD Statement Checklist 151
Appendix 2: Research Ethics Board Approval 152
Appendix 3: Registry form for UOHI Cardiac CT database 153
Appendix 4: Report Form of UOHI Cardiac CT Registry 156
Appendix 5: Survey for Model Predictors Pre-specification 157
Appendix 6: Coronary CT Angiography Evaluation for Clinical Outcomes: An International Multicenter (CONFIRM) Registry Centers 158
Appendix 7: Multivariable Model with Interaction Term (Full Model) 159

Appendix 8: Trace Plots of the Age in the Imputed Final Validation Datasets 160
Appendix 9: Trace Plots of Continuous Variables in the Imputed Derivation Datasets (reduced model)..... 161
Appendix 10: ROC Curve of the Developed Score from the Full Model in the Preliminary Vlidation Dataset (complete case analysis)..... 162

List of Tables

Table 1: Brief description of published tools to predict CAD 28
Table 2: Methodological aspects of the published tools to predict CAD 32
Table 3: Summary of the methodological characteristics of the published tools 37
Table 4: Definition of the main clinical variables in the derivation dataset 44
Table 5: Candidate predictors for inclusion in the multivariable model 49
Table 6: Clinical Characteristics of derivation dataset from UOHI cardiac CT registry 70
Table 7: Summary statistics of the candidate predictors in the derivation dataset 72
Table 8:Summary of responses to survey for predictors pre-selection 75
Table 10: Initial multivariable regression model using model pre-specification (Full Model) 78
Table 11: Final multivariable clinical model using model pre-specification (Full Model) 78
Table 12: Model performance measures 79
Table 13: Point scoring system for prediction of obstructive coronary artery disease 81
Table 14: Levels of risk score and the corresponding predicted risk of CAD 82
Table 15: Operating characteristics of each score threshold in derivation dataset 86
Table 16: Proportion of patients in each risk category and the proportion of confirmed obstructive CAD 87
Table 17: Characteristics of patients in preliminary validation dataset 92
Table 18: Distribution and missing proportions of model predictors and outcome in preliminary validation dataset 93
Table 19: Missing proportions of our model predictors and outcome as stratified by the center in the preliminary validation dataset 95
Table 20: Characteristics of the final validation dataset 96
Table 21: Characteristics of patients in final validation dataset vs excluded centers 98
Table 22: Characteristics of case mix and outcome of derivation and validation datasets 100
Table 23: AUC of ROC and Hosmer-Lemeshow statistic for each imputed validation dataset 102
Table 24: Operating characteristics of the score at low and high risk thresholds in validation dataset 102
Table 25: Score performance in derivation and validation datasets 103
Table 26: Classification accuracy of the developed score in derivation and validation datasets 104
Table 27: Candidate predictors and association with outcome 106

Table 28: Final list of predictors to be included in a multivariable model based on univariate screening 108
Table 29: Variable selection for each imputed dataset by selection mode 110
Table 30: Model parameters of the final multivariable model by univariate screening 110
Table 31: AUC of ROC curve and Hosmer-Lemeshow statistic for the reduced model in each imputed derivation dataset 112
Table 32: Frequency of predictors selection in the bootstrap models 115
Table 33: Pooled estimates of multivariable model in bootstrap samples vs derivation dataset 115
Table 34: Risk score points derived from the final multivariable model using the univariate screening 116
Table 35: Risk score thresholds with corresponding risk of obstructive CAD 117
Table 36: Operating characteristics of risk score thresholds derived from the model developed using univariate screening 118
Table 37: Comparison of the developed models 122

List of Figures

Figure 1:Schematic presentation of the methodological flow for development of our two models ... 41
Figure 2: Flow chart of derivation dataset preparation ... 67
Figure 3:Area under the curve for ROC curve of the developed model (Full Model) ... 79
Figure 4: AUC of ROC for the score in the derivation dataset ... 84
Figure 5: Proposed algorithm for clinical use of the developed risk score ... 88
Figure 6: Flow chart of preparation of final validation dataset ... 90
Figure 7: AUC of ROC curve of the developed score in the first imputed validation dataset ... 101
Figure 8: Calibration plot of expected vs observed risk of the score in derivation and validation datasets ... 104
Figure 9: ROC of the reduced model in the first imputed derivation dataset ... 113
Figure 10: ROC for the risk score derived from the reduced model ... 117

CHAPTER 1: INTRODUCTION

In this thesis we derived and validated a clinical prediction model to predict the presence of obstructive coronary artery disease (CAD) in patients who have zero calcium score based on the coronary computed tomography angiography. In this chapter, we will discuss the importance of developing such a clinical prediction model, state the objectives of this thesis and highlight the chapters the thesis.

1.1: Importance of Prediction of Obstructive CAD

CAD is a common medical problem associated with a significant morbidity and mortality[1,2]. Coronary artery calcification as quantified by Agatston calcium score is a marker of underlying coronary artery disease associated with future cardiovascular events and mortality (see section 2.1). A calcium score of zero is associated with a favorable outcome with lower rates of cardiovascular events and mortality; however, obstructive CAD cannot be ruled out especially in symptomatic patients. This residual risk of having obstructive CAD in patients with calcium score of zero lead to uncertain diagnostic and prognostic value of a calcium score of zero.

There are several clinical prediction tools to predict presence of coronary artery disease based on clinical presentation of chest pain and risk factors for CAD. The available tools for prediction of coronary artery disease are limited by the heterogenous populations and variable definitions of the outcome. Most of the published models were derived and validated in cohorts with relatively high prevalence of CAD (see Section 2.3 for details). Only one model incorporated coronary artery calcium as a predictor, but none of these models were developed in the population with zero calcium score which has a lower risk of cardiovascular events and mortality. Other limitations to most of the existing models include: different population characteristics with major changes in prevention guidelines over time which impact the incidence rate of CAD; emergence of non-invasive and accessible diagnostic modalities including coronary computed tomographic angiography (CTA); and methodological limitations pertaining to management of missing data and model specification as we will discuss in the next chapter in details.

Use of coronary CTA has grown recently because of the accumulating evidence regarding its diagnostic accuracy compared to the gold standard modality (invasive coronary angiography (ICA)); prognostic added value and feasibility. The highest diagnostic performance for coronary artery CTA is in its ability to rule-out presence of CAD (high negative predictive value). In practice, quantification of calcium in the coronary arteries is usually done as a part of chest CTA prior to gated images of the coronary CTA. It has a lower radiation, easy to perform and does not need iodinated contrast.

The radiation dose used for coronary calcium scan is estimated to be 1-3 mSv for each scan, with an extra dose 1-3 mSv to perform a prospective coronary CTA (versus up to 13 mSv to perform a retrospective coronary CTA)[3,4] (see Section 3.3.2). Nowadays, with prospective coronary CTA as the predominantly used coronary CTA protocol, performing coronary calcium scan alone could reduce radiation exposure by 50%.

Given the inability to rule out obstructive CAD with calcium score of zero , non-generalizable and limited existing tools for prediction of coronary artery disease , availability and feasibility of calcium score quantification in today's practice and to reduce exposure to unnecessary extra radiation by obtaining a coronary CTA, we sought to derive a clinical prediction model to predict the presence of obstructive coronary artery disease among patients who underwent coronary calcium scan for evaluation of coronary artery disease and have a zero-calcium score. Accurate prediction of obstructive coronary artery disease based on clinical variables in patients with no coronary artery calcification, who have a very low risk of cardiovascular events and mortality, would help clinicians to identify patients with high probability of obstructive coronary artery disease. Identification of this group with zero calcium score and low probability of obstructive CAD would help preventing this group from undergoing extra unnecessary radiation; while, patients with high probability could proceed to get the appropriate testing.

1.2 Objectives of The Thesis

Our main objective of this work is to develop and validate a clinical model to predict the presence of obstructive CAD among patients who have zero calcium score as measured by Agatston score. Other objectives of this work are:

1. Development of two clinical models using two distinct approaches. The first approach we used was developing the model by pre-specification of model predictors based on input from matter experts. We called this model the Full Model and the approach used was called (Model Pre-specification Approach) for the sake of simplicity. For the second model we used an approach that utilizes the available data to drive the selection of the final model predictors through univariate analysis of all candidate predictors and including variables meeting a certain p value to enter a multivariable model. The final model is developed through an automated process of step-wise selection of variables with statistical significance. We called this model; the Reduced Model and the approach was called (Univariate Screening Approach).
2. Comparing the two developed models in term of discriminative performance and clinical utility.
3. Deriving a clinical prediction rule in form of a point score from both models to use in clinical practice.
4. External validation of the risk score developed using the full model in a non-overlapping dataset and internal validation of the reduced model through performing bootstrapping.
5. To examine the key statistical differences between the two developed models and the implication on the final model used to derive the score.

1.3: Description of the Thesis Chapters

The thesis is composed of 5 chapters. In this chapter (Chapter 1), we described the importance of developing a model to predict presence of obstructive CAD in patients with zero calcium score, the objectives of the thesis and highlights of the chapters of the thesis.

Chapter 2 presents a background about coronary artery disease, a review of the current clinical models to predict coronary artery disease and the available literature about the diagnostic and prognostic role of coronary artery clarification, and an overview of the clinical prediction models including the different methodological aspects of the development of clinical prediction models.

In Chapter 3, we present the methods followed to develop and validate our models including data preparation and summary statistics of the derivation dataset, methods of developing

our models using the two approaches with an overview of each method, method of deriving the clinical prediction rule (score), the descriptive statistics of the validation dataset including handling of missing values , the method of validation of the score derived from the full model in the external validation dataset and method of the internal validation of the model derived using the univariate screening approach.

Chapter 4 presents the results of summary statistics of derivation and validation datasets, results of different steps of multivariable modelling process by the two approaches and the validation results of both models. We also present in this chapter the developed clinical prediction rule and the diagnostic characteristics at each level of the point scoring system.

In Chapter 5, we summarize our findings of the final multivariable models, their performance and the derived clinical rule from each model. We compare the two models in term of final list of predictors, performance measures and the resulting rules. We discuss our findings in context of the published work pertaining to the clinical prediction models for obstructive CAD and the role of coronary artery calcification. We highlight some of the methodological issues regarding the approaches we used for development of our models, the impact of the key differences between the two newly developed models and the implications on the final model used to derive the score. In this chapter, we also discuss some of the strength and limitations of our study. Finally, we present our conclusion, knowledge gaps and future direction research.

CHAPTER 2: BACKGROUND

2.1: Coronary Artery Disease

Coronary artery disease (CAD) is the most common form of heart disease and results from a prolonged multifactorial process that ultimately lead to build up of atherosclerotic plaques in the coronary arteries. This process starts with a cascade of events includes lipid accumulation in vascular all cells, leukocyte recruitment and accumulation, formation of focal lesions, smooth muscle migration and death and calcification of newly formed plaques [5,6]. It is estimated to be present in 6.3% (7.4% for men and 5.3 for women) among adults ≥ 20 years old in United States, based on data collected between 2011 and 2014 by National Health and Nutrition Examination Survey (NHANES) [7]. The life time risk of developing CAD in adults with ≥ 2 risk factors is 37.5% for males and 18.3% for females [7]. Cardiovascular disease, including CAD; is the leading cause of death in 27% of men and 32% of women worldwide according to the World Health Organization Global Burden of disease update 2004[2].

2.1.1: Risk Factors for CAD

There are several factors that have been identified that contribute to the development of CAD. A study that evaluated men and women from three cohort studies found that approximately 90% of patients with CAD have a prior exposure to at least one of these risk factors: high blood pressure or current use of blood pressure lowering medications; high total cholesterol or use of cholesterol lowering medications; cigarette smoking; and diagnosis of diabetes mellitus[8]. In a case-control study that recruited participants from 52 countries, nine modifiable risk factors were identified as potential targets to control and hence reduce the risk of coronary artery events by 90% [9] . Six out of nine factors are associated with increased risk of myocardial infarction including hypertension, diabetes mellitus, abdominal obesity, raised Apo-lipoprotein B to Apo-lipoprotein A ratio, smoking and psychosocial stressors. Physical activity, daily consumption of fruits and vegetables and regular alcohol intake were associated with reduced risk of myocardial infarction. Poor physical capacity as defined by the inability to achieve a heart rate target and ST-segment depression on electrocardiogram during exercise doubled the risk of developing coronary artery

disease among men with 10-year predicted risk of 20%; while each metabolic equivalent (MET) increment reduced the risk of developing CAD by 13%[10].

Family history of premature coronary artery disease is defined as occurrence of coronary artery disease in male first-degree relative at age < 55 years or in female first degree relative at age < 65 years. It has been considered as a risk factor for CAD for long time. The significance of family history as a risk factor was observed among patients with parental family history that is not premature (i.e. occurrence at age>55 years)[11]. The effect of family history is explained partially by the fact that most of coronary artery disease risk factors are heritable like hypertension, high cholesterol levels and diabetes; and partially by the influence of family behaviors in diet, smoking and lack of physical activity[12–15]. There is a persistent 1.5 to 2-fold relative risk of family history even after adjusting for other coexisting risk factors in multiple studies of parents, siblings, twins and second degree relatives [11,16–18]. The importance of family history as a risk factor was demonstrated across multiple ethnic backgrounds including White Europeans, Japanese, Hispanics, and African Americans [19–21]. It was reported that 75% of patients with premature coronary artery disease have a family history of premature coronary artery disease, despite the prevalence of family history of premature coronary artery disease in general population which is estimated to range from 14 to 35% [22,23]. The family history of premature CAD is usually self-reported and subject to recall bias. Multiple analyses have addressed this issue and reported an overall excellent negative predictive value (0.90%) but fair positive predictive value of (26%-67%) for validated premature coronary artery disease death or myocardial infarction[18,22].

2.1.2: Diagnosis of CAD

Conventional coronary angiography is considered the gold standard diagnostic modality of coronary artery disease. It can visualize the lumen in two dimensions. This modality is limited by the risks of invasive procedure including mortality, arrhythmia, bleeding, vascular injuries and the excess amount of contrast and radiation used for anatomical assessment of coronary arteries [24]. Coronary computed tomography angiography (CTA) using a multi-detector has emerged as a non-invasive modality that is increasingly used for diagnosis of CAD. Several clinical studies have been conducted and showed a high diagnostic accuracy of the current coronary CTA technology. The area under receiver-operating curve (AUC) from multiple studies for diagnosing obstructive

CAD ranges from 0.93 to 0.98 [25–30] . In practice, coronary CTA is used commonly to rule out CAD in patients with low to intermediate probability of coronary artery disease[31,32].

2.1.3: Coronary Artery Calcification

Coronary artery calcification (CAC) on pathological specimens was found to be correlated with death from heart disease since renaissance-era [33]. Identification of CAC on fluoroscopy in the 1960s and 1970s was indicative of underlying obstructive coronary artery disease and adverse cardiac outcomes [34,35]. In the 1990s, quantification of coronary artery calcification on unenhanced computed tomography started and emerged as a non-invasive technique to identify clinical and subclinical coronary artery atherosclerotic disease[36,37]. Multiple methods for quantification of coronary artery calcification on cardiac computed tomography were developed and called calcium scores. These methods include Agatston score, volume score, mass score and others[33]. Agatston score was introduced in 1990 and remains the gold standard for quantification of CAC in clinical and research setting due to availability of abundant data across age, gender and races; as well as, large clinical studies of risk stratification are based on Agatston score [33].

2.1.3.1: Clinical Significance of Coronary Artery Calcification

Coronary artery calcification is associated with underlying coronary artery disease and cardiac events across different populations. In a large cohort study of around 11000 asymptomatic men and women, 22-96 years old underwent coronary artery calcium scanning; there was a strong and graded association between coronary artery calcification and coronary artery events. For example , the hazard ratio for coronary events in the form of non-fatal myocardial infarction or death from any coronary event was 8.7 among men and 6.3 among women with a CAC score of 400 or more over a median follow up time of 3.5 years[38].In the Prospective Army Coronary Calcium Project, there was an increase in the risk of coronary events by a factor of 12 in men and women 40-45 years old with coronary calcification after a median follow up of 3 years[39]. In the Multi Ethnic Study of Atherosclerosis (MESA) cohort, doubling coronary artery calcification in four major ethnicities was associated with a 25% increase in the probability of major coronary events (myocardial infarction or death from coronary events) or any other coronary event independent of other risk factors[40]. In the Rotterdam population-based study of elderly asymptomatic patients, the adjusted relative risk for coronary events were 3.1, 4.6 and 8.3 for

coronary calcium score of 101 to 400, 401 to 1000 and more than 1000 respectively compared to a score below 100[41]. In a meta-analysis of studies evaluating the risk of atherosclerotic cardiovascular disease (ASCVD) in low risk women, presence of CAC was associated with increased risk of ASCVD and modest improvement in risk stratification compared to conventional risk factors for CAD[42]. CAC was found to be associated with increased prevalence of obstructive CAD, all cause and cardiovascular mortality in Japanese patients[43].

Prediction of future events is not confined to the coronary artery calcification score at a given time only, but rather progression of calcification also was found to predict events. Budoff *et. al.* reported significant increase in risk of all-cause mortality with progression of calcification despite the method used to assess the progression[44]. These findings were validated in another study assessing the association between progression of coronary artery calcification and future coronary events in a cohort from the Multi Ethnic Study of Atherosclerosis (MESA). Rates of coronary artery disease events were found to have a graded and linear relationship with increase of coronary calcification more than 100 units as measured by Agatston score. However, among patients with no coronary calcification at baseline, any calcification on follow up was associated with significant increase in rates of coronary events[45]. In patients with diabetes, CAC predicts all-cause mortality and cardiovascular events [46]. In hypertensive patients, there was a positive association between the amount and extent of coronary calcification and blood pressure values[47] . It was found to be a significant predictive of fatal and non-fatal coronary artery events, even after long follow up time of 14 years, which indicate its independent role in developing such outcome[48].

2.1.3.2: Clinical Significance of Calcium Score of Zero

The significance of a calcium score of zero was investigated in multiple studies of asymptomatic and symptomatic participants. Raggi et. al. reported very low annual coronary event rates of 0.11% for calcium score of zero compared to 4.8% annual events rate for calcium score of 400 or more in asymptomatic patients[49]. LaMonte et al reported an event rates of 0.4 and 0.7 per 1000 person-year for men and 0.7 for women per 1000 person-year for a calcium score of 0 compared to events rate of 8.7 per 1000 person-year for men and 6.3 per 1000 person-year for a calcium score of 400 or more in a cohort of more than 10000 asymptomatic patients followed up for 3.5 years[38]. This low event rate associated with zero calcium score was consistent with results

from other studies like Prospective Army Coronary Calcium Project, St Francis Heart Study and MESA [39,40,50]

Among symptomatic patients, coronary calcification score of zero was associated with similar low event rates. In a study of 192 patients with chest pain referred for coronary calcium scanning and followed up for a total of 50 months, calcium score of zero was associated with an annual cardiovascular event (including hard events) rate of 0.6%[51]. In a prospective study enrolled patients with chest pain in the emergency room setting, negative Electron Beam CT including zero calcium score in presence of negative biomarkers and low to intermediate ECG findings has a negative predictive value of 94 to100% [52]. The study suggested that patients can be safely discharged from emergency room with no further testing needed. Two small prospective studies enrolled symptomatic patients showed no cardiac events at all in patients with zero calcium score over 2-3 years follow up[53,54].Valenti et al. reported similar mortality rates among diabetic and non-diabetics with calcium score of zero up to five years, but beyond five years ; mortality was significantly higher among diabetics[55]. Authors attribute that partially to the progression of calcification after five years. In a recent meta-analysis of studies conducted on patients with acute chest pain, zero calcium score was associated with low subsequent risk of major adverse cardiovascular events[56].

The prevalence of obstructive coronary artery disease in patients with zero calcium score was found to vary widely depending on the population investigated, technology used and the definition of the obstructive coronary artery disease. In a mixed population of acute and chronic chest pain, obstructive (Luminal narrowing of ≥ 50% on multi-detector CT scans) CAD was found in 7% of patients with zero calcium score (Rubenstien Am J Card 2007). Analysis of patients with no coronary artery calcification recruited prospectively in CORE64 study showed that obstructive coronary disease (≥50 % stenosis) can be found in up to 19%, with a low sensitivity and negative predictive value of zero calcification to rule out obstructive coronary artery disease (45% and 68 % respectively)[57]. However, analysis of participants with zero calcium score from CONFIRM (Coronary CT Angiography Evaluation of Clinical Outcomes: An International Multicenter) registry revealed a prevalence of obstructive coronary artery disease of 3.5% and a lower overall risk of developing major cardiac events [58]. Recently, Mittal et. al. reported a very low prevalence

of obstructive coronary artery disease of 1.4% and a lower event-rate compared to those with coronary calcification in a cohort mostly of asymptomatic or have atypical symptoms [59].

2.1.3.3: Coronary Calcification and Cardiovascular Risk Modification

When added to a conventional risk factors, CAC improved the performance of prediction models significantly and led to reclassifying individual risk. In Rotterdam study, more than 50% of patients in intermediated risk for cardiovascular events in 10 years per Framingham risk score were either reclassified as high or low risk when CAC was added to the risk score. The authors used an empiric cut off limit of 50 and 615 units based on Agatston score for low and high risk respectively[60]. Adding CAC to Framingham score and Adult Treatment Panel III (ATP III) score resulted in reclassification of 21.7% and 30.6% with Framingham risk of 6-20% and ATP III risk score of 10-20% respectively into high risk group which is defined to be >20%; and c statistics improved significantly by adding CAC to traditional risk models [61]. In MESA, adding CAC to Framingham risk score resulted in reclassification of 26% of the sample to either a higher or lower risk class with a net reclassification index of 0.26 [62]. Recently, Kelkar et. al. reported a net reclassification improvement of risk for women by 0.155 (P=0.002) and for men by 0.094 (P=0.03) when CAC was added to Framingham risk score in a cohort of 9715 asymptomatic patients with an average follow up time of 14.6 years [63] .

2.2: Clinical Prediction Models

2.2.1: Overview

A clinical prediction model combines a set of characteristics related to the patient, disease or treatment to estimate probability of certain outcome or to investigate which treatment is effective[64]. They are also referred to as clinical prediction tools or clinical prediction rules. In the earlier years, clinical predictions were based on the experience of senior respected clinicians and used to be called clinical aphorisms[65]. Today, these models are created using multivariable statistical methods to help clinicians in decision making process based on estimation of probabilities to avoid potential biases associated with traditional clinical decision process.

Establishing a clinical prediction rule passes through four distinct phases. These phases include derivation, validation, impact analysis and implementation. The derivation phase is a

process of identification of predictors and formulation of them to assess the probability of the outcome. In the validation phase, the rule is tested in a separate cohort. In impact analysis phase, the rule's usefulness is measured in clinical setting in term of cost-benefit, patient satisfaction and resources allocation. Finally, in the implementation phase; the widespread usefulness , acceptation and adoption of the rule by clinicians is evaluated [66].

Clinical prediction rules have multiple advantages over the conventional clinical decision-making methods. First, statistical models are likely to accommodate many more factors than the human brain can consider. Second, statistical models are likely to be reproducible on similar datasets while human brain is exposed to certain degree of inconsistency. Third, many statistical models have been found to be more accurate than clinical judgement alone when it comes to making clinical decisions [66]. Despite these advantages, the use of clinical prediction rules in every-day practice seems to be limited. It has been proposed that there is an implied lack of knowledge of how to use these rules[66].Other limitations to the use of clinical prediction rules are related to existence of more than one rule for the same clinical problem, diversity of the methods by which these rules were developed and difficulty in comparisons of these rules [66]. Methodologically, statistical modelling is encountered with some challenges, mainly the model uncertainty and sample size. Model uncertainty arises from incomplete pre-specification of predictors prior to fitting the available data[67,68]. However; statistical methods to investigate uncertainty, such as bootstrapping, are available and can be helpful in model development.

2.2.2: Fundamentals of Clinical Prediction Models' Development

In the following sections, we will present some of the fundamentals for clinical prediction models, including study design, predictor selection, outcome definition and measurement and common statistical procedures used in modelling.

2.2.2.1: Role of Study Design in Clinical Prediction Models

Data used for development of clinical prediction models are obtained from multiple types of studies, such as cohort studies, case-control studies and cross -sectional studies depending on the goal of the prediction rule. Cohort studies, either prospective or retrospective; have been the prototype study design for derivation of prediction models. Prospective studies have the

advantages of prespecified inclusion and exclusion criteria which will enable consistent definition of predictors, as well as a precise assessment and classification of outcomes. These studies usually have sufficient sample size and undergo periodic quality checks. On the other hand, retrospective studies are limited by incorrect identification of subjects, poor quality data recording and missing data despite the low cost and feasibility of conducting such studies. A prospective nested case-control studies are attractive designs for prediction of rare outcomes. If the outcome is well-defined, then the bias is minimized with random sampling of the controls used for comparisons; but retrospective assessment of predictors is the main limitation of such approach [69]. Overall, development of prediction models, like other clinical research, are better conducted using a prospective design whenever possible.

2.2.2.2: Model Predictors

Strong predictors are required for a prediction model to perform well. Strength of a predictor is a function of association between the predictor and the outcome; and distribution of the predictor [70]. The degree of association of predictor with the outcome is usually obtained from previous work or clinical relevance based on content expert's best knowledge[71]. For key predictors, prospective pre-selection is preferred. Predictors that are known to have a little or no contribution to the outcome should be excluded unless relevant to the application in practice, which is usually determined by content experts[72].

Plausible interactions between predictors should be carefully chosen and determined either a priori or during the development of the model as some interactions are found to be an important predictor and contribute to the model performance. Correlation between different predictors is common and attention should be paid to the fact that predictors with high correlation may be excluded based on the strength and relevance of the predictors due to little information added to the model. Predictors can be categorized based on the nature of the problem under investigation according to the proposed categories such as demographics, disease severity and presentation, historical elements of the disease, comorbidities, physical assessment findings and functional status, and finally subjective health status and quality of life[70] .Since clinical prediction models are meant to serve as efficient and practical tools for clinicians, cost and availability of predictors in real life is important in the process of predictor selection. Variables such as demographics, basic clinical assessment elements and laboratory tests should be considered before invasive and

advanced diagnostic tests. Advanced diagnostic tests or procedures can be added to basic clinical models to assess their incremental value, for example; adding coronary calcification score to basic clinical models to predict cardiovascular events [60]. It is important for the definition of predictors to be consistent with definitions used by the end users[70]. There is a concern when definition of predictors varies between research settings and clinical practice settings which need to be corrected for in the prediction model.

The clinical model's specification is not well defined as no clear rule for selection of predictors and factors such as the clinical content, data availability and statistical significance are usually considered. In general, there are two approaches to predictor selection. First, is to define a list of predictors a priori based on knowledge from previous work or clinical content experts [70,73]. The second approach is based on selection of the variables included in the final model based on their statistical significance either in univariable or multivariable analysis[70,73]. We will discuss these approaches in detail in the next chapter.

2.2.2.3: Outcomes in Clinical Prediction Rules

Outcomes of diagnostic or prognostic significance are the most commonly modelled outcomes in clinical prediction rules like mortality and presence, progression or recurrence of certain diseases or clinical events. Other prognostic outcomes beyond mortality and non-fatal clinical outcomes have been modelled depending on the goal of the clinical prediction model such as patient centered outcomes (e.g. functional status) and wider range burden outcomes (e.g. days absence from work) [70]. Although it seems easy to define mortality as an outcome, it sometimes becomes an issue to define cause-specific mortality rather than all-cause mortality and relative versus conditional survival should be estimated in such case. Definition of diagnostic outcomes in clinical prediction rules is based on results of a diagnostic test and need to be done according to standard reference. In case such a reference is absent, relevance of diagnostic test to clinical features or future events can be used and adjudicated by content experts [74].

Composite end points are common in cardiovascular medicine research, which combines mortality with non-fatal events; to increase the effective sample size and thus the power of statistical analysis [70]. For example, the Framingham risk model (Wilson model) predicts incident cardiovascular disease, defined as fatal or non-fatal myocardial infarction, sudden cardiac death or

angina pectoris[75]. When used as an outcome in modelling, the prevalence of the individual components of the composite outcome should be calculated and weight of each component in developing the composite outcome should be clarified.

2.2.2.4: Analytical Methods for Clinical Prediction Rules

The modeling method depends mostly on the nature of the question under investigation, prediction outcome and type of data available to derive the predictors. There are fundamental considerations when choosing a statistical method for modeling, including suitability of the method to fit the structure of the data available; robustness of the method to deal with hidden problems within the data such as skewed distribution; and the appropriateness of the mathematical formulas for modeling and extendibility of the model to be used on different datasets. The type of outcome being modeled is by far the most important determinant of the statistical method of modelling.

To model continuous outcomes, Linear regression model is used as a reference statistical method [70,76]. There are two important assumptions when modelling continuous outcomes using the linear regression model. First is that the residuals are normally distributed and secondly, they do not depend on observed variable value [70] .

For binary outcomes, logistic regression model is the most commonly used statistical method. It has the flexibility to incorporate both categorical and continuous predictors, non-linear transformations and interaction terms [70]. The binary outcome Y is related to a linear combination of predictors and regression coefficients using a logistic function to restrict the prediction to (0,1).

For categorical outcomes, dichotomization can be done and then modelling can be conducted using logistic regression model. Several reports have discussed the use of polytomous logistic regression models to model three or more outcomes simultaneously [70,77]. Ordinal outcomes are common in clinical practice .These outcomes can be modelled as continuous outcomes or simplified to binary outcomes and use logistic regression models [70].

For outcomes that are modelled during a specific follow up time, survival analysis is appropriate. In such case, follow up time may be incomplete for some subjects in the analysis and these cases are censored. Cox proportional hazard is the most often used model for survival

outcomes [78]. It is like logistic regression model conditioned to outcome occurrence time with hazard function used at baseline and indicates the risk of outcome during follow up time [79].

2.2.3: Methodological Considerations for Clinical Prediction Models

There are few statistical considerations that are important for the process of developing practical, robust and valid clinical decision tools. These considerations are important to balance model performance, complexity and ability to perform on different populations. We will discuss some of these considerations including over-fitting, handling of missing data and dealing with continuous predictors.

2.2.3.1: Over-fitting

Overfitting refers to fitting a prediction model with too many degrees of freedom or fitting a model with large number of parameters, which performs well in the cohort used for deriving the model but does not in a cohort of new subjects[70]. Overfitting leads to spurious association between predictors and the outcome and will reduce the model reliability, which is ability of the model to predict the outcome in future observations[73]. Therefore, the number of all initial parameters considered for the model, which includes the intercept and all candidate predictors (including interaction terms) needs careful attention in relation to the sample size [73]. Multiple studies have shown that a fitted regression model is likely to be reliable if the ratio of outcome events to the number of parameters ,Event Per Variable (EPV) is at least 10 events per variable [80].

Overfitting is a major issue in development of clinical prediction models and arise from model uncertainty [70]. This uncertainty results from deriving the model based on the available data rather than specification of the model a priori, while the analytical method assumes it is pre-specified and ignores this uncertainty[67,68]. Selection bias can result if the model was fit using available dataset and a set of predictors instead of choosing a representative sample from the population and fitting a prespecified model [81,82]. In addition, overfitting increases the chance of including predictors that have a large effect only; this is known as “testimation” bias because predictors are tested first [68,70]. Uncertainty may extend to the parameters in fully-prespecified models with overfitting[83]. Overfitting in clinical prediction models can lead to optimism in model performance which is the difference between the model performance and the true performance in

the underlying population[84]. Optimism in model performance can be presented by testing for trade-off between model complexity and model performance through internal or external validation.

In a systematic review of the methods and reporting of clinical prediction models published in 2008 in six large clinical journals, most of the published models did not report the number of events per variable, and more than half of these reports were found to have less than 10 events for each predictor; this increased the risk of overfitting in these models[85]. To minimize overfitting in clinical models, number of events per variable as a rule of thumb should be 10 or more. This number is calculated by dividing the number of observed events by the number of regression coefficients including the interaction terms[70,86,87].Achieving such number of events per variable is sometime a challenge especially when modelling rare outcomes.

Shrinkage, another approach to account for and minimize model overfitting, refers to shrinking regression coefficients usually to zero. This results in moving poorly calibrated predictions towards the average risk. Shrinkage could improve the accuracy of these predictions and reduce the model optimism when applied on new subjects [88]. In the simplest approach, a shrinkage factor estimate can be calculated and applied using the shrinkage estimate by J. C. van Houwelingen and La Cisse. This estimate is obtained by calculating the ratio of the difference between the likelihood ratio statistics of the model and predictors degree of freedom to the likelihood ratio statistics[83]. Shrinkage can be incorporated in the model fitting initially which results in better calibrated predictions[73]. This process is referred to as penalized regression, in which constraints are applied on the estimates of regression coefficients and results in smaller estimates. Penalized regression is considered the best approach to complex models with limited data[73]. Ridge and Lasso methods are the most popular for penalized regression (Shrinkage)[88]. In ridge regression, the sums of squared regression coefficients are restricted by a certain threshold to maximize the predictive ability of the model in a cross-validation process. The process is similar in Lasso, however instead of restricting the sums of squared regression coefficients to certain threshold, the sums of the absolute regression coefficients are restricted to a certain threshold. This allows for eliminating some of the predictors due to restriction to zero.

2.2.3.2: Mechanisms and Handling of Missing Data

Incomplete data is a common problem in epidemiological research and can affect the validity of research results by leading to inefficient analysis, biased estimates of association and

difficult interpretation of results[70,73]. Missing data issues are more frequently encountered when data on some predictors are not available, but outcome data can be missing. Knowledge of the extent and mechanism of the missing data is important to handle it appropriately in clinical prediction modelling; to reduce the risk of bias and the effect on models' validity. Most of the standard statistical software packages conduct complete case analysis and delete subjects with missing values for any potential predictor[89,90]. In available case analysis, subjects with missing values of specific predictors are deleted but not for missing values of other covariates. Both analyses ignore information about subjects who have data on some but not all predictors and lead to inefficient statistical analysis. Interpretation of clinical prediction models with missing data can be difficult due to variability between estimates in univariate and adjusted analyses which could be due to either missing values of some subjects in multivariable analysis or correlation between predictors. Comparisons of p values and of performance measures of both models when number of subjects are different are cumbersome and limit a model's interpretation. Finally, presence of systematic differences between subjects with complete data and other with missing data represents a significant source of bias in model development [70].

There are three mechanisms for missing data including Missing Completely at Random (MCAR), Missing at Random (MAR) and Missing Not at Random (MNAR) or informative missing. Understanding the underlying mechanism of missing data is important for statistical analysis which rely on assumptions regarding these mechanisms[70]. In MCAR, the reasons for missing the data are not related to any of the subjects' characteristics or responses including the missing value[73]. However, in MAR; data are not missing at random but rather the probability of missing depends on the value of another variable that was measured and subjects with missing values are randomly different from others. The informative missing (MNAR) is a mechanism where missing data for a variable is a function of data that are not available, such as tendency of extreme values to be missing. In contrast to the MCAR and MAR, MNAR is called nonignorable non-responses and considered the most difficult to handle in the analysis [73].

Characterizing the pattern of data missing as a part of exploratory data analysis is a standard approach in statistical modelling. The two most frequently used techniques to explore missing data pattern are binary logistic models and recursive partitioning for predicting the probability that a variable is missing[73]. To avoid the standard complete case analysis during modeling and the

subsequent issues of biased estimates as discussed earlier, two approaches are available statistically to deal with missing data. These two approaches include the maximum likelihood and multiple imputation, which are the most commonly used approaches in clinical prediction modeling[70].

Imputation is substitution of missing data with plausible data to allow for a complete data set analysis using standard statistical procedures[70]. Missing values can be estimated simply from the mean for continuous predictors or from the most frequent category for categorical predictors [91]. This technique does not account for the potential correlation between predictors, underestimates the association between predictor and outcome and ignores variability between predictor values among subjects [70]. To overcome these issues, conditional mean imputation and regression imputation have been used to build an individual predictive model for a predictor with missing values based on data of other predictors and the outcome in multivariable models with expected values imputed [73].

In multiple imputation, a certain number of completed data sets are created by imputing the missing values for each data set using an imputation model. Missing values are imputed by approximation to the distributional relation with the available information[70]. This method yields a specific conditional distribution for each variable with missing data using available information[92]. Since some predictors in imputation model used to impute missing values could have missing values too, data augmentation methods are used as a solution with up-front imputation step for the missing data and a second step, which draws new estimates for the model parameters [70]. This process continues until convergence and repeated to obtain the number of desired completed data sets. Number of imputed data sets varies according to uncertainty about the predicted missing values from the observed data [70].

After creating completed data sets, analysis is conducted for each data set in a standard way, then an analysis of all completed datasets is combined to obtain estimates of regression coefficients and performance. Point estimates are simply the average of point estimates of the imputed data sets and the variance estimates is the average of variance within each data set and between these sets[70].

2.2.3.3: Handling Continuous Predictors

There is a tendency in clinical research and in modelling more specifically to dichotomize continuous variables for ease of analysis and interpretation in clinical practice[93]. Others argue that dichotomization of continuous variables would result in biased assessment of the association between the predictors and the outcome due to loss of information. On the other hand, assuming a linear relationship between continuous predictors and the outcome means that the strength of association is the same for all possible values of the continuous predictor[70].

Dichotomization leads to loss of power and information due to extreme rounding. It was estimated that dichotomization of continuous variables results in a loss of one third of the data[94]. Dichotomization instead of categorization of continuous variables into multiple groups results in under-estimation of risk distribution across different groups. Furthermore, dichotomization of a continuous variable in regression models would lead to residual confounding compared to adjustment of the continuous variable[95]. Dichotomizing more than one variable in clinical modelling can lead to serious problems related to which variables are significant and the over-all predictivity of the model. These issues are due to the potential correlation between these dichotomized variables. Correlation of two dichotomized model may result in spurious association of one of these two predictors and the outcome as well as multiple interactions in the model[95].

The choice of a cut point for dichotomization is usually determined a prior based on either a clinical significance or findings from earlier similar research. However; it is common to use the median of the sample as the cut point for dichotomization[95]. Use of median of the sample as a cut point is associated with different cut points in different samples and makes it difficult to compare different studies. The search of optimal cut point for dichotomization of a continuous variable is not recommended as it is a data driven approach and a source of bias which can lead to over-estimation of the predictor-outcome association[95].

In general, dichotomization of continuous variables is not recommended given these concerns. Instead, researchers should follow methodologically sound approaches for treatment of continuous variables in modelling.

To avoid inaccurate assumptions of a linear relationship between a continuous predictor and an outcome, and dichotomization of continuous predictors in clinical models; alternative

approaches have been proposed to model continuous predictors using some non-linear functions. These approaches include polynomials, fractional polynomials and splines functions[73].

2.2.4: Measurement of Model Performance

A model can be judged by its overall performance, discriminative ability and calibration. The overall performance of a model reflects the difference between the predicted outcome and the observed one. Explained variation between observation and prediction is measured in prediction models by R^2 score, in particular, linear models[70]. Brier score is another overall-performance measure that is used for binary outcome and measures mainly the difference between predictions and observations. However; it is dependent on the incidence of the outcome [73].

The ability of prediction models to discriminate those with outcomes from those without is an important aspect of model performance. There are multiple measures for discriminatory ability of models but concordance (*c*) statistics is the most commonly used for generalized linear models[70]. In models with binary outcomes, receiver operating characteristics (ROC) curve is formed by plotting of the sensitivity against 1-specificity for certain number of predicted probabilities. Area under the curve (AUC) is similar to c statistics and can be used for measurement of discrimination in binary outcomes[70].

2.2.5: Validation of Developed Models

The purpose of developing clinical prediction models is to be able to predict valid outcomes in new patients for making proper clinical decisions. To assess this, it is crucial to test the performance of the newly developed model with a different data set. Validation is a process of assessment of the validity of the new model predictions in new patients and can be internal or external.

2.2.5.1: External Validation

Validation of the newly developed model on external data set is important to support the generalizability of the model. External validation considers the differences between the original data set and validation dataset in terms of historic, geographic, methodological and spectrum aspects[96]. This method allows for testing of the transportability across these aspects. Historic transportability assesses the performance of the model in a more recent population with a recent diagnostic methods or therapeutic approaches. Geographic transportability investigates the new

model performance in a sample from a different geographic area with all pertinent place-related features. Methodological transportability refers to the performance of the model in a data that was collected and processed using different methods. Spectrum transportability tests the model performance in different spectrums of the disease (i.e. testing models that was derived from emergency room patients on data collected from ambulatory care or validating a model that was derived from data of early stage disease on data from advanced stage of the same disease) [96].Fully independent validation refers to external validation by independent investigators. Full external validation has been found to under-estimate the model results in several validation studies[97].

Results of validation of well-developed models are often favorable, however; poor validation can be explained by several factors. Inadequate development of the new model due to small sample size or single center data is a major factor, but more importantly, the unfavorable validation could be a result of true underlying differences between the derivation and validation settings. Fundamental differences usually arise between the cohorts of derivation and validation if definition of predictors and outcome differed significantly. Historic transportability of the developed models is a challenge as definitions of outcomes and predictors are dynamic, especially those based on diagnostic tests and not representing hard clinical outcomes or well-defined clinical conditions [97].

2.2.5.2: Internal Validation

Internal validation emphasizes on the internal validity (reproducibility) of the developed model. The same dataset of derivation is used for validation through data splitting, cross validation and resampling or bootstrapping techniques.

Splitting the original data set of derivation is a straightforward technique that mimic the idea of external validation[73]. The original data set is usually divided at random into a derivation and validation subsets. Although it is a classical approach, there are issues related to the reliability of the development and validation of clinical models using this technique. Complete random splitting may lead to variation in the outcome occurrence rate and the distribution of the predictors between the two sub-sets. Moreover, with split-sample method for validation, a significant amount of data is excluded while available for derivation. Development of a model from a smaller sub-group of the data leads to a less stable results compared to deriving the model from the complete

data set, as well as validation in a smaller data sub-set may yield less reliable model performance[70]. Use of data splitting is not recommended due to these drawbacks and availability of computer-based alternative solutions [70].

Cross validation is a modification of the split-sample method in which the complete data set is divided into one large group for derivation and a smaller group for validation. This process is repeated multiple times to have all observations used for both derivation and validation. Overall validation results is calculated as the average of the individual validations[98]. Although the larger group is usually used for derivation, the need for large number of repeats to get stable results is a challenge and less efficient.

Bootstrapping is a general process by which resampling and replacement from the underlying population is conducted multiple times, often 100-200 times (Samples). Larger number of repeats may be required to reach stable estimates which is facilitated by the current advancement in computers [70]. Each time, the sample size is of the same as the original sample size [99]. For validation purposes, with each bootstrap resampling, a prediction model is developed and evaluated in both the bootstrap sample and the original sample [70]. In simple bootstrap validation, the estimate of the likely performance of the final model on new subjects is estimated from the averages of the indexes of performance measures repeatedly of the original sample[73]. In the enhanced bootstrap validation, optimism is calculated as the difference in model performance between the bootstrap sample and the original ones. Optimism is subtracted from the model performance in the original sample. Bootstrap validation aims at validation of the process of deriving the original model through measuring the optimism. Bootstrap-based validation yields stable estimates of the performance due to correction of the optimism and use of the same sample size for both derivation and testing of the model[70]. It accounts for all sources of model uncertainty especially variable selection based on simulations[100]. It has been found to work reasonably well in studies with large number of predictors like in genetic research studies[70]. Bootstrap-based validation and other resampling techniques are used if an automated selection of predictors, such as step wise technique, used only. This indicates that full models or models with an automated stepwise variable selection can be validated using these techniques[70]. It is also difficult to repeat some of intermediate steps done during modelling if bootstrap-based validation is desired. For

example, collapsing of categories of predictors, assessment of linearity by visual plots, testing for interaction and truncation and omission of outliers and influential observations [70].

2.2.6: Model Presentation

Clinical prediction models provide estimates of the absolute risk of the predicted outcome based on set of predictors usually patient characteristics. For an individual patient, this will result in a spectrum of predictions that usually require the intercept for interpretation[70]. For clinicians, it is practical to present the model in a way that facilitate the decision making based on the model predictions. There are several formats for presentation of clinical prediction models.

The simplest format for presentation of clinical prediction models is the regression formula of the model [70]. Estimating the individual predictions is usually conducted in two steps. First, predictors are calculated by multiplying the value of continuous predictor by the its regression coefficient and the code of categorical variable by its regression coefficient. Second, the linear predictor is translated to units on the outcome scale[70]. It is important to have a clear definitions and units of the different predictors in the model. Regression formulas can be implemented electronically in health facility systems, web-pages or hand-held devices applications.

Nomograms are graphical presentation of clinical prediction models that require similar calculation of the same steps as for regression formulas[70]. It is essential to calculate the linear predictors accurately. In the chart of monograms, there is a reference line that has a range of score points from zero to a certain value that represent a maximum value for the score for any predictor. Each predictor then has a line with its score based on the model. The last line represents the total points score with corresponding predicted probability. The user add score of each predictor manually and the sum is the total score for the individual patient. Nomograms were used widely in pre-computerized solutions era[70].

Point score chart is another simple format for presentation of clinical prediction models. Regression coefficient is used as a scale for the new score after rounding. The smallest regression coefficient is used to serve as the common denominator and other variables are assigned points by dividing their regression coefficients by that denominator[101]. In another approach proposed by Sullivan et. al., a reference value is defined for each predictor and the difference between the reference value and the value of each category of the predictor is calculated. This difference defines

how far is a certain category of the predictor from the reference value of that predictor. Then, the regression coefficient is multiplied by the calculated difference between the reference value and the category value to obtain how far is each category from the base category in regression units. This will assign a value of zero to the reference value. After that, a common denominator is used to assign a scoring points to different predictors[102]. For regression coefficients of continuous predictors, they can be kept as a regression coefficient for a continuous predictor which should be understood by the user and used properly. Alternatively, continuous predictors can be dichotomized or categorized, and score can be calculated for each category by multiplying the mid- point of that category by the regression coefficient[102]. Categorization of continuous predictors is usually done for simplification of the score use, however; it can lead to loss of information and increase in the uncertainty of the developed model [70].

2.3: Clinical Prediction Models for Coronary Artery Disease

Over the last few years, there had been a growing interest in the development of clinical prediction models in general and in the field of cardiovascular health. In a recent review by Wessler et al., around 800 models of cardiovascular disease prediction were published over 20 years with more than 200 models in the field of coronary artery disease[103]. This growth in development and publication of these models is important for better understanding of different diagnostic and therapeutic outcomes, however; it highlights several issues related to the methods of development, calibration, validation and application of these models. For example, in the report by Wessler et. al.; only 63% of the published models reported a quantification of discrimination like c-statistics, and even less (around 36%) reported a measure of calibration [103]. Validation of published models either internally or externally and comparison to other models was limited in the published models which represents a major challenge in the dissemination and application of the developed models [103,104].

More recently, Damen et.al. conducted a systematic review that focused on clinical prediction models of cardiovascular disease risk, including data related to the outcome predicted, method of derivation, model performance assessment and calibration and validation [105]. Authors reported that there is an abundancy of models that predict coronary artery disease but 64% of these models were not externally validated and only 19% were validated by independent investigators[105]. Moreover, there is a significant variation in the definition of the predicted outcomes, with most

being fatal and/or non-fatal cardiovascular disease, as well as the population from which the model was derived. This variation leads to difficulty in implementation of such models in every-day practice.

2.3.1: Overview of Models Predicting CAD

Diamond and Forrester published one of the first tools to predict the probability of coronary artery disease in patients with chest pain[106]. Since then, several tools have been developed specifically to predict the probability of coronary artery disease in symptomatic and asymptomatic patients. This includes studies by Pryor et al. who published Duke Clinical Scores of significant and severe coronary artery disease[107,108]. Morise et.al. published a model to estimate probability of CAD in men and women with suspected CAD[109]. The model by Diamond-Forrester was extended and updated by Genders et. al.[110]. Two models to rule out CAD were published in 2010 by Gencer et.al. and Bosner et. al.[111 112]. Genders et. al. published a model that incorporated calcium score as a predictor [113]. Yang et. al. published a model from our institute to predict high risk CAD based on coronary CTA findings[114].

A brief description of these tools is provided in Table 1. Other models predicting different outcomes related to clinical presentation of chest pain especially in emergency room like safety of discharge or risk of adverse outcomes have been developed, however; they are not the scope of this review. In the following sections, we will review the basic aspects of the development of these tools. Table 2 display the methodological aspects of the published models.

2.3.1.1: Derivation Cohorts, Setting and Design

There is a significant variation among the published clinical prediction models of coronary artery disease in term of setting and cohorts used for derivation of these models. Diamond-Forrester reviewed the literature to estimate diagnostic accuracy of historical elements and different diagnostic tests, and developed their approach using Bayes theorem of conditional probability[106]. Duke Clinical Scores were derived retrospectively from a database of patients referred for coronary angiogram and then validated using data in outpatient setting[107,108,115].The model by Morise et. al. was developed retrospectively from a cohort of patients referred for exercise stress test as a work up of a presentation with chest pain[109]. Two models were developed

retrospectively from data on patients in general practice[111,112]. Two studies developed prediction tools from data collected prospectively from patients presented with chest pain and were sent for conventional coronary angiogram or coronary CTA(T. S. S. Genders et al., 2011; Tessa S S Genders et al., 2012). Yang et. al. developed a model to predict high risk CAD retrospectively in all patients referred to have coronary CTA. We excluded models developed from patients with acute chest pain in emergency room as this population is not within the scope of this review.

This variation in setting and design is important due to differences in disease prevalence between different clinical settings, as patients encountered in primary care practice are likely to have lower prevalence of the disease compared to those who have been referred to assessment by cardiology or specialists in general. Referral bias is likely to occur in models derived from databases of patients referred for coronary angiography compared to patients from general practice. Moreover, practitioners should consider the type of cohort and clinical setting for appropriate use of these models in daily practice.

2.3.1.2: Predictors Selection and Model Specification

The selection process of predictors to be included initially for analysis in these models is not justified in the publications, however, it seems to be largely driven by available data and clinical input supported by the evidence regarding the importance of different predictors. For example, there is a consistency across the available models by including certain predictors such as symptoms, mainly typical chest pain, and demographics, such as age and sex.

Since the description of angina in 1786 by William Heberden[116], clinicians considered chest pain as the sole diagnostic clue of underlying coronary artery disease, however; diagnostic utility of chest pain alone is limited[117]. Classification of chest pain according to typicality is a key step in determining the pre-test probability of coronary artery disease [3,106]. In a review of the bedside diagnosis of coronary artery disease, Chun et. al. reported that typical chest pain is the most important (LR=5.8, 95% CI 4.2-7.8) among all bed-side predictors of coronary artery disease in patient with chronic stable chest pain.

There is a variation among the published models in term of predictors included. However; most of them included all or some of the risk factors for coronary artery disease (diabetes, hypertension, dyslipidemia, smoking and family history), prior history of coronary artery disease

or other features related to location, duration and nature of chest pain (Table1). Risk factors for coronary artery disease such as diabetes, hypertension, dyslipidemia and smoking were associated with a modest improvement in model performance when added to Diamond-Forrestor, which originally included age, sex and chest pain typicality (Genders et. al. BMJ 2012).

While univariable analysis to estimate the association between individual predictors and outcome was done in five studies to determine on initial predictors for inclusion in multivariable models[107,108,111,112,114], two studies included all initial predictors simultaneously in multivariable modeling based on being risk factors for CAD from available data [109,113].

Stepwise selection was used for multivariable modelling in two studies (by Pryor et al. for prediction of severe coronary artery disease and Bosner et. al.)[108,111]. In the original model by Diamond-Forrestor, selection of predictors was based on pooled accuracy data of historical conditions using the Bayes theorem of conditional probability[106]. In the updated version of this model by Genders et. al., it was extended to include older age without adding other predictors[110].

Most of predictors in prediction models for coronary artery disease are dichotomous as they represent pre-existent risk factors for the outcome. However, continuous predictors, in particular, age was dichotomized in four studies[106–108,112].

Table 1: Brief description of published tools to predict CAD

Study	Year	Country	Setting Design	Outcome	Predictors	Model Performance
Diamond-Forrester	1979	USA	Literature review using Bayes theorem to define the diagnostic accuracy of historical conditions	Likelihood of coronary artery disease in symptomatic patients	Age Type of chest pain (Typical, Atypical or Non-cardiac)	NA
Pryor et. al.	1983	USA	Cardiac catherization referral (all patients) Retrospective	Probability of significant coronary artery disease (≥75% stenosis of any major epicardial coronary artery) on coronary angiogram	Age Sex Type of chest pain History of MI Hyperlipidemia Smoking Diabetes ST-T wave changes	NA
Pryor et. al.	1991	USA	Cardiac catherization referral (all patients) Retrospective	Likelihood of severe coronary artery disease (defined as ≥75 stenosis of the left main or all three major epicardial coronary arteries) on coronary angiogram.	Age Sex Type of chest pain Frequency of chest pain History of MI Hyperlipidemia Smoking Diabetes Blood pressure Carotid bruit	NA

Study	Year	Country	Setting Design	Outcome	Predictors	Model Performance
Morise et. al.	1997	USA	All symptomatic patients referred for exercise stress test for cardiac work up Retrospective	≥50% stenosis of at least one coronary artery on coronary angiogram. Severe CAD: ≥70% stenosis of at least 2 coronary arteries) on coronary angiogram	1.Age 2.Gender 3.Chest pain type 4.Estrogen status 5.Diabetes 6.Hypertension 7.Hyperlipidemia 8.Smoking 9.Family history 10.Obesity	NA
Gencer et. al.	2010	Germany & Switzerland	Patients were enrolled prospectively if they have chest pain in a general practice setting	To rule coronary heart disease in patients with chest pain in primary care setting at one year after initial assessment using records and self-reporting	1.Age-Sex category 2.Known risk factor (None, 1-2, >=3) 3.Known history of cardiovascular disease 4.Duration of chest pain 5.Sub-sternal area of chest pain 6.Precipitation with exertion 7.Absence of tenderness	AUC:0.94 R^2: 0.536

Study	Year	Country	Setting and Design	Outcome	Predictors	Model Performance
Bosner et. al.	2010	Germany	Patients were enrolled prospectively if they have chest pain in a general practice setting	Prediction of coronary artery disease presence during the index assessment as decided by a panel based on the collective results of work up	- Age/Sex - Known prior vascular disease - Pain worse during exercise - Pain not reproducible by palpation - Patient assumes cardiac pain	AUC: 0.87
Genders et.al.	2011	Europe &USA	Multicenter patients referred for work up of chest pain using non-invasive and invasive studies	Updating model by Diamond-Forrester	- Age - Male sex - Chest pain Typicality (Typical, atypical, non-cardiac)	AUC: 0.82
Genders et. al.	2012	USA & Europe	Multicenter patients referred for work up of chest pain using non-invasive and invasive studies	Obstructive CAD defined as ≥50% stenosis of at least one vessel on coronary angiogram or coronary CTA	1.Age 2.Male sex 3.Chest pain (Typical with and without DM, atypical, Non-specific) 4.Diabetes 5.Hypertension 6.Dyslipidemia 7.Smoking 8.Coronary calcium score	AUC: 0.88

Study	Year	Country	Setting and Design	Outcome	Predictors	Model Performance
Yang et. al.	2015	Canada, USA, Europe and South Korea	International multicenter cardiac CT registry (CONFIRM) registry	Prediction of probability of high risk coronary artery disease defined as	1.Age 2.Sex 3.Symptoms (Typical angina, atypical/non anginal pain, asymptomatic) 4.Diabetes 5.Hyperlipidemia 6.Current smoking 7.Peripheral arterial disease 8.Hypertension 9.Family history of CAD	AUC: 0.76

Overall, available models to predict coronary artery disease relied heavily on the available data mostly on risk factors and statistical significance more than a pre-specified list of predictors given the well-known association between these risk factors and the outcome predicted.

2.3.1.3: Definition of Predicted Outcome

Definition of coronary artery disease in the published models vary across these models depending on the setting and source of the data. As we discussed earlier in section, coronary angiogram is considered the gold standard for diagnosis of coronary artery disease. While Clinical Duke Score used more than 75% luminal stenosis of at least one major epicardial coronary artery on coronary angiogram as a definition for the predicted outcome of significant coronary artery disease[107], studies by Genders et al. defined the predicted outcome as at least one coronary artery luminal stenosis of 50% or more on coronary angiogram[110,113].

Table 2: Methodological aspects of the published tools to predict CAD

	Diamond-Forrester 1979	Pryor et. al. 1983	Pryor et. al. 1991	Morise et.al. 1997	Bosner et. al. 2010	Gencer et. al. 2010	Genders et. al. 2011	Genders et. al. 2012	Yang et. al. 2015
Outcome clearly defined	✓	✓	✓	✓	✓	✓	✓	✓	✓
Predictors clearly defined	✓	✓		✓	✓	✓	✓	✓	✓
Study participants clearly described		✓	✓	✓	✓	✓	✓	✓	✓
Statistical analyses described	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sample Size									
Adequate information to assess Events per Variable	✓	✓	✓	✓					
Overfitting mentioned or discussed					✓				
Missing Data									
Compete case analysis					✓	✓	✓		✓
Imputation								✓	
Not Reported	✓	✓	✓	✓					
Modelling Continuous Predictors									
Categorization						✓			
Model Specification									
Variables Pre-specified				✓			✓		
Univariate Screening		✓	✓		✓	✓			✓
Stepwise Selection			✓		✓				
Model Performance									
AUC (ROC)					✓	✓	✓	✓	✓
Calibration						✓	✓		
Sensitivity &Specificity		✓	✓	✓		✓	✓	✓	✓
Validation									
Internal						✓			
External		✓	✓	✓	✓	✓	✓	✓	✓

In the most recently published model by Genders et. al., coronary CTA was used in some participants as a non-invasive assessment for chest pain and subsequently a coronary angiogram if needed based on the results of coronary CTA. Two studies developed models to predict severe and high-risk coronary artery disease. In the study by Pryor et. al., outcome of severe coronary artery disease was defined as 75% or more stenosis of left main coronary artery or all three major epicardial coronary arteries on coronary angiogram[108]. Yang et. al. defined high risk coronary artery disease as presence of coronary CTA findings of at least 50% luminal stenosis of the left main, or 50% or more stenosis of all three major epicardial coronary arteries or two major coronary arteries including proximal left anterior descending artery[114]. Outcomes predicted by models developed in general practice setting were defined differently. In the study by Gencer et. al., outcome was determined in a delayed assessment up to one year after the initial assessment and categorized it into diagnostic groups including coronary heart disease without details on the definition of coronary heart disease and whether anatomic assessment of coronary arteries are required or not[112]. Similarly, the study by Bosner et. al. relied on a panel consists of a cardiologist, a primary care physician and a research staff member to decide on the presence of coronary artery disease at the consultation index in a delayed-type reference standard[111]. In the absence of definition of outcome that is based on anatomic characterization of coronary artery disease by coronary angiogram using catherization or computed tomography imaging, risk of outcome misclassification is high.

2.3.1.4: Over-fitting

The number of events per variable (EPV) is an important measure of model over-fitting as discussed previously (Section 2.2.3.1). Although this was not discussed explicitly in the published models to predict coronary artery disease, estimation of EPV was difficult in all. Except for the model by Diamond-Forrestor, we think that all 8 models achieved the minimum EPV of 10. This is likely due to relatively higher prevalence of the disease being modelled compared to other clinical conditions and availability of advanced diagnostic modalities. Models derived from cohorts with higher prevalence of the disease, such as databases of coronary angiogram or CTA, have a larger sample size and likely have high EPV compared to models derived from general practice. In the study by Genser et. al. and Bosner et. al., the risk of overfitting because of borderline EPV cannot be eliminated.

2.3.1.5: Missing data

Depending on the magnitude and the method of handling them, missing data can be a source of bias in modelling as discussed previously (see section 2.2.3.2). In three studies for prediction of coronary artery disease, proportion and handling of missing data were not reported [107–109]. We assume that these models were derived based on complete case analysis. In four studies by Genders et. al., Bosner et. al., Gencer et. al. and by Yang et. al.; patients with missing values on the predictors were excluded as reported by the authors [110–112,114]. Multiple imputation technique of around 2% of the data was conducted in the study by Genders et. al. [113]. In this study, outcome of obstructive coronary artery disease on coronary angiogram was imputed from CTA data as an auxiliary variable in 34% and 6% of patients with and without obstructive coronary artery disease on CTA respectively. While multiple imputations revealed obstructive disease on coronary angiogram in 65 to 77% of patients with obstructive disease on CTA, non-obstructive disease on coronary angiogram was found in 97 to 98.4% across imputations of missing data of patients with non-obstructive disease on CTA. These findings correspond well with the high negative predictive value and reduced positive predictive value of CTA in diagnosis of coronary artery disease.

Overall, stating the proportion of missing data and the method of handling this issue in the published models is suboptimal in three studies (study by Morise et.al. and the two studies by Pryor et.al.) and complete case analysis is assumed in these three studies. The other four studies reported the missing proportion; however, it was small in three of these studies. Multiple imputations were used in one study by Genders et.al. to account for the large proportion of missing data on coronary angiogram in the derivation cohort (study by Genders et.al. 2012).

2.3.1.6: Model performance

Reporting of model performance measurements varied in four studies of the published models predicting coronary artery disease [110–114]. Discriminatory ability of these models by c-statistics was the most commonly reported measure. In one study by Gencer et. al., authors reported area under the curve (AUC) for discrimination and R^2 for overall performance of the model. They mentioned that assessment of goodness of fit was assessed using Hosmer-Lemeshow statistic, but it was not reported in detail. We observed a significant time-related variation in reporting measurements of model performance. Models' performance measurements are likely to be

discussed and reported in the more recent publications [110–114]. This could be explained by the advancements in statistical methods for modelling and subsequently presence of standardized approach to development and publication of prediction models like the TRIPOD statement[118].

2.3.1.7: Validation

Except for Diamond-Forrestor model, all published models for prediction of coronary artery disease in our review underwent external validation and results were published at the same time. Datasets used for validation were defined from the same institute using referral date to separate derivation and validation cohorts in studies by Pryor et. al. and Morise et. al.[107–109]. Four studies validated their models on a completely different external data sets[110–112,114]. In one study by Genders et. al., a cross validation was used by omitting a data-subset with low prevalence of the disease from derivation data and using it for validation[113]. In one of these studies, study by Gencer et. al. , both internal validation using bootstrapping, and external validation were done [112].

Model performance in the validation cohorts was reported sub-optimally with a general description of whether the model was validated or not in the three studies, namely the two studies by Pryor et.al and the study by Morise et. al. [107–109]. Three studies reported performance of models in the validation cohorts without emphasis on that performance compared to the performance of the model in the derivation cohort[110,113,114]. In two studies, Bosner et. al. reported an improvement in c-statistics from 0.87 to 0.9 in the validation cohort; and Gencer et. al. reported c-statistics of 0.752 in external validation cohort compared to 0.94 in the derivation and internal validation cohorts[111,112].

2.3.1.8: Summary of Published Models

There are nine published tools for prediction of coronary artery disease. Diamond-Forrester approach was the known first to be used widely in clinical practice and it was developed in non-traditional way as discussed earlier. Other available models were developed more classically with fundamental differences in the derivation cohorts, model specification, outcome definition and methods of dealing with the statistical issues surrounding these model's development (see Table 2). We present a summary of some methodological characteristics of these models in Table 3.

After reviewing the published models to predict presence of coronary artery disease, three major conclusions are identified. First, the clinical setting and derivation cohort played a major role in the variation in the definition of the outcome and how the diagnosis of coronary artery disease was made. Almost all these models were derived from symptomatic patients but models that were derived from general practice differed remarkably from those derived from specialized setting, such as referral to cardiac diagnostic test including exercise stress test, coronary angiogram or coronary CTA. Secondly, older models are likely to be less transparent in term of dealing with statistical issues of model development regarding dealing with missing data and model performance. Finally, one model only utilized calcium score by adding it as a predictor in a model to predict presence of obstructive CAD in symptomatic patients, but none were developed from patients with zero calcium score, in which the prevalence of obstructive CAD is very low as discussed previously (see section 2.1.2.4).

2.3.2: Clinical Utility of Calcium Score of Zero in Prediction of CAD

CAC has been known for long time as a marker of underlying coronary artery disease and associated with increased risk of cardiac events including mortality as discussed previously in section (2.1.4.1). Multiple reports have shown the incremental value of coronary artery calcification in risk modification when added to the conventional cardiovascular risk prediction tools [60–63]. Coronary calcium score is recommended by the international professional societies to be used in addition to the conventional cardiovascular risk scores to improve the classification of individual risks given its incremental value and feasibility of routine use[119,120].

A calcium score of zero is associated with a low probability of developing cardiovascular events, however , absence of coronary calcification does not rule out presence of obstructive CAD especially in symptomatic patients [58]. Currently, only one of the published clinical prediction models of CAD incorporated coronary artery calcification data, which improved the discriminatory performance of the model from 0.79 to 0.88 when included as a predictor[113]. Given the diagnostic uncertainty of the calcium score of zero and the wide availability and easy accessibility of such diagnostic technique, it is important to have a clinical tool that can be added to the calcium score of zero to improve its clinical utility in daily practice.

Table 3: Summary of the methodological characteristics of the published tools

Characteristic	Number of models (%)
Outcome clearly defined	9 (100)
Predictors clearly defined	9 (100)
Study participants clearly described	8 (89)
Statistical analyses described	100(100)
Sample Size	
Adequate information to assess Events per Variable	8(89)
Events per Variable	8(89)
Overfitting mentioned or discussed	1(11)
Missing Data	
Compete case analysis	4 (44)
Imputation	1 (11)
Not Reported	4 (44)
Model Specification	
Variables Pre-specified	2 (22)
Univariate Screening	5 (55)
Stepwise Selection	2 (22)
Model Performance	
AUC (ROC)	5(55)
Calibration	2 (22)
Sensitivity &Specificity	7 (63)
Validation	
Internal	1 (11)
External	8(88)

CHAPTER3: METHODS

3.1: Overview of Standard Methodology for Developing Clinical Prediction Models

Despite the surge in development and publication of clinical prediction models over the last few years, the clinical adoption and acceptance by the end-users are still limited[85]. Several methodological issues and quality of reporting of most of these models play a role in limiting their clinical usefulness.

To improve the methods of development, reporting and dissemination of clinical prediction models, there is a significant demand for a consensus on methodological framework or evidence-based guidelines for development of clinical prediction models.

Currently there are multiple published individual expert opinions and reviews that attempt to provide an overview of the appropriate methodology for development of clinical prediction models. However, two published documents that can serve as a guide for such purpose, namely : Prognosis Research Strategy (PROGRESS) 3: Prognostic Model Research; and Transport Reporting of multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): The Tripod Statement [121,122].

In PROGRESS 3, authors outline three distinct phases for prognostic modelling research which can be extrapolated to other predictive models (e.g. diagnostic predictive models). These phases include model derivation, external validation of the developed model and assessment of impact of the model on clinical practice[121]. The reliable model, as described by PROGRESS 3 document, is the model that is derived from a large and high-quality dataset, conducted based on a study protocol with a sound statistical procedure and validated on an external dataset from a different location. After identification of the model, the document recommended assessment of the impact of the developed model on clinical practice if disseminated and adopted. Since a clinical model's performance may wane over time due to advances in diagnostic procedures and survival change, it is recommended to consider updating the existing models by adding new predictors as needed instead of developing new models[121].

The TRIPOD statement aims at improving transparency of reporting of studies developing, validating or updating clinical prediction models regardless of the type of the study or the nature

of the modeled outcome [122]. The statement includes a checklist of 22 items that covers areas required to be reported appropriately in clinical prediction model reports (Appendix1). These areas are grouped into 5 traditional groups commonly used for reporting health research, title and abstract, introduction, methods, results and discussion.

In the methods section, the statement emphasizes on the need for details on the study design, dates of recruitment and end of follow up and the setting of the study participants (e.g. primary care vs emergency room etc.) for both derivation and validation. Definition of outcome predicted and all predictors, including how and when the assessment of the outcome and measurement of the predictors were conducted should be clearly reported. Researchers should explain how the sample size was calculated and the method of handling missing data. Statistically, the statement requires a detailed description of all statistical procedures used in the analysis, how predictors were handled including selection of final list of predictors in the model, procedures used for assessment of model performance and any internal validation technique used. For validation on external dataset, the statement emphasizes on description of any differences between validation and derivation datasets.

Results of clinical prediction models should be reported in informative details including number and clinical characteristics of participants with and without the outcome, the full prediction model with regression coefficients, intercepts and confidence intervals. The report should include explanation of how to use the model and model performance with confidence intervals[122].These two documents, PROGRESS3 and TRIPOD Statement, represent a framework for researchers during development, validation and reporting of clinical prediction models.

In the following sections, we present the methodology we followed to develop, validate and report our model.

3.2: Overview of Methods Used for This Study

To develop a clinical tool to predict the presence of obstructive CAD as stated in the objective of the thesis (Section 1.2), we prepared our datasets and performed the explanatory analysis and summary statistics then developed two distinct models. The first model, we called The Full Model, which was developed through pre-specification of model predictors based on input from clinical experts. The second model was called The Reduced Model, and it was

developed using statistical significance for predictors selection in univariate analysis and stepwise selection of final model predictors in multivariable model.

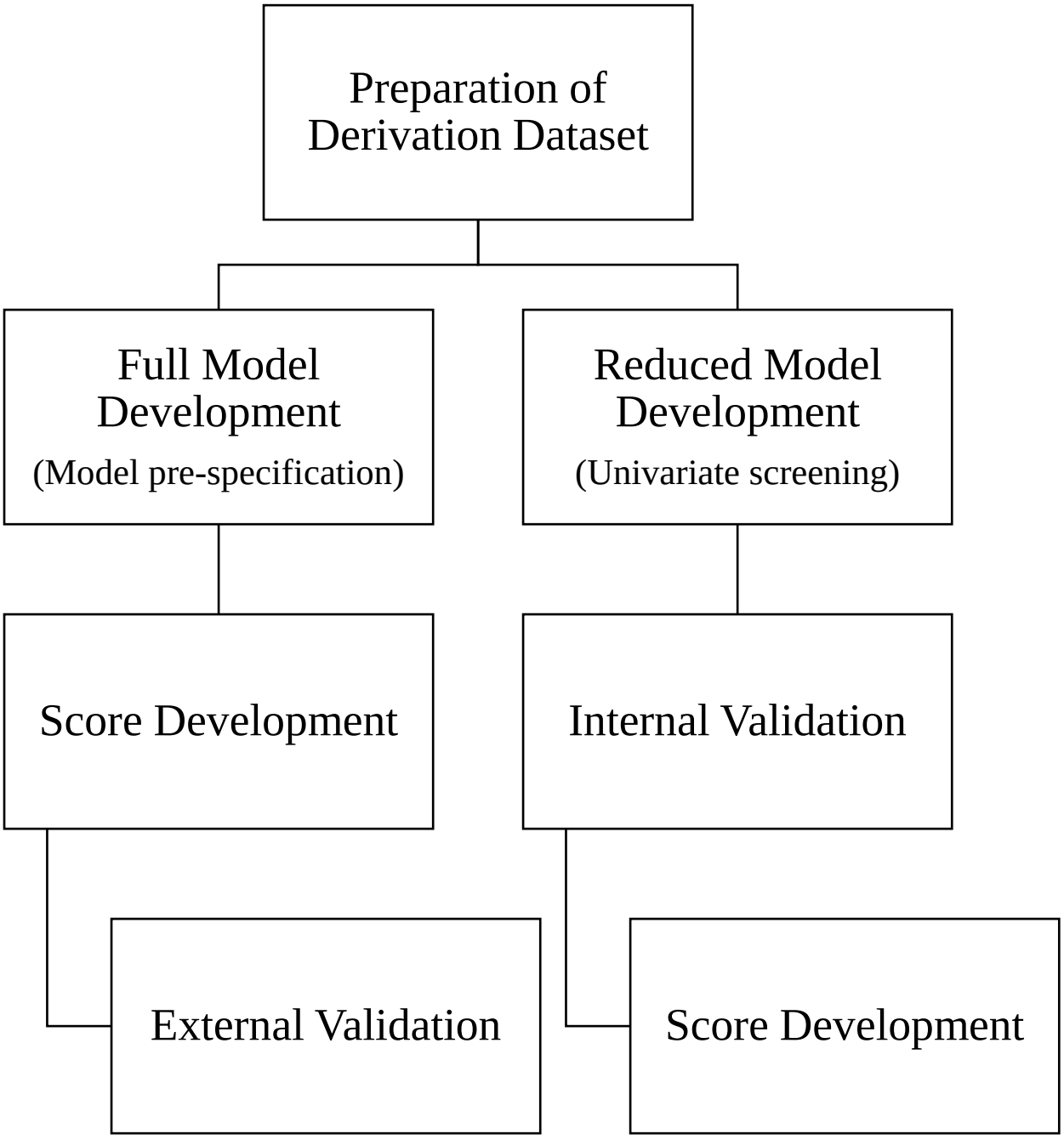
All data preparation and analysis were conducted using the statistical software SAS 9.4© (SAS institute, Cary, NC, USA.). A schematic representation of our methodological steps is shown in Figure 1. The study protocol was approved by the Ottawa Health Science Research Ethics Board (Appendix 2).

3.3: Data Source for Model Derivation

We obtained the dataset for derivation of the model from the cardiac computed tomography (CT) registry at Ottawa University Heart Institute (Chow JACC 2010). Cardiac CT registry at Ottawa University Heart Institute is a prospective registry that started in 2004 and includes all patients referred for coronary artery calcium score and coronary CTA. It covers a referral area that has a population of over a million. All referrals for cardiac CT in this catchment area are sent to this institute by family physicians and specialist.

All patients referred for the test are interviewed by a trained nurse prior to the test to collect information regarding demographic data, past medical history and risk factors, symptoms, family history, social history and habits, medications and prior coronary artery investigations. This interview is conducted as per a standardized registry form (Appendix3). After that, baseline blood pressure, heart rate, body weight and height are recorded, and ECG is performed. Subsequently, the patient proceeds to the test if there is no contraindications and the form is submitted electronically after adding the technical data of the test. Images are interpreted after acquisition by a trained physician and a separate form of test report including, but not limited to, coronary calcification and findings of coronary arteries either normal or the degrees of stenosis is recorded and submitted electronically to the central database (Appendix 4). The final report form must be approved by the reading attending physician. Quality of the submitted forms are audited periodically by a database specialist.

Figure 1:Schematic presentation of the methodological flow for development of our two models



3.3.1: Definition of Variables

All variables were recorded based on patient reporting, reviewing of medical charts if available or measurement at the time of interview. For past medical history, risk factors for coronary artery disease or findings of previous coronary arteries investigations, information was obtained from the medical electronic records or charts if available or as reported by the patient. Information on medications was mostly obtained from patients' personal documentation of their medications. Patients were questioned about their symptomatic status despite the reason for the referral. Questions regarding symptomatic status includes chest pain presence, duration, frequency, and triggering and relieving factors. All variables recorded based on patient reporting were recorded as an answer to questions by "yes", "no" or "unknown". Baseline heart rate, blood pressure, weight and height were measured during the interview as a part of the assessment in the registry. Definition of the main variables collected in the registry are presented in Table 4.

3.3.2: Image Acquisition and Interpretation

After the clinical interview, patients receive oral medication to slow their heart rate according to an algorithm designed a priori as a part of the procedure protocol. Slow and regular heart rate is important for optimal image quality in coronary CTA. After that, patients are put on the gantry table with intravenous access for the contrast. Imaging is performed thereafter in three phases. The first phase is the scout topogram to allow for accurate prescription of scanning field. Following that, a non-contrast with low power of scanning is done to minimize radiation for calcium score images. Coronary CTA is then performed if no significant calcification that may preclude accurate assessment of coronary arteries due to certain types of artifacts. The decision to proceed to coronary CTA after coronary calcium scan is made by the interpreting physician usually based on visual assessment. Coronary CTA is then performed with ECG-gated image acquisition either prospectively or retrospectively depending on the protocol decided prior to the exam time.

In the retrospective coronary CTA, the scanner is continuously imaging the heart during the whole cardiac cycle in a helical pattern. Phases are then selected and reconstructed and sent for interpretation. In the prospective imaging, the scanner is triggered by ECG to acquire images only during a certain phase of the cardiac cycle. This is in turn leads to significant reduction in radiation dose. Most of the coronary CTA protocols are prospective given the low radiation dose

with prospective imaging. The acquired images are then converted to digital images using the appropriate software. Interpretation of the scan is done by a staff physician either a specialized cardiologist or radiobiologist as per the Society of Cardiovascular Computed Tomography (SCCT)[123].

3.3.3: Definition of the Outcome

The outcome of obstructive coronary artery disease is defined as ≥50% stenosis of any major epicardial coronary artery[123]. Quantification of stenosis on coronary CTA is usually based on visual assessment of the interpreter, however; in some cases, measuring the diameter of the artery and obtaining the ratio of stenosis segment to normal artery can help to quantify the stenosis. Several studies have shown that degree of maximal stenosis on coronary CTA has a good correlation with findings of stenosis on invasive angiograms with greater standard deviation[123]. Quantitative findings on coronary CTA can predict stenosis on invasive coronary angiograms within 25% on comparative studies. This cut limit of 50% stenosis was used as obstructive coronary artery disease definition which is the definition of moderate stenosis or more according to the guidelines by the Society of Cardiovascular Computed Tomography(SCCT)[123].

3.3.3.1: Diagnostic Accuracy of CTA for Diagnosis of Coronary Artery Disease

Since coronary CTA emerged to play a role in diagnosis of coronary artery disease, several studies were conducted to evaluate its diagnostic accuracy across populations and using different technologies. In a prospective multicenter study (ACCURACY trial) by Budoff et. al., coronary CTA had a NPV of 99% for obstructive CAD when compared to conventional coronary angiogram[26]. Miller et. al. conducted another multicenter prospective study and showed that coronary CTA has a sensitivity and specificity of 83% and 90% respectively for identifying obstructive CAD compared to conventional angiogram[27].

Table 4: Definition of the main clinical variables in the derivation dataset

Variable	Type	Definition
Age	Continuous	Age in years
Sex	Dichotomous	Male, Female
Indication for the test	Categorical	The clinical indication (e.g. chest pain, shortness of breath or palpitations etc..)
Height	Continuous	Measured in centimeters during interview
Weight	Continuous	Measured in kilograms during interview
Body mass index (BMI)	Continuous	Calculated from height and weight
Prior diagnosis of CAD	Categorical (Yes, No, Unknown)	Any confirmed or suspected diagnosis of CAD on a prior assessment based on medical chart or patient reporting
Prior diagnosis of myocardial infarction	Categorical (Yes, No, Unknown)	Any confirmed or suspected diagnosis of myocardial infarction on a prior assessment based on medical chart or patient reporting
Prior coronary angiogram	Categorical (Yes, No, Unknown)	Any confirmed prior coronary angiogram based on medical chart or patient reporting
Prior coronary arteries revascularization	Categorical (Yes, No, Unknown)	Any confirmed prior coronary angioplasty, stents or bypass graft surgery based on medical chart or patient reporting
Family history	Categorical (Yes, No, Unknown)	History of CAD in 1st degree relative at age <55 years for males, and <65 years for females
Hypertension	Categorical (Yes, N, Unknown)	Prior diagnosis of hypertension (blood pressure >140/90 mmHg) based on medical chart or patient reporting; or treatment with blood pressure lowering medications
Diabetes	Categorical (Yes, No, Unknown)	Prior diagnosis of diabetes based on medical chart or patient reporting; or treatment with insulin or oral hypoglycemic agents
Hyperlipidemia	Categorical (Yes, No, Unknown)	Prior diagnosis of hyperlipidemia based on medical chart or patient reporting; or treatment with lipid lowering medications

Variable	Type	Definition
Smoking status	Categorical	Current smoking, ex-smoking, never smoked or unknown
Peripheral artery disease	Categorical (Yes, No, Unknown,)	Prior diagnosis of peripheral arterial disease based on medical chart or patient reporting
Cerebrovascular disease	Categorical (Yes, No Unknown,)	Prior diagnosis of cerebrovascular disease based on medical chart or patient reporting
Chest pain	Dichotomous (Yes, No)	History of chest pain as reported by the patient during the interview
Quality of chest pain	Dichotomous (Yes, No)	Presence of retrosternal chest heaviness
Duration of chest pain	Continuous	Measured in minutes for each episode of chest pain
Frequency	Continuous	Number of times of chest pain
Triggered chest pain	Dichotomous (Yes, No,)	Chest pain is triggered by exertion or emotional stress
Relieved chest pain	Dichotomous (Yes, No)	Chest pain is relieved by rest or use of nitro spray
Shortness of breath	Dichotomous (Yes, No)	History of shortness of breath as reported by the patient
Palpitations	Dichotomous (Yes, No, Unknown)	History of palpitations as reported by the patient
Syncope	Dichotomous (Yes, No, Unknown)	History of shortness of breath as reported by the patient
Medications: - Aspirin - Beta blockers - Calcium channel blockers - Angiotensin converting enzyme inhibitors - Nitrates - Lipid lowering medications	Dichotomous (Yes, No Unknown,)	Treatment with any of the medications as reported by the patient or documented on medication list of the patient

Variable	Type	Definition
Baseline systolic and diastolic blood pressure	Continuous	Blood pressure measured during the interview
Baseline heart rate	Continuous	Heart rate counted during the interview
Electrocardiogram (ECG) Findings - Normal - ST segment elevation - ST segment depression - Q wave - T wave inversion	Dichotomous for each level (Yes, No, Unknown)	Findings of ECG performed during the interview as interpreted by the physician interpreting the coronary CTA.
Probability of CAD	Continuous	Calculated after the assessment by the nurse during the interview using updated Diamond-Forrestor model[110]

In a multicenter multivendor prospective study from Europe , coronary CTA was found to have a sensitivity for identifying obstructive CAD of 99% , a specificity of 68%,PPV of 86% and NPV of 97% [25]. Authors concluded that coronary CTA may over-estimate the degree of stenosis especially in a population with high prevalence of CAD. Multiple pooled analyses of the published studies to investigate the diagnostic accuracy of coronary CTA found an over-all high accuracy for CTA in diagnosis of CAD with the highest performance in low to intermediate pre-test probability[30,124–126].

3.4: Data Preparation and Explanatory Analysis

The data set from UOHI Cardiac CT registry included patients enrolled between March 1st, 2006 and July 31st, 2016 and have calcium score of zero since we aim at predicting the risk of obstructive coronary artery disease in patients with no coronary calcification on cardiac CT. We also excluded all patients with previous coronary artery disease, coronary artery bypass graft or previous percutaneous coronary intervention. The final data set included 4,903 patients and 31 variables obtained from the clinical assessment by history, physical examination and electrocardiographic (ECG) findings.

3.4.1: Data Exploration and Variables Recoding

A list of candidate variables for model derivation was prepared based on the content knowledge and data availability. The registry database was designed to include all possible data which can be obtained during the interview prior to performing coronary artery CTA. List of the candidate variables, their types (after variables recoding) and the method of measurement of variables are presented in Table 5.

We examined each candidate variable distribution to assess for characteristics of data skewness, sparseness and overall shape of distribution. Although the quality of our data set is well-controlled, we examined for presence of extreme values and values that have no biological plausibility of continuous variables using descriptive statistics like univariate procedure and histograms. These values were corrected when possible and set to be missing otherwise. We also investigated for abnormal entries of categorical variables by constructing frequency tables and plots of these variables. Using frequency tables, distribution of categorical variables was assessed.

Diagnosis of diabetes mellitus is represented in the data set by two independent variables, diabetes mellitus type one and diabetes mellitus type two. Patients' knowledge regarding type of diabetes during interview is crucial to avoid misclassification, however, they may not be accurate as some patients would not know the type and others may not be labeled accurately especially if they developed diabetes mellitus in early adulthood. Relying on the type of treatment to decide on the type of diabetes, as patients on oral agents are usually diagnosed with type two diabetes, is not appropriate since many patients with diabetes type two are on insulin due to inadequacy of oral agents.

All categorical variables with values (responses) of (YES, NO and UNKNOWN) were dichotomized to two categories (YES, NO) by recoding the observations with "UNKNOWN" value (responses) to have "NO" value as they are unlikely to have the risk factor or the medication without any documentation or prior knowledge. We collapsed diabetes mellitus to one variable given the fact that risk of obstructive coronary artery disease is not different between the two entities and the possibility of misclassification of diabetes mellitus type.

We created three variables describing chest pain. Typical chest pain which is defined as presence of retrosternal chest pain that is triggered by exertion or emotional stress and relieved by

rest or use of nitroglycerine spray. Atypical chest pain was defined as retrosternal chest pain that is either triggered by exertion or emotional stress, or relieved by rest or nitroglycerine use. The non-cardiac chest pain is any chest pain that is not clearly triggered and relieved. Information regarding presence of chest pain, trigger of chest pain by exertion or emotional stress and relieve of chest pain by rest or nitroglycerin spray are collected during the interview prior to performing the coronary CTA. Due to the difficulty to differentiate atypical chest pain from non-cardiac chest pain in clinical practice in our experience and to simplify the approach to chest pain types as predictors, we combined the variables of atypical chest pain and non-cardiac chest pain into one variable, atypical/non-cardiac chest pain.

Coronary artery stenosis in our data set is recorded as a percentage of luminal narrowing in any coronary artery segment as mild (<50% stenosis), moderate (50-69% stenosis), significant (≥70) or occluded (100% stenosis). Otherwise, the artery or the segment is normal, absent or non-evaluable. There are 17 segments of coronary arteries assessed in our database as recommended by the guidelines from the Society of Cardiovascular CT (Appendix 4)[123]. In general, coronary artery segment is non-evaluable if the test interpreter is not able to delineate the segment well during the test for several reasons, usually severe calcification, motion or any other technical artifact. The risk of obstructive coronary artery disease is increased if the reason for non-evaluable artery is severe calcification which is not the case in our cohort. Absent coronary artery or any segments are usually normal variants and there is no evidence of association between variations of coronary artery anatomy and the risk of obstructive coronary artery disease. We dichotomized the coronary arteries findings in our database into either an obstructive CAD which was defined as presence of any segment with a luminal stenosis of ≥50% including total occlusion or non-obstructive CAD, which is absence of any segment with a stenosis of ≥50% (normal coronary artery segments or maximum stenosis that is < 50%).

Table 5: Candidate predictors for inclusion in the multivariable model

Variable	Type	Measurement Method
Age	Continuous	Age in years
Sex	Dichotomous	Male, Female
Indication for the test	Categorical	The clinical indication as stated in test request by the treating physician (chest pain, shortness of breath, palpitations, others)
Body mass index (BMI)	Continuous	Calculated from height and weight during the interview prior to the test
Family history	Dichotomous (Yes, No)	As patient report or from the medical chart including the request for the test
Hypertension	Dichotomous (Yes, No)	As patient report or from the medical chart including the request for the test
Diabetes	Dichotomous (Yes, No)	As patient report or from the medical chart including the request for the test
Hyperlipidemia	Dichotomous (Yes, No)	As patient report or from the medical chart including the request for the test
Current Smoking	Dichotomous (Yes, No)	As patient report or from the medical chart including the request for the test
Typical chest pain (if meets all three conditions: 1. Retrosternal chest pain 2. Triggered by exertion or stress 3. Relieved by rest or nitroglycerine spray use)	Dichotomous (Yes, No)	As patient report or from the medical chart including the request for the test
Atypical/ Non-cardiac chest pain (if meets one or two of three conditions, as in typical chest pain, including number)	Dichotomous (Yes, No)	As patient report or from the medical chart including the request for the test

Variable	Type	Measurement Method
Shortness of breath	Dichotomous (Yes, No)	As patient report or from the medical chart including the request for the test
Palpitations	Dichotomous (Yes, No)	As patient report or from the medical chart including the request for the test
Medications: - Aspirin - Beta blockers - Calcium channel blockers - Angiotensin converting enzyme inhibitors - Lipid lowering medications	Dichotomous (Yes, No)	As patient report or from the medical chart including the request for the test
Baseline systolic blood pressure	Continuous	Measurement during the interview
Baseline systolic blood pressure	Continuous	Measurement during the interview
Baseline heart rate	Continuous	Counting of heart rate during the interview
Ischemic electrocardiogram (ECG) Findings	Dichotomous (Yes, No)	Findings of ECG performed during the interview as interpreted by the physician interpreting the CTA
Probability of CAD	Continuous	Calculated after the assessment by the nurse during the interview using updated Diamond-Forrester model[110]

3.4.2: Statistical Summarization

For descriptive purposes, we calculated the proportion of the categorical variables and the means with standard deviation of continuous variables for those with and without obstructive coronary artery disease.

We compared the association between different categorical variables and the outcome between the two groups using the chi-square test. For continuous variables, we used t-test to compare the association between these variables and the outcome between these two groups.

3.4.3: Missing Data

The proportion of missing values of each variable in the list of predictors or the outcome was calculated using frequency tables for categorical variables and univariate analysis for continuous variable. We investigated for the mechanism of missing data and if there is any significance of the missing values on the association between predictors and outcomes by comparing this association among patients with completed data and those with any missing value.

3.5: Overview of Model Derivation Approaches

To derive a model for prediction of obstructive coronary artery disease among patients with no coronary artery calcification, we developed two models using two distinct approaches. First approach is based on a full pre-specified model and the second is based on a reduced model using statistical significance of the predictors through a combination of univariate screening of the predictors and stepwise selection of predictors in a multivariable model. Model pre-specification approach emphasizes minimizing the statistical overfitting associated with data-driven modelling but instead; predictors are defined in priori based on content knowledge and available literature. In the following two sections, we will review some of the features of each of the two approaches.

3.5.1: Model Pre-specification (Full Model) Approach

Pre-specification of a model predictors based on a prior knowledge or clinical content is recommended in clinical modelling whenever possible[70,73]. This includes specification of predictors, method of transformation of continuous predictors, categories of categorical variables and possible interactions.

Knowledge about the content of modelling is essential for pre-selection of the main effects (predictors) [70]. This knowledge can be obtained from a prior work or clinical experience which is important in reflection of the end user understanding of the problem being investigated. Specifying

a list of predictors for a clinical model should consider the effective sample size and number of events to avoid overfitting.

Distribution and frequencies of the pre-specified list of predictors should be assessed to investigate for missing values and need for imputation, skewness and similarity between variables and possibility of combining these variables. Variables with missing values can be imputed or deleted especially if not important [70].

3.5.2: Univariate Screening (Reduced Model) Approach

Another approach of specification of clinical prediction models is developing a reduced model by selection of predictors from a list of potential predictors based on their statistical significance. Knowledge about the content topic is less influential in deciding which predictors should be included in the modelling process [73]. Two methods are used for selection of predictors to be included in the final model, univariate screening of predictors and stepwise selection of predictors during multivariable modelling[73].

Modelling reduced models in clinical research is common nowadays. Several arguments are in favor for reduced models. The most prevalent is smaller model is more practical and more likely to be used while larger and more complex models are not. Larger models are likely to have collinearity and strong correlation between included variables, however; this issue can be investigated and resolved using available statistical techniques. Finally, including predictors based on statistical significance is likely to lead to a significant predictor that was not known to be, which make such process as an exploratory process. In the next two sections, we present an overview of the two statistical techniques of selecting predictors in modelling.

3.5.2.1: Univariate Predictors Screening

In this approach, the relationship between individual predictors and the outcome is assessed first for inclusion in a regression model. Based on univariate criterion usually a preset p value, a variable is considered for multivariable modelling [70]. This practice aims in most of the cases to reduce the number of included variables or to overcome the lack of knowledge regarding predictors of clinical importance that should be included in the model by exploring the univariable association between predictors and outcome. Eliminating certain predictors at earlier stages before

multivariable modelling and reduction of computational burden associated with large dataset with many predictors are key features of using this method.

3.5.2.2: Step-wise Selection of Predictors

Step-wise selection is an automated process of variable selection to be included in a final model based on significance which is determined by a criterion (e.g. Wald test or score statistics in logistic regression models). It is the most popular technique used to select predictors for inclusion in a multivariable clinical model development and includes only statistically significant predictors [70]. Step-wise selection can be forward where a significant predictor is added to a model with no predictors initially. Backward step-wise selection is accomplished by eliminating non-significant predictors from a full model. A combination of both forward and backward step-wise selection can be done iteratively. Step-wise selection is helpful in modelling as an objective tool that yield smaller models that are easy to use and identify unknown predictors based on statistical significance [70].

3.5.2.3: Issues of Statistically-Driven Predictors Selection

Selection of predictors in clinical modelling based solely on statistical significance either by using a conditional p-value in a univariate analysis or using step-wise selection in a multivariable modelling has few concerning issues. Step-wise selection may lead to multiple combinations that are similar which indicate unstable selection of predictors[127]. Instability of selection could be exaggerated by small sample size. Estimation of predictors contribution in clinical models after testing them either in univariate models or by step-wise selection may be a source of biased estimation of coefficients and events per variable [70]. In stepwise selection, all possible combinations of predictors are tested for the model with best fit prior to estimation of predictions. As a result, the distribution of the regression coefficients estimation is difficult to quantify or interpret due to misspecification of variability and highly variable predictions, which leads to exaggeration of p values[70,128,129]. Results of regression models by stepwise selection are biased and yields usually high coefficients with small standard errors and p values [73]. Although multivariable models should control for confounders, but automated selection of variables is associated with residual confounding [70]. In many occasions, quality of the derived model using

stepwise selection is lower than that of a full model when validated either internally or externally [80].

3.6: Model Development - Pre-Specification Approach (Full Model)

After we prepared our data for model derivation, performed our exploratory analysis including statistical summarization and assessment of variables distribution, we derived our proposed model using the model pre-specification approach in three phases; Predictors specification by content experts, multivariable modelling and finally assessment of model performance.

3.6.1: Model Pre-Specification

We specified model predictors a priori in two stages. First, we reviewed literature to identify the predictors included in models predicting coronary artery disease in general. Second; we surveyed 10 practicing cardiologists, who are considered content-experts and represent the targeted end-users for such model (Appendix 5 for the survey). Based on the number of outcome events in our dataset, 8 parameters were the maximum number could be included in development of the new model. In our survey, we asked the practitioners to select a list of 10 predictors from a larger list of candidate predictors (Table 5), including some of those in previous models and other predictors considered clinically relevant. If other predictors were thought to be relevant, the surveyed practitioner was given the liberty to add to the list of ranked predictors. We pooled all responses and selected a final list of 8 predictors with the highest number of votes (as selected by surveyed cardiologists).

3.6.2: Multivariable Modelling

After specifying a list of predictors, multivariable modelling process was performed using logistic regression procedure in SAS 9.4© (SAS institute, Cary, NC, USA.). There was no need for any imputation as our derivation dataset was complete for the list of predictors selected for multivariable modelling. Continuous predictors were modelled as continuous variables without dichotomization or transformation.

The initial model included the final list of selected predictors based on the votes by the surveyed cardiologists. After development of the full model, deciding on subsequent steps for simplifying of the model by excluding any of the initial predictors was planned to be determined by weighing the predictive value of the predictor as measured by chi-square test in the model versus the clinical importance to be included in the final model.

Interaction between sex and typical chest pain was assessed. Several interaction terms were created and included in the modelling process depending on the possible interactions in such clinical context.

3.6.3: Assessment of Model Performance

For assessment of the developed model, we assessed the model discrimination ability and calibration. Assessment of model performance was conducted using statistical software SAS 9.4© (SAS institute, Cary, NC, USA.).

To assess the discriminative ability of the newly developed model, we constructed the receiver operating characteristics (ROC) curve, which plots the sensitivity against 1- specificity for a consecutive probability of having obstructive coronary artery disease[70,73]. The area under the curve (AUC) was used as a measure for discriminative ability of the model. It corresponds to the concordance statistics (c) which is commonly used to measure discrimination of linear models[70].

Hosmer-Lemeshow test was used for assessment of goodness-of fit of the newly developed model. We grouped observations by decile of predicted probabilities. The sum of the predicted probabilities in each decile is considered the expected outcome which is compared to the observed outcome for each decile (total of 10 groups) using χ^2 with reporting of p values. Adequate goodness of fit has a p value that is >0.05 [70].

3.7: Model Presentation

Clinical prediction models estimate the predicted probability of a certain outcome based on multiple characteristics (predictors). Interpretation of such predictions needs further knowledge regarding that outcome most of the time. Presentation of the model in a user-friendly format that could enable users to make decisions is crucial for dissemination and clinical feasibility. These formats of model presentation can serve as a decision rule which guide the proper next action based

on the predicted probability by the model[70]. There are several formats for presentation of clinical prediction models including regression formulas, nomograms and point score system and charts (for details see section 2.2.6). For presentation of our models, both full and reduced; we elected to present them in points scoring system format for its simplicity to use and interpret in a busy clinical practice and familiarity of clinicians with such scoring system.

3.7.1: Development of the Scoring System

After obtaining the final model, we used the regression coefficients to derive the points assigned for each predictor. We divided each categorical variable regression coefficient by the smallest absolute value of a dichotomous regression coefficient in the multivariable model [70102]. The results were rounded to the nearest integer. This value represents a weighing factor of the predictor in the scoring system. We assigned a scoring point for each predictor based on the weighing factor in 1:1 ratio.

3.7.1.1: Dealing with Continuous Predictors in The Score

Age was treated as a continuous variable in the multivariable model. In the score , age was categorized in 6 groups based on the observed probability of obstructive CAD in clinical practice and prior use of these age groups in clinical models[114]. Age groups are <30 years, 30 to 39 years, 40-49 years, 50-59 years ,60-69 years and ≥70 years.

To obtain a weighing factor for each class, first we obtained the regression coefficients for each class by calculating the mean age for patients in each class in the derivation cohort and multiplied it by the regression coefficient for age in the multivariable model. Then, the resulting coefficient of the age class <30 was taken as a reference and subtracted from the resulting coefficient of each age group to obtain a relative coefficient [102]. The relative coefficient for each age group was divided by the absolute smallest regression coefficient of a dichotomous variable in the multivariable model as for other variables to get the weighing factor for each class in the score.

3.7.2: Assessment of Score Performance

Each patient in the derivation data set has a total-points score based on their summation of the points for each predictor in the model. We calculated the predicted risk estimates for outcome of obstructive coronary artery disease using logistic regression model based on the total point score as a predictor. We constructed the receiver operating characteristics (ROC) curve for the score in the derivation cohort, which plots sensitivity against 1-specificty for probabilities of developing the outcome of obstructive CAD at each score threshold.

For assessment of the discriminative performance of the developed score, we calculated the area under the ROC curve of the score. We plotted the predicted and the observed probabilities for each threshold of the score to assess the calibration of the developed score.

3.7.3: Diagnostic Characteristics of Score Threshold

We calculated the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (LR+) and negative likelihood ratio (LR-) at each cumulative point score threshold to assess the score classification ability of the patients based on the probabilities into low, intermediate and high-risk groups. We calculated the proportion of total patients and patients with confirmed obstructive CAD in each group to assess for the predictive accuracy of each risk group.

3.8: External Validation of the Score Derived from the Full Model

The main goal of validation is to assess whether the predictions of the derived model (score) can accurately predict the responses or outcomes in a different group of subjects that was not used in the process of model development (see section 2.2.5). Model failure to validate in different population can be a result of several reasons. These reasons are either related to derivation techniques like overfitting or to major differences between the derivation and validation cohorts in term of variables definitions, measurements or coding or study inclusion criteria[70]. To avoid these potential causes of failure, we chose a cohort that have similar outcome measurement technique and variables' definitions as our derivation cohort.

We performed an external validation using a data set from CONFIRM (Coronary CT Angiography Evaluation for Clinical Outcomes: An International Multicenter) registry[130].

3.8.1: Overview of Validation Cohort

CONFIRM (Coronary CT Angiography Evaluation for Clinical Outcomes: An International Multicenter) registry is an open-labeled prospective multi-center study of coronary CTA in different clinical contexts[130]. In this study, referred patients with or without coronary artery disease or asymptomatic patients for risk stratification were enrolled prospectively from 12 sites around the world (Appendix 6). Information related to demographics, targeted medical history and cardiovascular risk profiles were collected prospectively and recorded in a site-specific form. Coronary calcium score with or without coronary CTA were done and interpreted according to the guidelines of Society of Cardiac CT[123]. The forms and reports from different centers were merged all together at a later stage.

The CONFIRM registry contains two sections. Section 1 comprises 27,125 patients from 12 sites enrolled between January 2004 and May 2010 and was locked in October 2010. This section served as the test sample. Section 2 comprises 4,682 patients from 5 sites enrolled between July 2005 and October 2010 (database locked in May 2011) and served as validation sample. Institutional review board approval was obtained at each center[131].For our validation , we used the test sample of the registry.

3.8.2: Data Preparation of Validation Dataset

Data from CONFIRM registry was shared with us as our institute is a part of the registry. Observations with calcium score of zero were extracted. We included patients with no known diagnosis of coronary artery disease, prior per cutaneous coronary intervention or prior coronary artery bypass grafts. We excluded patients with known heart failure, significant valvular heart disease, congenital heart disease, heart transplant or suspected of having diagnosis of acute coronary syndrome as an indication for the test. Further, we excluded data from our institute to avoid any possible overlapping data. The remaining data set represents the primary validation dataset and was used for further explanatory analysis.

3.8.3: Exploratory Analysis and Variables Recoding

The validation data set includes data on demographics, risk factors for coronary artery disease, clinical history and physical examination findings and findings of coronary artery CTA. Data on variables included in our model are available in the validation dataset. Variables were defined and coded in a similar way to the definition and coding of variables in the derivation data set (see table 5).

We examined each variable distribution to assess for characteristics of data skewness, sparseness and overall shape of distribution. We examined for presence of extreme values and values that have no biological plausibility of continuous variables using descriptive statistics like univariate procedure and histograms. These values were corrected when possible and set to be missing otherwise. We also investigated for abnormal entries of categorical variables by constructing frequency tables and plots of these variables. Using frequency tables, distribution of categorical variables was assessed.

Coronary artery disease risk factors were collected as dichotomous variables with responses reported by the patient as present or absent risk factor. Chest pain typicality was defined similar in both derivation and validation cohorts. Typical chest pain is a retrosternal chest pain that is triggered by exertion or emotional stress and relieved by rest or Nitro spray use. Atypical chest pain was defined by lack of either the triggering or the relieving factor; and non-cardiac pain was defined by lack of both triggering and relieving factors. Atypical and non-cardiac pain were collapsed to one class of atypical or non-cardiac chest pain for the noticeable clinical overlap between the two classes. Categorical variables with responses of "unknown" values were set to "No" when present as in the derivation dataset.

Coronary arteries anatomy findings as assessed by coronary CTA were classified as normal, mild non- obstructive CAD (luminal stenosis of less than 50%) or obstructive CAD (luminal stenosis of 50% or more). To dichotomize the outcome, obstructive CAD was defined as any luminal coronary artery stenosis of 50 % or more as per the guidelines by SCCT and hence the normal and mild non-obstructive disease classes were collapsed into one category [123].

3.8.4: Dealing with Missing Values in Validation Dataset

Proportion of missing values of each variable in the list of predictors or the outcome was calculated using frequency tables for categorical variables and univariate analysis for continuous

variable. Missing pattern was examined using MI procedure (zero imputations) in the statistical software SAS 9.4© (SAS institute, Cary, NC, USA.).

We assessed for association between the magnitude of missing data and the site to identify institute with largest proportion of missing values. There were sites with large proportion of missing data especially on chest pain and on criteria used to define chest pain typicality. These sites were excluded from the final data set for validation if missing values on any of our model predictors exceed 20% of all observations in that site. These excluded sites were deemed of low quality on data collection due to the significantly large missing proportions on variables that are known to be important in such registry.

After exclusion of the centers with large missing values on any of the model predictors, the final validation dataset included 4,290 patients from 8 centers around the world. We compared centers included in the final validation dataset with those excluded in term of clinical characteristics using Fisher exact test for categorical variables and t-test for continuous ones. To account for the effect of potential imbalance in sample size between the two groups, we used standardized difference between means of continuous variables and standardized difference between proportions of categorical variables if needed when compare the two groups.

3.8.4.1: Multiple Imputations of Missing Values in Final Validation Dataset

After exclusion of centers with large missing proportions that exceed 20% for any of the predictors in our proposed multivariable model, the final validation dataset has some residual missing values. We performed multiple imputations by chained equation approach to produce 10 imputed datasets. We used MI procedure in in the statistical software SAS 9.4© (SAS institute, Cary, NC, USA.).

3.8.5: Score Validation

After completion of validation dataset preparation which included exploration of variables distribution, frequency, identification of extreme and non-sense values and excluding of centers that have significant proportion of missing values; a final sub-group from CONFIRM registry of

total 4,290 observations was obtained for the validation of the developed score derived by model pre-specification (full model) approach.

For validation of the score in the external validation dataset, we first assessed the relatedness of case mix and outcome of both derivation and validation cohort. Secondly, we assessed the performance the developed score in the validation cohort. Finally, we calculated the diagnostic characteristics of the selected risk thresholds and the proportion of patients in each risk category in the validation cohort.

3.8.5.1: Relatedness of Derivation and Validation Cohorts

For evaluation of relatedness between the derivation and validation cohort, we compared the case mix of both cohorts for variables and outcome occurrence. We used proportions for categorical variables and means, standard deviation and ranges for continuous variables[132].

3.8.5.2: Assessment of Score Performance in Validation Dataset

After evaluation for relatedness of the derivation and validation datasets, we assessed the diagnostic performance of the developed score to predict obstructive CAD in the validation dataset, using AUC for the ROC curve that plots sensitivity against 1-specificty for probabilities of developing the outcome of obstructive CAD at each score threshold. We compared AUC of the developed score in both derivation and validation datasets. For assessment of calibration of the score in the validation dataset, we plotted the observed versus predicted probabilities for each score threshold.

3.8.5.3: Diagnostic Characteristics of the Score Thresholds in Validation Cohort

As a part of score validation, we calculated the diagnostic characteristics of the selected score threshold to define the risk. Sensitivity, specificity, positive and negative predictive values and likelihood ratios for the thresholds used to classify the risk of developing outcome were calculated in validation cohort and compared to the ones in the derivation cohort.

To evaluate for predictive accuracy of each risk class, we calculated the proportion of all patients and patients who developed obstructive CAD in the class.

3.9: Model development - Reduced Model

To derive a reduced model by selection of the predictors during modelling based on their statistical significance, we prepared a list of candidate predictors based on the available data (Table 5). We compared the association between each of the potential predictors and the outcome between patients with and without obstructive CAD (see section 3.3.2). After that, modelling process was conducted in three phases, multiple imputations of missing values of candidate predictors; univariate screening for statistically significant variables and a multivariable modelling phase.

3.9.1: Multiple Imputations

Despite the completeness of predictors in the full model in our derivation dataset, some of the candidate predictors that are potential for our reduced model have missing values. We performed multiple imputations by chained equation approach to produce 10 imputed datasets. We used MI procedure in the statistical software SAS 9.4© (SAS institute, Cary, NC, USA.). The multiple imputations were performed using an imputation model that included the analysis variables (predictors and outcome of interest) and auxiliary variables , which are variables that are not part of the prediction model but could contribute some information about the missing values[133]. Imputation dataset is produced after 100 iterations of the imputation procedure. Continuous variables were treated as continuous in the imputation model, while categorical variables were rounded up to the nearest plausible categorical variable. We assessed the stability of the imputations using the trace and autocorrelation plots. Trace plot graph the regression coefficients of the imputation model throughout iteration of data augmentation, and autocorrelation plot depicts the difference between regression coefficients for different lag intervals of data augmentation procedure [133].

3.9.2: Univariate Screening

In the phase of univariate screening, we performed a univariate analysis for each variable in the list of candidate variable to test the association between each variable and the outcome of obstructive coronary artery disease. We used chi square test with correction for continuity if needed or Fisher's exact test where appropriate for categorical variables. For continuous variables,

we used two tailed t-test. Variables with a p value of 0.05 or less was included in a multivariable logistic regression.

3.9.3: Multivariable Modelling- Reduced Model

In the multivariable modelling phase, all variables with p value of 0.1 or less were included in an initial full model. We performed an automated selection of predictors included the final model using backward, forward and stepwise selection technique to compare the predictors selected by all three methods.

Backward stepwise selection was chosen to develop the final reduced model because of its reasonable technique. Elimination of predictors for a full model is favored over adding predictors to a null model (model with no predictors) especially if collinearity between predictors is present[73].

3.9.4: Assessment of Model Performance

To assess the performance of the multivariable model to predict the presence of obstructive CAD using univariate screening approach (reduced model), we used the same measures used for assessment of the performance of the model developed using the model pre-specification approach (full model) as discussed previously (see section 3.4.3). We used AUC the ROC curve to assess the discriminative ability of the developed model and Hosmer-Lemeshow to assess the goodness of fit. We plotted the predicted probabilities grouped in deciles versus their observed risk of the outcome to assess the calibration of the model.

3.10: Internal Validation- Reduced Model

For validation of the model developed by univariate screening, we performed internal validation using the bootstrap technique, since some of the predictors in the final multivariable model were not available in the external validation dataset.

3.10.1: Bootstrapping Technique

Bootstrapping technique was used for resampling of 1000 datasets with the same number of observations as the original dataset using PROC SURVEYSELECT procedure in SAS 9.4©

statistical software package (SAS institute, Cary, NC, USA.). We used unrestricted random sampling method which is simple resampling with replacement. Then, we fitted the initial multivariable model with the backward selection in each of the bootstrap dataset to assess for the consistency of the predictors' selection across all bootstraps. We calculated the frequency of times each predictor was selected. We pooled the estimated regression coefficients and model performance in each dataset.

3.11: Model Presentation

Clinical prediction models estimate the predicted probability of a certain outcome based on multiple characteristics (predictors). Interpretation of such predictions needs further knowledge regarding that outcome most of the time. Presentation of the model in a user-friendly format that could enable users to make decisions is crucial for dissemination and clinical feasibility. These formats of model presentation can serve as a decision rule which guide the proper next action based on the predicted probability by the model[70]. There are several formats for presentation of clinical prediction models including regression formulas, nomograms and point score system and charts (for details see section 2.2.6). For presentation of our models, both full and reduced; we elected to present them in points scoring system format for its simplicity to use and interpret in a busy clinical practice and familiarity of clinicians with such scoring system.

3.11.1: Development of the Scoring System

After obtaining the final model, we used the regression coefficients to derive the points assigned for each predictor. We divided each categorical variable regression coefficient by the smallest absolute value of a dichotomous regression coefficient in the multivariable model [70102]. The results were rounded to the nearest integer. This value represents a weighing factor of the predictor in the scoring system. We assigned a scoring point for each predictor based on the weighing factor in 1:1 ratio.

3.11.1.1: Dealing with Continuous Predictors in The Score

Age was treated as a continuous variable in the multivariable model. In the score , age was categorized in 6 groups based on the observed probability of obstructive CAD in clinical practice

and prior use of these age groups in clinical models[114]. Age groups are <30 years, 30 to 39 years, 40-49 years, 50-59 years ,60-69 years and ≥70 years.

To obtain a weighing factor for each class, first we obtained regression coefficients for each class by calculating the mean age for patients in each class in the derivation cohort and multiplied it by the regression coefficient for age in the multivariable model. Then the resulting coefficient of the age class <30 was taken as a reference and subtracted from the resulting coefficient of each age group to obtain a relative coefficient [102]. The relative coefficient for each age group was divided by the absolute smallest regression coefficient of a dichotomous variable in the multivariable model as for other variables to get the weighing factor which was.

3.11.2: Assessment of Score Performance

Each patient in the derivation data set has a total-points score based on their summation of the points for each predictor in the model. We calculated the predicted risk estimates for outcome of obstructive coronary artery disease using logistic regression model based on the total point score as a predictor. We constructed the receiver operating characteristics (ROC) curve for the score in the derivation cohort, which plots sensitivity against 1-specificty for probabilities of developing the outcome of obstructive CAD at each score threshold.

For assessment of the discriminative performance of the developed score, we calculated the area under the ROC curve of the score. We plotted the predicted and the observed probabilities for each threshold of the score to assess the calibration of the developed score.

3.11.3: Diagnostic Characteristics of Score Threshold

We calculated the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (LR+) and negative likelihood ratio (LR-) at each cumulative point score threshold to assess the score classification ability of the patients based on the probabilities into low, intermediate and high-risk groups. We calculated the proportion of total patients and patients with confirmed obstructive CAD in each group to assess for the predictive accuracy of each risk group.

CHAPTER 4: RESULTS

In this chapter, we present the results of data exploration and preparation of both the derivation and validation datasets, multivariable modelling of both the full and reduced models, results of the assessment of the models' performance, the point score system derived from each model with their performance measure and the decision rule thresholds, and finally; the external validation of the score derived from the full model and the internal validation of the reduced model.

4.1: Derivation Dataset Preparation

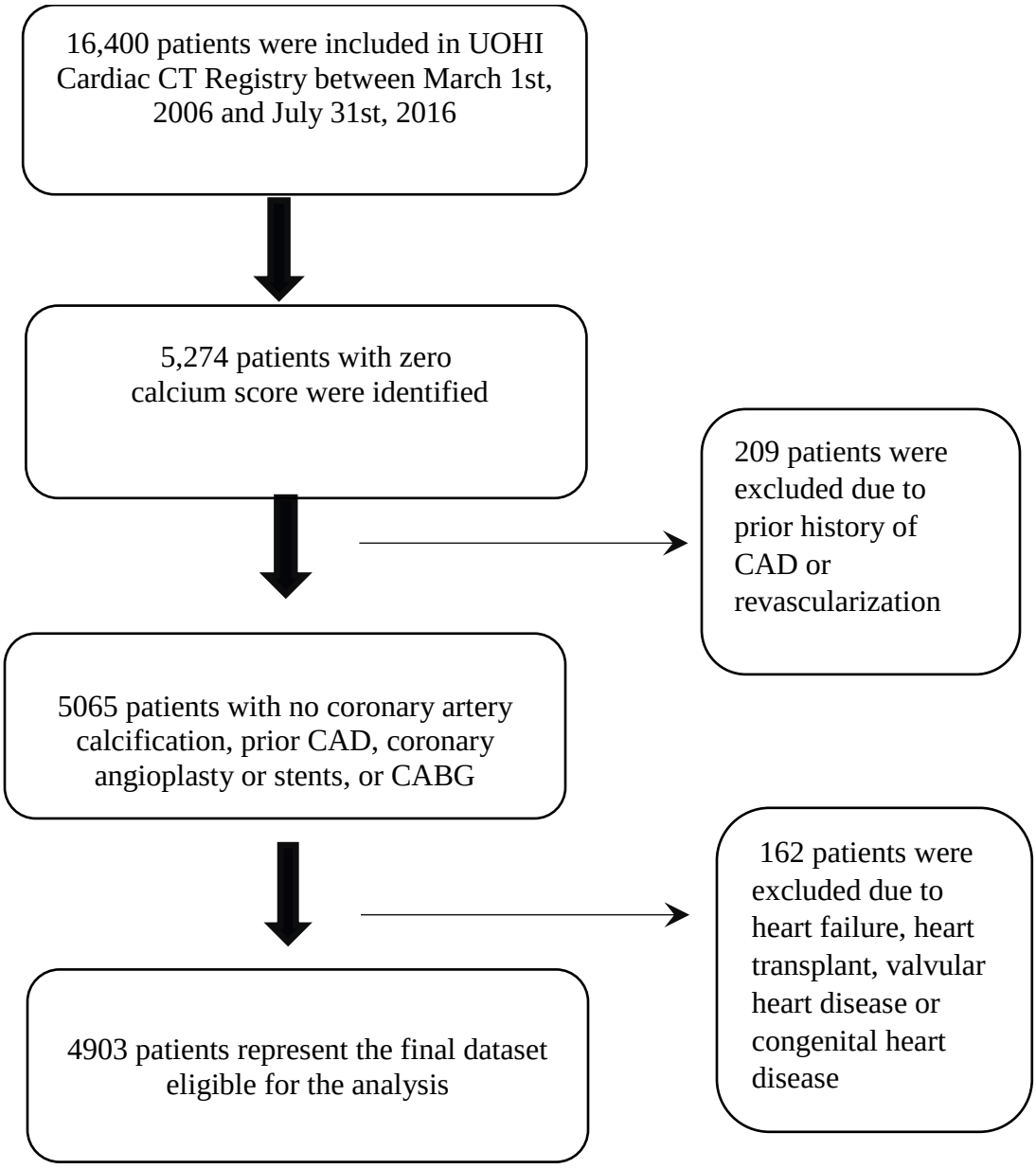
The dataset used for derivation of our model was obtained from UOHI cardiac CT registry, and included a total of 16,400 patients enrolled between March 1st, 2006 and July 31st, 2016. We identified 5,274 patients with calcium score of zero (32%) among all patients enrolled in this period. After exclusion of patients with prior coronary artery disease, myocardial infarction, coronary artery by-pass graft surgery or coronary artery angioplasty or stents (n=209, 3.9%), a total of 5,065 patients were eligible for our analysis to derive a clinical prediction model to predict obstructive coronary artery disease. We further excluded patients who had "Unknown" status of any of the risk factors for coronary artery disease. There was a total number of 162 patients (3%) who were excluded due to unknown status of CAD risk factors.

The final dataset for our analysis included a total number of 4,903 participants with zero calcium score and no prior history of coronary artery disease, coronary artery intervention or coronary artery by-pass graft surgery. Preparation flow chart is presented in Figure 2.

4.1.1: Characteristics of Participants in Derivation Dataset

Characteristics of the patients enrolled in the UOHI cardiac CT registry included in the final derivation dataset is presented in Table 6. The majority were females, 58% versus 42%, and the mean age was 53 years (SD:10.3). The outcome of obstructive coronary artery disease, as evaluated by coronary artery CTA; was found in 112 patients (2.28%). Chest pain work up was the indication of the test in 2,793 patients (57%), followed by work-up for shortness of breath in 1,233 patients (25%).

Figure 2: Flow chart of derivation dataset preparation



*Abbreviations: CAD, coronary artery disease; CABG, coronary artery bypass graft;

Positive family history was the most prevalent risk factor in our dataset and was found in 2250 participants (45.89%). Diagnosis with hyperlipidemia and hypertension were found in 2037(41.5%) and 1892(39%) participants respectively. Other risk factors were less prevalent with only 665(13%) were current smokers, 466 (9.5%) participants with diagnosis of diabetes mellitus and 208(4.2%) with peripheral arterial disease.

More than 67% of the patients in our dataset have at least one medication with a cardiovascular effect or effect on lipids prescribed. Aspirin was taken on daily basis by 2073 participants (41.55%), beta blockers by 1648 participants (33.61%) and lipid lowering agents by 1457 participants (30%).

Symptomatic status of the patients in our dataset at the time of the clinical interview, just before the test; revealed 3,071 (63%) with a history of chest pain; 1,104 participants (23%) with a chest pain that is triggered by physical exertion or emotional stress and 614 participants (13%) with chest pain that is relieved by nitro spray or rest. Shortness of breath and palpitations were prevalent symptoms in patients included in our dataset reported in 2996 (61%) and 2826 (58%) participants respectively. The mean pre-test probability of coronary artery disease as calculated by the updated Diamond-Forrester prediction rule was 22.6% which is in the intermediate range [110].

Findings of physical assessment during the clinical interview before performing the test revealed that mean systolic and diastolic blood pressure, and mean heart rates of our study patients are within normal values. The body mass index of the derivation cohort is 29.6 (SD 6.3) which indicates that our study patients are overall overweight. Electrocardiographic changes that are considered ischemic changes were documented in 37% of the participants, when EKG was performed during the interview.

4.1.3: Summary Statistics of Predictors and Outcome

Table 7 provide a list of the candidate predictors for modelling of obstructive coronary artery disease based on the available data collected as a part of the ongoing registry of cardiac CT at UOHI and clinical input from content experts, and their summary statistics. The summary results of the analysis of candidate variables including their type, frequency, distribution, and proportion

of missing values. We observe from Table 7 that data is complete for most of the candidate predictors.

4.2: Results of Full Model (Prespecified Model Approach)

This approach of model development used a pre-specification of a list of predictors based on input from content experts or end users and is conducted in three phases. Model pre-specification, imputation of missing values of the specified predictors if needed, multivariable modelling and assessment of performance of the final model.

4.2.1: Model Specification

Pre-specification of predictors to be included in a multivariable model was based on surveying a group of practicing cardiologists who are the main end-user group of our tool. We asked the surveyed cardiologists to list 10 predictors from the list of the candidate predictors (Table 5). We received responses from 6 cardiologists representing 60% response rate to our survey. The responses of our surveyed cardiologists are summarized in Table 8.

All respondents (100%) selected age, gender, chest pain and diabetes mellitus to be included in the model. Hypertension was chosen by 5 respondent cardiologists (83%). Family history, hyperlipidemia and smoking were selected by 4 respondent cardiologists (67%). The rest of the candidate variables were selected by 50% or less of the respondent cardiologists. Variables including height, weight, baseline blood pressure both systolic and diastolic and heart rate were not selected by any of the respondents.

Table 6: Clinical Characteristics of derivation dataset from UOHI cardiac CT registry

Characteristics	**N (%)** **Mean (SD)**
Age	53(10.31)
Male	2,058 (41.97)
Female	2,845 (58.03)
Indication	
Chest pain	2,793 (56.96)
Shortness of breath	1,233 (25.05)
Atrial fibrillation	239 (4.87)
Others	638 (13.01)
Family History	2,250(45.89)
Hyperlipidemia	2,037(41.55)
Hypertension	1,892 (38.59)
Diabetes Mellitus	466 (9.50)
Current Smoking	665 (13.56)
Medications	
Aspirin	2,073(42.28)
Beta Blockers	1,648(33.61)
Calcium channel blockers	478(9.75)
Angiotensin Converting enzyme inhibitors (ACE inh)	746(15.21)
Lipid lowering agents	1,457(29.71)

Characteristics	**N (%)** **Mean (SD)**
Symptoms	
Typical chest pain	3,071 (62.64)
Atypical chest pain	1,104 (22.52)
Shortness of breath	2,996 (61.11)
Palpitations	2,826 (57.64)
Syncope	1,729 (35.26)
Pre-test probability of CAD	22 % (23.00)
Assessment prior to test	
Systolic blood pressure (mmHg)	130(18.00)
Diastolic blood pressure (mmHg)	78(10.08)
Heart Rate (bpm)	68(12.13)
Body mass index (BMI)	29 (6.34)
Creatinine	77 (19.15)
ECG findings	
Normal	2992 (61.02)
ST segment or T wave changes	1324 (27.00)
Other ischemic changes	487(9.93)
Other changes	100 (2.04)
Obstructive coronary artery disease	
Yes	112 (2.28)
No	4791(97.72)
Total	4903

Table 7: Summary statistics of the candidate predictors in the derivation dataset

Predictor	**Scale**	**Range /Distribution (Categories)**	**N (%)** **Mean (SD)**	**Missing** **n (%)**
Age	Continuous	(15,86)/ Normal	53 (10.3)	0
Gender	Categorical	Male Female	2058 (42) 2845 (58)	0
Chest pain typicality	Categorical	- Typical - Atypical/Non-cardiac - No chest pain	392(8) 2679(55) 1832 (37)	0
Shortness of breath	Categorical	Yes No	2996 (61) 1907 (39)	0
Palpitations	Categorical	Yes No	2826 (58) 2077 (42)	0
Family history	Categorical	Yes No	2250(46) 2653(54)	0
Hypertension Dx	Categorical	Yes No	1892 (39) 3011(61)	0
Hyperlipidemia Dx	Categorical	Yes No	2037(42) 2866(58)	0
Diabetes mellitus Dx	Categorical	Yes No	466 (9) 4458(91)	0
Current smoking	Categorical	Yes No	665 (14) 4238 (86)	0
Treatment with Aspirin	Categorical	Yes No	2073(42) 2830 (58)	0
Treatment with Beta blockers	Categorical	Yes No	1648(34) 3255 (66)	0

Predictor	**Scale**	**Range/Distribution (Categories)**	**N (%) Mean (SD)**	**Missing n (%)**
Treatment with CCB	Categorical	Yes No	478(10) 4425(90)	0
Treatment with Lipid lowering agents	Categorical	Yes, No	1457(30) 3446 (70)	0
Treatment with ACE inhibitors	Categorical	Yes No	746(15.2) 4157(85)	0
Pre-test probability of CAD	Continuous	0.3%,100% /Normal	21 % (23)	5(0.1)
BMI	Continuous	14,66/ Normal	29 (6.3)	2(0.04)
Systolic BP (mmHg)	Continuous	78,227/ Normal	130(18)	10(0.2)
Diastolic BP (mmHg)	Continuous	18,140 / Normal	78(10)	13(0.27)
Heart Rate (bpm)	Continuous	38,124/ Normal	68(12)	11(0.22)
Ischemic ECG changes	Categorical	Yes No	1705(35) 3198(65)	0
Abbreviations: ACE, Angiotensin converting enzyme; BP, blood pressure; Dx, diagnosis; ECG, Electrocardiogram;				

Treatment with aspirin and treatment with lipid lowering agent variables were selected by three respondents (50%), but treatment with beta blockers was selected by two respondents (33%) and none selected treatment with nitrates or treatment with ACE inhibitors.

We obtained the final list of 8 predictors to be included in the prespecified model from the predictors with the highest number of votes based on the respondent cardiologists surveyed. The

list of the top 8 predictors to be included in the multivariable model is presented in Table 9. Data are complete on predictors specified for multivariable model and there was no need for imputation.

4.2.2: Multivariable Modelling

After defining the final list of the predictors specified to be included in a multivariable model to predict obstructive coronary artery disease, an initial multivariable model was derived using logistic regression model in the statistical software SAS 9.4© (SAS institute, Cary, NC, USA.). We included the top 8 predictors selected by the surveyed cardiologists to enter the multivariable model (Table 9).

4.2.2.1: Initial Multivariable Logistic Regression Model

The initial logistic regression model included the variables selected in advance by the surveyed cardiologists which were age, gender, chest pain typicality, hyperlipidemia, hypertension, family history, diabetes mellitus and current smoking.

As shown in Table 10, typical chest pain has the highest odds ratio of 2.89 (CI:1.68-4.98) followed by male sex (OR:2.23, CI:1.48 - 3.38). These two variables have the strongest predictions for presence of obstructive coronary artery disease after controlling for other predictors in the multivariable model. Hyperlipidemia had a significant association with outcome of obstructive coronary disease with an odds ratio of 1.884 (CI:1.253-2.833). Age as a continuous predictor, hypertension, family history, diabetes mellitus and current smoking do not have significant association with the risk of developing obstructive coronary artery disease in this model. However, regression coefficients of these predictors showed an insignificant incremental risk of having obstructive CAD except for current smoking; which showed a paradoxical association to the known effect of increasing the risk of obstructive coronary artery disease. The univariate analysis of current smoking association with the outcome of obstructive CAD, current smoking was insignificantly associated with an increased risk of the outcome. The same incremental risk of the outcome of obstructive CAD was observed in the univariable and multivariable analysis of current smoking in the whole dataset including patients with and without coronary artery calcification. Due to this association between current smoking and the outcome which is not clinically plausible, we elected to exclude smoking both current or previous from the final model.

Table 8:Summary of responses to survey for predictors pre-selection

Predictor	**Responder 1**	**Responder 2**	**Responder 3**	**Responder 4**	**Responder 5**	**Responder 6**	**Total Votes (%)**
Age	✓	✓	✓	✓	✓	✓	6(100)
Sex	✓	✓	✓	✓	✓	✓	6(100)
Probability of CAD		✓	✓			✓	3(50)
BMI	✓	✓		✓			3(50)
Family history	✓		✓	✓		✓	4(67)
Hypertension	✓	✓	✓		✓	✓	5(83)
Diabetes	✓	✓	✓	✓	✓	✓	6(100)
Hyperlipidemia	✓	✓	✓		✓		4(67)
Smoking	✓	✓		✓			3(50)
Typical Chest pain	✓	✓	✓	✓	✓	✓	6(100)
Atypical chest pain					✓		1 (17)
Shortness of breath			✓	✓			2(33)
Palpitations							0
Syncope							0
Baseline systolic BP							0
Baseline diastolic BP							0
Baseline heart rate							0

Predictors	Responder 1	Responder 2	Responder 3	Responder 4	Responder 5	Responder 6	
Treatment with Aspirin	✓				✓	✓	3(50)
Treatment with beta blocker					✓	✓	2(33)
Treatment with ACE inhibitors'							0
Treatment with CCBs							0
Treatment with lipid lowering gents			✓	✓		✓	3(50)
Ischemic ECG changes		✓			✓		2(33)
Abbreviations: ACE, Angiotensin converting enzyme; BMI, body mass index; BP, blood pressure;							

Table 9: Final list of predictors to be included in the final multivariable clinical model (Full Model)

Variable	Type	Number of votes (%)
Age	Continuous	6 (100)
Gender	Categorical	6 (100)
Chest pain	Categorical	6 (100)
Diabetes mellitus	Categorical	6 (100)
Hypertension	Categorical	5 (83)
Hyperlipidemia	Categorical	4 (67)
Family History	Categorical	4 (67)
Smoking	Categorical	4 (67)

4.2.2.2: Final Multivariable Logistic Regression Model

After exclusion of current smoking, the regression coefficients, standard errors, p values and odds ratios with their confidence interval for the final model using the full model approach is presented in Table 11. Typical chest pain, gender and hyperlipidemia were significant predictors of obstructive coronary artery disease in a multivariable model after controlling for all other variables.

4.2.2.3: Interaction

Although interaction terms were not included in our survey, we decided to test for interaction between chest pain typicality and age and interaction between chest pain typicality and gender given the possibility of such interaction based on clinical knowledge. There was no significant interaction in both cases. Regression coefficients, standard errors and p values are shown in Appendix 7.

4.2.3: Model Performance

We assessed the performance of our developed model from two aspects. Discrimination ability and goodness of fit. For discrimination ability, we used the AUC of ROC curve (c statistics) and for goodness of fit, we used Hosmer-Lemeshow statistics.

The ROC curve plots true positive rate versus false positive rate at different thresholds (Figure 3). The AUC of ROC curve (which corresponds c statistics) for our developed model is 0.68 (0.63-0.73) which indicates a fair ability of the discriminate between those developed the outcome versus those who did not.

Hosmer-Lemeshow statistics has a chi-square statistic of 4.08, for a degree of freedom of 8 with *p* value of 0.85, which indicates an adequate fit of our developed model for the derivation dataset. Table 12 summarizes the measures of model performance.

Table 10: Initial multivariable regression model using model pre-specification (Full Model)

Parameter	Regression Coefficient	Standard Error	P value	Odds Ratio	Confidence Interval
Intercept	-5.2617	0.6542	<.0001		
Age	0.00861	0.0105	0.4120	1.009	0.988-1.030
Sex (Male)	0.8038	0.2106	0.0001	2.234	1.478-3.376
Typical Chest pain	1.0615	0.2783	0.0001	2.891	1.675-4.988
Hyperlipidemia	0.6333	0.2082	0.0023	1.884	1.253-2.833
Hypertension	0.2109	0.2012	0.2947	1.235	0.832-1.832
Family history	0.2329	0.1970	0.2372	1.262	0.858-1.857
Diabetes Mellitus	0.1636	0.2955	0.5798	1.178	0.660-2.102
Current Smoking	-0.1235	0.2881	0.6683	0.844	0.502-1.555

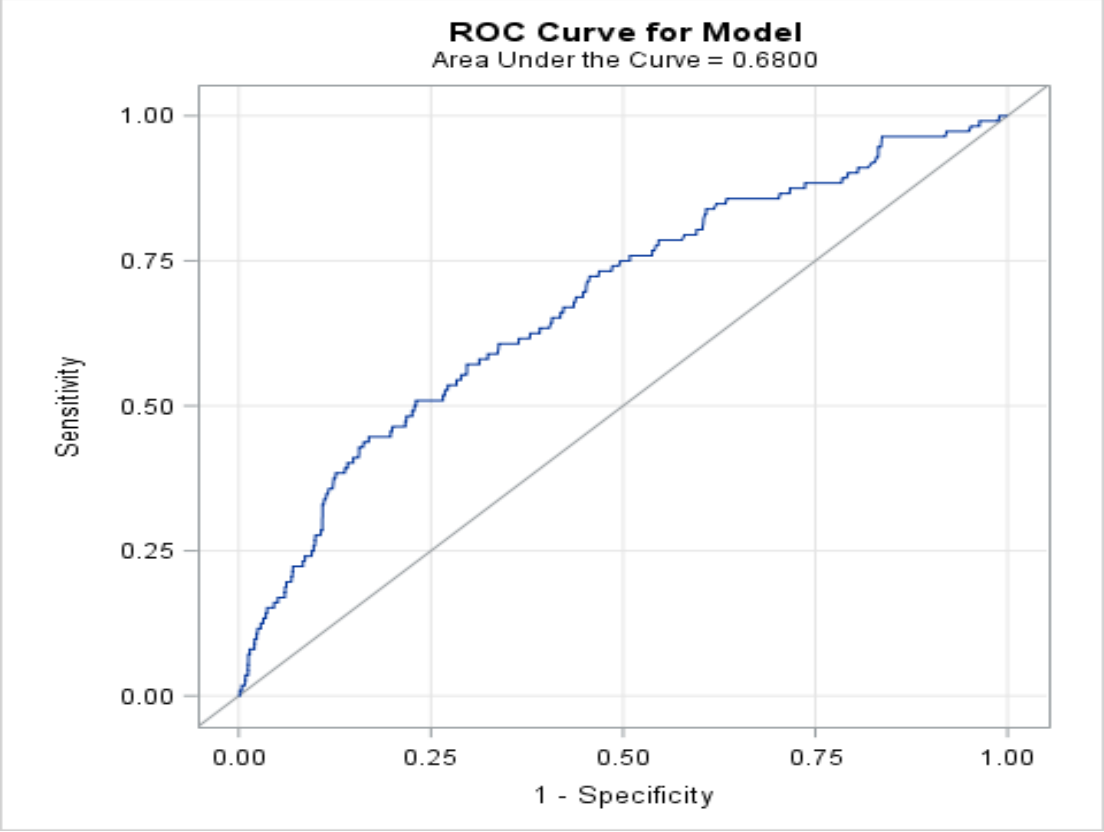
Table 11: Final multivariable clinical model using model pre-specification (Full Model)

Parameter	Regression Coefficient	Standard Error	P value	Odds Ratio	Confidence Interval (OR)
Intercept	-5.3101	0.6457	<.0001		
Age	0.00935	0.0104	0.3665	1.009	0.989 - 1.030
Sex (Male)	0.8024	0.2105	0.0001	2.231	1.477 - 3.370
Typical chest pain	1.0549	0.2779	0.0001	2.872	1.666 - 4.950
Hyperlipidemia	0.6276	0.2077	0.0025	1.873	1.247 - 2.814
Hypertension	0.2091	0.2011	0.2984	1.233	0.831 - 1.828
Family history	0.2331	0.1971	0.2370	1.262	0.858 - 1.858
DM	0.1597	0.2954	0.5889	1.173	0.657 - 2.093

Table 12: Model performance measures

Measure	Value
AUC of ROC (c statistics)	0.68
Hosmer-Lemeshow goodness of fit • Chi-square • Degrees of freedom • *p* value	 4.08 8 0.85

Figure 3:Area under the curve for ROC curve of the developed model (Full Model)



4.3: Model Presentation- Score Development

After development of the multivariable clinical prediction model, we present the model in a point score format to enable practitioner to use effectively in practice and to serve as a clinical prediction rule.

4.3.1: Scoring System Points

Each variable in the final multivariable model was assigned a point score based on the result of dividing the regression coefficient of the variable by the smallest regression coefficient of a dichotomous variable in the multivariable model which was diabetes mellitus. The point score for each variable is presented in Table 13. Each point score is obtained after rounding up or down to the nearest whole number. We observe that variable of typical chest pain had the highest score points with 5 points followed by being male variable with 4 points, then hyperlipidemia diagnosis variable with 3 points. Other risk factors for coronary artery disease, family history, diagnosis with hypertension and diagnosis with diabetes mellitus were assigned one point for each. For age variable, we dichotomized it to simplify its use in the clinical risk score as described in the previous chapter (section 3.8.1).

The total points score ranges from the lowest of zero to the highest of 20 points. The corresponding expected and observed risk of diagnosis with obstructive coronary artery disease for each level of the newly developed risk score are presented in Table 14. The lowest predicted risk of a patient with a score of zero is 0.45%, representing the score for an asymptomatic female below the age of 30 years without any risk factor for coronary artery disease. The highest predicted risk of obstructive coronary artery disease based on our model was 17.9% corresponding a risk score of 19. There were no observations in our dataset with the highest possible score of 20 points.

Table 13: Point scoring system for prediction of obstructive coronary artery disease

Variable	**Scoring Point**
Age (years)	
< 30	0
30-39	1
40-49	2
50-59	3
60-69	4
70 or more	5
Male	4
Female	0
Typical chest pain	5
Risk factors	
Hyperlipidemia	3
Hypertension	1
Family history	1
Diabetes mellitus	1

Table 14: Levels of risk score and the corresponding predicted risk of CAD

Score	Predicted risk (%)	Observed risk (%)	Number of patients
0	0.45	0	11
1	0.54	2.56	39
2	0.67	0.52	191
3	0.8	0.5	401
4	1.0	1.78	505
5	1.2	0.51	391
6	1.5	1.63	552
7	1.8	1.69	711
8	2.3	2.11	617
9	2.7	2.26	443
10	3.4	2.57	350
11	4.1	7.3	274
12	5	4	201
13	6	3	103
14	7.3	14	44
15	8.7	17.39	23
16	10.6	4.2	24
17	12.7	5	20
18	15	1	2
19	17.9	0	1

4.3.2: Score Performance

We assessed the performance of the developed score in terms of discrimination and goodness of fit as we did with model to confirm the results of performance assessment of the model and investigate for any discrepancy that could result from developing the score. For assessment of the discriminative ability of the developed score, we used the AUC of ROC curve and for goodness of fit, we used Hosmer-Lemeshow statistic.

The ROC curve plots true positive rate versus false positive rate at different thresholds (Figure 3). The AUC of ROC curve for our developed score is 0.680 (CI: 0.630-0.730) which is consistent with the results of discriminative performance of the model developed. The score shows a good ability to discriminate between those developed the outcome versus those who did not.

Hosmer-Lemeshow statistic had a chi-square statistic of 7.5, for a degree of freedom of 8 with *p* value of 0.38, which indicates an adequate fit of the developed score for the derivation dataset as shown with the base model.

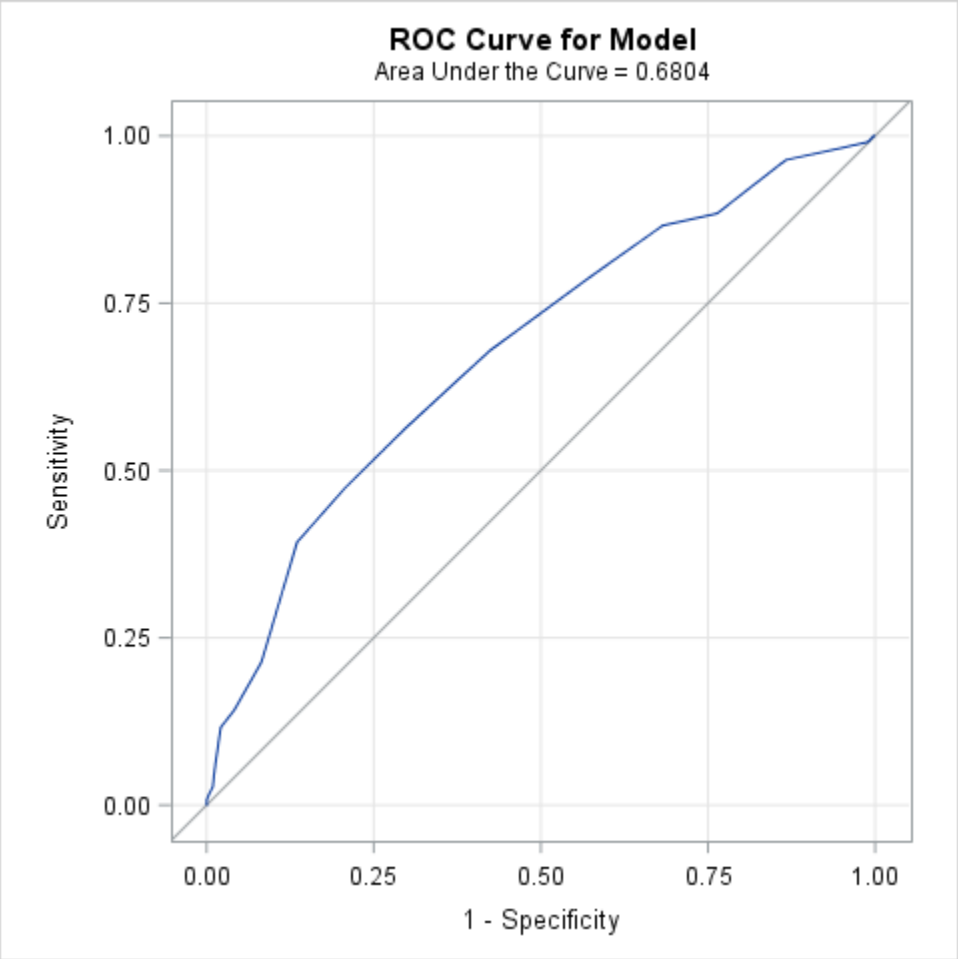
4.3.3: Operating Characteristics of Risk Score Levels

The sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (LR+) and negative likelihood ratio (LR-) of having obstructive coronary artery disease at each level of the developed risk score is presented in Table 15. The score sensitivity was better with the lower score thresholds, but the specificity was better with the higher score thresholds. All score thresholds have a had a negative predictive value ranging from 97.6% to 100% and a low positive predictive value ranging from 2.3% to 33%. Overall, lower thresholds have a low negative likelihood ratio and higher ones have a higher positive likelihood ratio.

4.3.4: Risk Classification of Score Thresholds

Based on the positive and negative likelihood ratios, we categorized risk into three groups. The low risk group were patients with score of ≤ 6. This group had a negative likelihood ratio of 0.421(CI: 0.210-0.633) and a high negative predictive value of 99% (CI:98.4%-99.5%). The high-risk group were those score ≥14 points on our score. This group had a positive likelihood ratio of 5.505(CI: 2.312-8.697) and a specificity of 98% (CI:97.4%-98.3%). The intermediate risk groups

Figure 4: AUC of ROC for the score in the derivation dataset



was those with a score from 7 to 13. Table 16 presents the risk categories and the proportion of patients with obstructive CAD in each category in derivation datasets.

4.3.4.1: Risk Re-classification of Intermediate Risk Group

We investigated the possibility of reclassifying individuals in the intermediate risk category based on the presence of typical chest pain alone. Individuals with typical chest pain in the intermediate risk group had a positive likelihood ratio of 1.268 (CI: 0.447-2.058) and a negative likelihood ratio of 0.966 (CI: 0.876-1.061) for obstructive CAD. These findings limit the use of typical chest pain to reclassify patients in the intermediate risk group.

4.3.5: Accuracy of Risk Classification

The proportion of confirmed obstructive CAD in the category with low risk (score ≤ 6) was 1.15 %. This indicates an accuracy of >98% for the classification ability of the developed score in the derivation cohort.

4.3.6: Proposed Clinical Algorithm Based on the Risk Score

Based on the risk categories of our developed risk score, we propose an algorithm for clinical use. For individual patients in the low risk category (score ≤6); testing can be terminated after obtaining calcium score of zero with 99% probability of absence of obstructive CAD (NPV 0.990 CI:0.984-0.995) and negative likelihood ratio of 0.421. (CI: 0.210-0.633). Individual patients in intermediate and high-risk categories should proceed to undergo coronary CTA (Figure 6).

4.4: External Validation of the Risk Score

For validation of our risk score derived from the full model, we used a dataset from the CONFIRM registry for cardiac CT[130]. An overview of CONFIRM registry was presented in chapter 3 (section 3.5.1).

4.4.1: Results of Validation Dataset Preparation

Figure 6 outline steps of the preparation of validation dataset. The initial dataset included a total of 27,125 patients enrolled between January 2004 and May 2010 from 12 centers around the world. We identified 8,835 patients with a calcium score of zero from 11 centers. After exclusion of patients with a known diagnosis of coronary artery disease, previous coronary bypass graft surgery or coronary artery intervention, and patients enrolled at our center (which is a participating center in the registry) to avoid any overlap, there were a total number of 8,021 patients in the preliminary dataset considered for exploratory analysis.

Table 15: Operating characteristics of each score threshold in derivation dataset

Score	Sensitivity	Specificity	PPV	NPV	PLR	NLR
0	1.00 (0.999-1.000)	0.0	0.023 (0.019-0.027)	-	1.00	--
1	1.00 (0.979-1.000)	0.002 (0.001-0.004)	0.023 (0.019-0.027)	1.000 (0.720-1.000)	1.00	--
2	0.99 (0.958-1.000)	0.010 (0.008-0.014)	0.023 (0.019-0.028)	0.980 (0.894-1.000)	1.001(0.982-1.020)	0.877(-0.953–2.699)
3	0.98 (0.948-1.000)	0.049 (0.044-0.056)	0.024 (0.019-0.028)	0.992 (0.970-1.000)	1.033(1.005-1.062)	0.358(-0.165-0.887)
4	0.96 (0.911-0.990)	0.133 (0.124-0.143)	0.025 (0.020-0.030)	0.994 (0.984-0.998)	1.112(1.068-1.157)	0.268(-0.006-0.542)
5	0.884 (0.810-0.937)	0.237(0.225-0.249)	0.026 (0.022-0.032)	0.989 (0.981-0.994)	1.158(1.073-1.242)	0.490(0.223-0.758)
6	**0.866 (0.789-0.923)**	**0.318 (0.305-0.331)**	**0.029 (0.023-0.035)**	**0.990 (0.984-0.995)**	**1.269(1.168-1.371)**	**0.421(0.210-0.633)**
7	0.786 (0.698-0.858)	0.431 (0.417-0.445)	0.031 (0.025-0.038)	0.990 (0.983-0.993)	1.381 (1.235-1.528)	0.497(0.309-0.865)
8	0.679(0.584-0.764)	0.577 (0.563-0.591)	0.036 (0.029-0.045)	0.990 (0.982-0.992)	1.605(1.380-1.829)	0.557(0.397-0.716)
9	0.563 (0.466-0.656)	0.703 (0.690-0.716)	0.042 (0.033-0.054)	0.987 (0.981-0.989)	1.895(1.555-2.235)	0.622(0.483-0.761)
10	0.473 (0.378-0.570	0.794 (0.782-0.805)	0.051(0.038-0.066)	0.985 (0.980-0.988)	2.292(1.798-2.786)	0.664(0.540-0.788)
11	0.393 (0.302-0.490)	0.865 (0.855-0.874)	0.064(0.047-0.084)	0.984 (0.980-0.987)	2.905(2.162-3.648)	0.702(0.591-0.813)
12	0.214 (0.142-0.302)	0.918(0.910-0.925)	0.057 (0.037-0.084)	0.980 (0.976-0.984)	2.606(1.591-3.620)	0.856(0.768-0.944)
13	0.143 (0.084-0.222)	0.958(0.952-0.964)	0.074 (0.043-0.117)	0.980 (0.975-0.983)	3.405(1.694-5.115)	0.894(0.8223-0.967)
14	**0.116 (0.63-0.190)**	**0.980(0.974-0.983)**	**0.114 (0.062-0.187)**	**0.979(0.975-0.983)**	**5.505(2.312-8.697)**	**0.903(0.839-0.967)**
15	0.063 (0.026-0.124)	0.987(0.983-0.990)	0.100 (0.041-0.195)	0.978 (0.973-0.982)	4.753(0.930-8.576)	0.950(0.902-0.998)

NLR: Negative likelihood ratio, NPV: Negative predictive value, PLR: Positive likelihood ratio, PPV: Positive predictive value

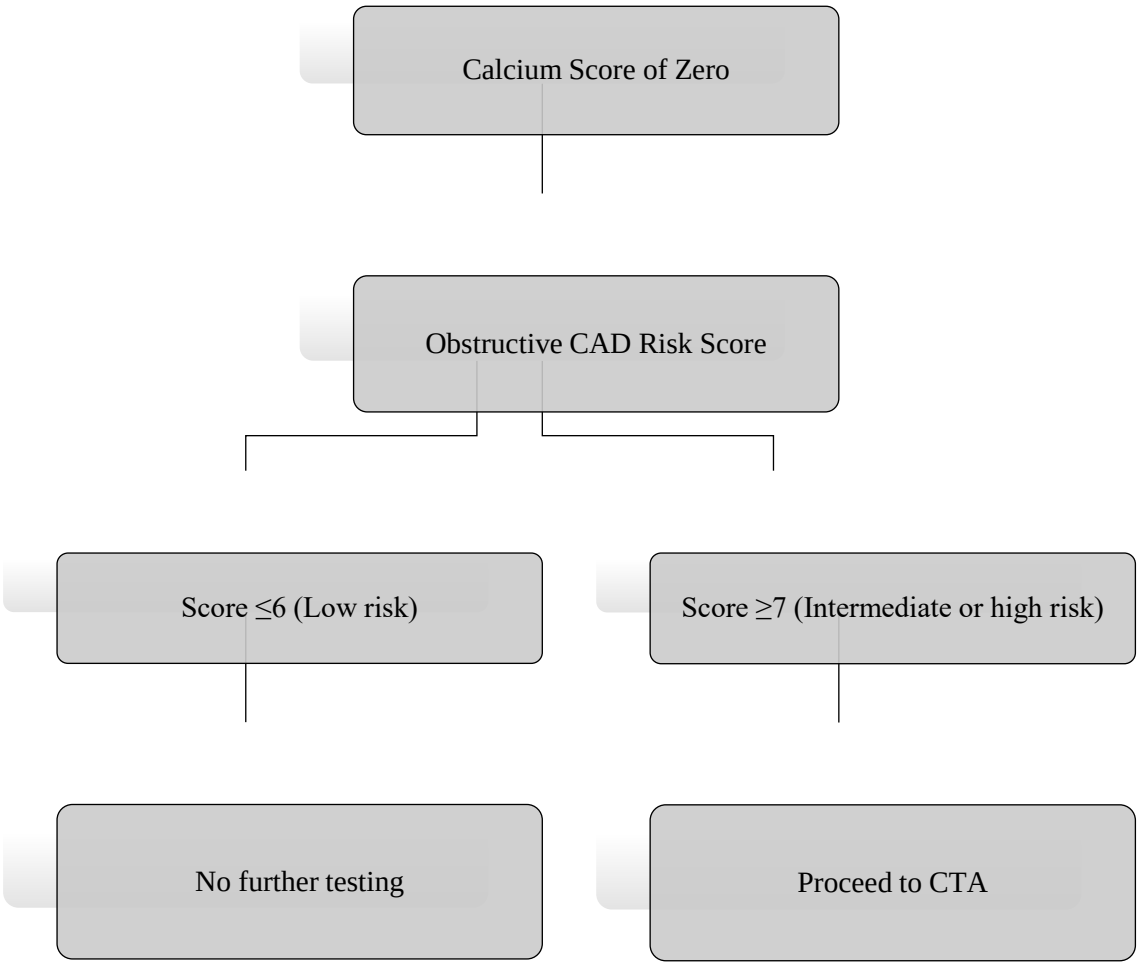
Score	Sensitivity	Specificity	PPV	NPV	PLR	NLR
16	0.027 (0.006-0.076)	0.991 (0.988-0.993)	0.064 (0.013-0.175)	0.978(0.973-0.982)	2.917(0.656-6.489)	0.982(0.950-1.014)
17	0.018(0.002-0.063)	0.996 (0.993-0.997)	0.087(0.012-0.280)	0.978(0.972-0.981)	4.074(-2.143-10.291)	0.986(0.960-1.0127)
18	0.009(0.002-0.049)	1.000 (0.998-1.000)	0.333 (0.008-0.906)	0.977(0.973-0.981)	21.38(-32.921-75.6980	0.992(0.973-1.010)
19	0.0	1.000(0.998-1.000)	-	0.977(0.973-0.981)	0.00	1.009(0.810-1.208)
20	--	--	--	--	--	--

NLR: Negative likelihood ratio, NPV: Negative predictive value, PLR: Positive likelihood ratio, PPV: Positive predictive value

Table 16: Proportion of patients in each risk category and the proportion of confirmed obstructive CAD

Risk Category	Derivation Cohort	
	Total Patients N (%)	**Confirmed Obstructive CAD N (%)**
Low risk ≤ 6	2090 (43)	24 (1.15)
Intermediate risk 7-13	2699 (55)	75 (2.80)
High risk ≥ 14	114 (2)	13 (11.40)

Figure 5: Proposed algorithm for clinical use of the developed risk score



The final dataset included 4, 290 patients after exclusion of one center with large missing proportions of data on symptoms (chest pain) and outcome (see next sections for details).

4.4.2: Exploratory Analysis of Preliminary Validation Dataset

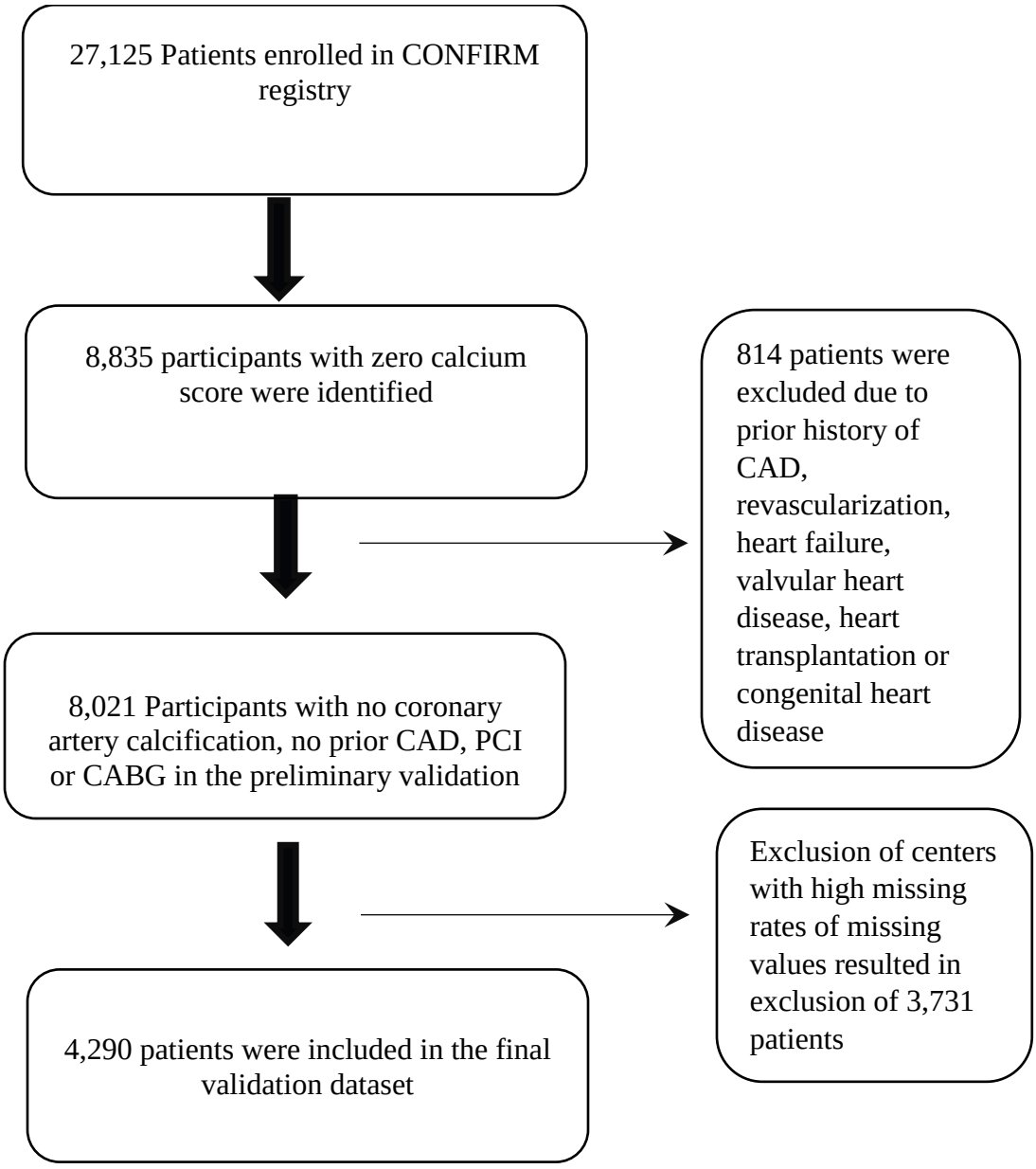
The preliminary dataset obtained by identification of patients with calcium score of zero, non-overlapping with patients in our derivation dataset and have no previous history of coronary artery disease, coronary artery bypass graft surgery, angioplasty or stents underwent an exploratory analysis. This group included a total of 8,021 patients from CONFIRM registry. The key characteristics of patients included in the preliminary validation dataset is presented in Table 17. Among patients included in the preliminary dataset; 4,095 were females (51%) with an overall age mean of 53 (SD :12). More than half of the patients have chest pain as a reason for the study. Diagnosis with hyperlipidemia was the most prevalent known conventional risk factor for coronary artery disease as it was reported in 3,573 patients (45%). Diagnosis with hypertension was next and reported in 3,071 patients (38.3%).

Medications profile of patients included in this dataset showed that 1,542 patients (19%) were taking aspirin at the time of the scan and 1,326 patients (17%) were taking lipid lowering agents. Medications with antianginal effect like beta blocker, calcium channel blockers and nitrates were reported by 1,153 patients (14.4%), 765 patients (9.5%) and 533 patients (7%) respectively.

More than half of the patients reported having chest pain. Typical chest pain that is triggered by exertion or emotional stress and relieved by rest or nitroglycerine was reported by 637 patients (8%). Atypical or non-cardiac chest pain was reported by 3493 patients (43.5%). The mean of pre-test probability of coronary artery disease as the cause of the symptoms was 32.5%. Data on parameters of physical assessment prior to the test was limited to body weight and height which was not unified in measurement units collected. There was no data on blood pressure, heart rate or ECG findings in this dataset.

The outcome of obstructive coronary artery disease, which is defined as luminal stenosis of 50% or more on coronary CTA as per the guidelines of SCCT[123] , was found in 419 patients (5%) .

Figure 6: Flow chart of preparation of final validation dataset



*Abbreviations: CAD, coronary artery disease; CABG, coronary artery bypass graft;

In contrast to the derivation dataset, the preliminary validation dataset had a significant proportion of missing values of the predictors and the outcome. Extent of the missing values, their pattern and how were they treated are discussed in the next section.

4.4.3: Predictors and Outcome in Preliminary Validation Dataset

Table 18 provides summary of the frequency, distribution and proportion of missing values for each of the predictors included in the final multivariable model (full model approach) and the outcome. We observe that data on all predictors included in the final multivariable model are available, however; unlike in the derivation dataset, the proportion of the missing values was high for the typical chest pain variable (predictor) and the outcome variable (21% & 10% respectively).

4.4.3.1: Missing Data in Preliminary Validation Cohort

Investigating the pattern of missing values in the preliminary validation dataset revealed that there are 6,357 patients with at least one missing value for at least a predictor representing 79% of all preliminary validation dataset.

We assessed the completeness of the predictors and the outcome as stratified by the site (Table 19) and found that there are two centers with more than 20% missing values of the variable of typical chest pain (predictor) and the outcome. These two centers are center number 6 and center number 8. The missing proportions in the other centers did not exceed 20%. We decided to exclude centers with > 20 % missing proportion in any of the model predictors or the outcome.

4.4.4: Final Validation Dataset

After we excluded the two centers with more than 20% missing values in any of our model predictors or the outcome, the final data set included 4,290 patients from eight centers in CONFIRM registry (Appendix 6). Characteristics of the final validation dataset is presented in Table 20. The mean age of patients in the final validation dataset is 52 years, with 49 % males (n=2.080).

Table 17: Characteristics of patients in preliminary validation dataset

Characteristics	**N (%)** **Mean (SD)**
Age	53(12.00)
Male	3,914 (48.80)
Female	4,095 (51.05)
Reason for Study	
Chest pain	4,051 (50.50)
Shortness of breath	642 (8.00)
Asymptomatic	500 (6.23)
Other reasons	301(4.00)
Risk Factors	
Family History	2,109(26.29)
Hyperlipidemia	3,573(44.50)
Hypertension	3,071 (38.28)
Diabetes Mellitus	721 (8.98)
Current Smoking	1,259 (15.70)
Medications	
Aspirin	1,542(19.22)
Beta Blockers	1,153(14.37)
Ca channel blockers	765 (9.53)
ACE inhibitors	568(7.08)
Lipid lowering agents	1,326(16.53)
Symptoms	
Typical chest pain	637(7.94)
Atypical/Non-cardiac chest pain	3,493(43.55)
No chest-pain	2,142(26.70)
Shortness of breath	1,080 (13.46)

Characteristics	**N (%)** **Mean (SD)**
Pre-test probability of CAD	32.5 % (27)
Body mass index (BMI)	26 (5.00)
Creatinine	81 (27.00)
Obstructive coronary artery disease	419 (5.00)
Total	8,021
ACE inh: Angiotensin converting enzyme inhibitors Ca channel blockers: Calcium channel blockers	

Table 18: Distribution and missing proportions of model predictors and outcome in preliminary validation dataset

Predictor	**Scale**	**Range**	**Missing percentage**
Age	Continuous	18-105	0.12%
Gender	Categorical	Male, Female	0.2%
Typical Chest pain	Dichotomous	Yes, No	21%
Family history	Categorical	Yes, No	2%
Hypertension	Categorical	Yes, No	1.3
Hyperlipidemia	Categorical	Yes, No	1.2%
Diabetes mellitus	Categorical	Yes, No	1.1%
Obstructive CAD	Categorical	Yes, No	10%

Work up of chest pain was the indication for coronary CTA in 47% (n=1,997). Diagnosis with hyperlipidemia was the most prevalent risk factor for CAD (41%) followed by hypertension (38%) and family history (26%). Typical chest pain was reported in 10% of the final validation dataset (n=430). The outcome of obstructive CAD was found in 5% of the patients in the final validation dataset (n=205).

4.4.4.1: Final Validation Dataset vs Data from Excluded Centers

When the final validation dataset and the patients in the excluded centers were compared (Table 21), the two cohorts were significantly different on aspects of the indication for the test, chest pain typicality on presentation, presence of shortness of breath on presentation, family history, hyperlipidemia, diabetes mellitus, the pharmacological treatments and the rate of outcome of obstructive CAD. The patients in the excluded centers had a higher risk profile as reflected by the significantly higher prevalence of risk factors for obstructive CAD and higher prevalence of the outcome of obstructive CAD.

4.4.5: Score Validation

The developed score from the full model was validated using the final validation dataset. We first investigated for the relatedness between the case mix and the outcome of the derivation and validation datasets by comparing the characteristics of both datasets, then we performed multiple imputations of missing values in the final validation dataset, assessed the discriminative ability and the goodness of fit of the score in the final validation dataset and finally compared the diagnostic characteristics of risk thresholds in both derivation and validation datasets.

4.4.5.1: Relatedness of Derivation and Validation Datasets

Table 22 shows the results of the relatedness of case mix and outcome of both the derivation and validation datasets. We compared the proportions of the categorical variables and means with standard deviations of continuous variables for both datasets.

Table 19: Missing proportions of our model predictors and outcome as stratified by the center in the preliminary validation dataset

Center	Number Of patients	Age	Sex	Chest Pain	Family History	Hyperlipidemia	Hypertension	Diabetes Mellitus	Obstructive-CAD
1	19	0	0	0	0	0	0	0	0
3	525	0	0	0.19%	0	0	0	0	0
4	613	0	0	0	0	0	0	0	0
6	606	0	1.16%	100%	1.16%	1.16%	1.16%	1.16%	0
7	121	0	0	20%	0	0	0	0	0
8	3125	0	0	31%		0.4%	1.25%	1.63%	25%
9	333	0	0	0	1.8%	0.6%	0	0	0
10	1561	0	0.32	10%	8.46%	4%	3.2%	1.22%	0.06%
11	741	0	0	0	0.54%	0.54%	0.54%	0.54%	0
12	377	0	0	0	1.3%	1.06%	1.06%	1.33%	0.27%

- Center 2 did not have patients with zero calcium score
- Center 5 was excluded to avoid overlap with the derivation cohort

The case mix of the two datasets shows significant differences between the two datasets, the derivation and validation; in term of age, percentage of males, indication, family history and the outcome of obstructive CAD. The mean age of patients in derivation dataset was 53 years vs 51 years in the validation dataset. The proportion of males in the derivation dataset was 42% vs 49% in the validation dataset. The family history of CAD was present in 46% of patients in the derivation cohort vs 38% in the validation cohort. The outcome of obstructive CAD was documented in 5% of patients in the validation cohort, while only 2% of the derivation cohort were found to have obstructive CAD.

Table 20: Characteristics of the final validation dataset

Characteristics	**N (%)** **Mean (SD)**	**Missing values percentage**
Age	52(12)	0.12%
Gender		0.15%
Male	2,080 (49)	
Female	2,205(51)	
Reason for Study		25%
Chest pain	1,997 (47)	
Shortness of breath	401 (9)	
Asymptomatic	500 (12)	
Other reasons	301 (7)	
Risk Factors		
Family History	1,581(37)	3.0 %
Hyperlipidemia	1,773(41)	2.0 %
Hypertension	1,648 (38)	1.0 %
Diabetes Mellitus	342 (8)	1.0 %
Current Smoking	695 (16)	1.0 %
Peripheral arterial disease	86(2)	34.0 %
Medications		
Aspirin	730(17)	38.0%
Beta Blockers	682(16)	38.0%
Calcium channel blockers	105(2)	55.0%
Nitrates	67 (2)	43.0 %
ACE inhibitors	417(10)	38.0%
Lipid lowering agents	437(11)	46.0 %

Characteristics	**N (%)** **Mean (SD)**	**Missing values percentage**
Symptoms		
Typical chest pain	430 (10)	4.0%
Atypical/Non-cardiac	2,248 (52)	
Shortness of breath	839 (20)	15.0%
Pre-test probability of CAD	31% (26)	4.0%
Body mass index (BMI)	28 (5)	3.0%
Creatinine	77 (23)	33.0%
Obstructive coronary artery disease		0.05%
Yes	205 (5)	
No	4,083(95)	
Total	4,290	
ACE inh: Angiotensin converting enzyme inhibitors Ca channel blockers: Calcium channel blockers		

In summary, the patients in validation cohort were younger and had higher prevalence of the outcome of obstructive CAD. There are no significant differences between the two cohorts in the risk factors except family history.

4.4.5.2: Imputation of Missing Values in the Final Validation Dataset

After exclusion of centers with large proportion of missing values on any of the predictors or the outcome of the developed model, there is residual missing values of some of the variables in the final validation dataset as presented in Table 20. Prior to score validation, we investigated for the missing pattern of the residual missing values and was completely at random (MCAR). We performed multiple imputations using chained equations to produce 10 imputed datasets. Trace plots of imputed datasets showed that there are no specific trends in means of missing values which indicate randomness of means between different iterations and good convergence (Appendix 8).

Table 21: Characteristics of patients in final validation dataset vs excluded centers

Characteristic	**Final Validation Dataset (N =4,290)** ***n* (%), mean (SD)**	**Excluded Centers Patients (N=3,731)** ***n* (%), mean (SD)**	**P value**
Age	51(12)	54(11)	0.0001
Male	2,080(49)	1,834(49)	0.56
BMI	28 (5)	25 (6)	0.0001
Indication			0.04
Chest pain	1,997 (47)	2,054(55)	
Shortness of breath	401 (9)	241(6)	
Others	801 (19)	338(18)	
Chest pain			
- Typical	430(10)	207(6)	0.07
- Atypical/Non-cardiac	2,248(52)	1,245 (33)	
- None	1,435 (33)	707 (19)	
Shortness of breath	839 (20)	241 (6)	0.0001
Hyperlipidemia	1,773 (41)	1,800(48)	0.0001
Hypertension	1,648 (38)	1,423 (38)	0.85
Family history	1,581 (37)	528(14)	0.0001
Diabetes mellitus	342 (8)	379 (10)	0.0001
Current Smoking	695(16)	564 (15)	0.16
Treatment with Aspirin	730(17)	812 (22)	0.0001
Treatment with beta blockers	682(16)	471(13)	0.0001

Characteristic	**Final Validation Dataset (N =4,290)** *n* **(%), mean (SD)**	**Excluded Centers Patients (N=3,731)** *n* **(%), mean (SD)**	**P value**
Treatment with Ca channel blockers	106(2)	659(18)	0.0001
Treatment with ACE inhibitors	417(10)	151 (4)	0.0001
Treatment with lipid lowering agents	473 (11)	568(15)	0.0001
Pre-test probability of CAD	31% (26)	36(27)	0.0001
Obstructive CAD	205 (5)	214(6)	0.0001

4.4.5.3: Score Performance in the Final Validation Dataset

We evaluated the performance of the developed score in the final validation dataset in by evaluating discriminative ability and the goodness of fit in each of the imputed data set. We used the area under the curve (AUC) of the ROC curve for evaluation of the discriminative performance and Hosmer-Lemeshow statistic for evaluation of the goodness of fit.

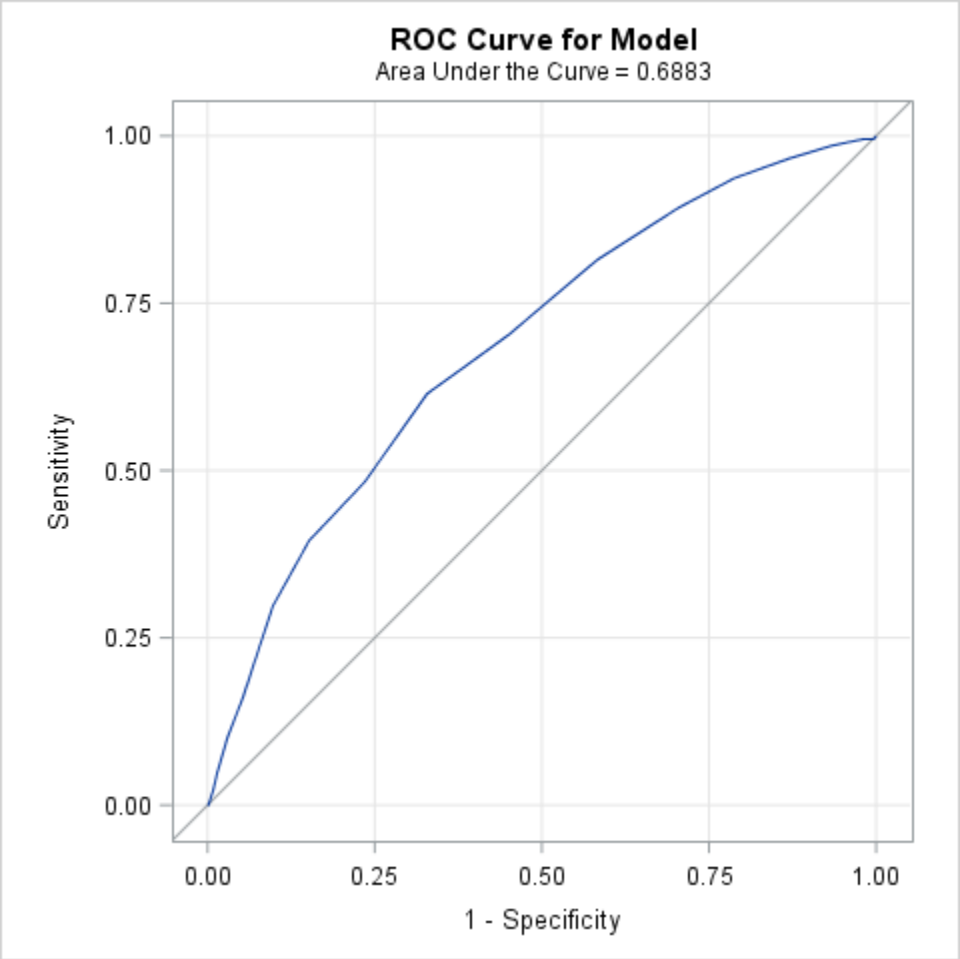
The ROC curve was plotted for the discriminative model performance in the first imputed validation dataset (Figure 5). The AUC of the ROC of the developed score in the first imputed dataset was 0.688 (CI 0.651-0.725).

The goodness of fit of the developed score as assessed in each of the imputed validation datasets by Hosmer-Lemeshow statistics indicated that the score fit adequately to the validation dataset (*p* value >0.05). Table 23 shows the AUC of ROC curve and the Hosmer-Lemeshow statistic for each imputed dataset.

Table 22: Characteristics of case mix and outcome of derivation and validation datasets

Characteristic	**Derivation Cohort (N=4903)** ***n*** **(%)** **mean (SD)**	**Validation Cohort (N =4,290)** ***n*** **(%)** **mean (SD)**	***p*** **value**
Age	53(10.3)	51(12)	<0.0001
Male	2,058 (42)	2,080 (49)	0.0027
Indication			<0.0001
Chest pain	2793 (57)	2,935 (48)	
Shortness of breath	370(8)	540 (9)	
Others	1740 (35)	463 (8)	
Chest pain			
- Typical	392(8)	430(10)	0.04
- Atypical/Non-cardiac	2679(55)	2,298(55)	
- None	1832 (37)	1,435 (35)	
Hyperlipidemia	2,037(42)	1,773 (42)	0.43
Hypertension	1,892 (39)	1,648 (39)	0.44
Family history	2,250 (46)	1,581 (38)	<0.0001
Diabetes mellitus	466 (9)	342 (8)	0.07
Current Smoking	665(14)	695(16)	0.02
Obstructive CAD	112 (2)	205 (5)	0.03

Figure 7: AUC of ROC curve of the developed score in the first imputed validation dataset



4.4.5.3: Diagnostic Characteristics of Risk Thresholds in Validation Dataset

We used the first imputed dataset to calculate the sensitivity, specificity, positive and negative predictive values, and the likelihood ratios (positive and negative) for the thresholds used of low and high-risk categories, score of 6 and 14 respectively. At score of ≤6 in the validation cohort, the negative likelihood ratio is 0.207(CI:0.087-0.327). The positive likelihood ratio of a score of ≥14 is 5.478(CI: -0.140 - 11.096) (Table 24).

Table 23: AUC of ROC and Hosmer-Lemeshow statistic for each imputed validation dataset

Imputation	Discriminative assessment AUC of ROC (Confidence Interval)	Goodness of Fit (Hosmer-Lemeshow) Chi-Square (*p* value)
1	0.688 (0.651-0.725)	6.48 (0.59)
2	0.686 (0.650-0.723)	6.57 (0.58)
3	0.690 (0.652-0.726)	8.02 (0.43)
4	0.689 (0.653-0.726)	8.26 (0.41)
5	0.689 (0.652-0.725)	7.60 (0.47)
6	0.687 (0.650-0.724)	7.01 (0.54)
7	0.689 (0.653-0.726)	6.97 (0.54)
8	0.689 (0.652-0.725)	6.27 (0.62)
9	0.689 (0.652-0.726)	6.87 (0.55)
10	0.688 (0.651-0.724)	6.63 (0.57)

Table 24: Operating characteristics of the score at low and high-risk thresholds in validation dataset

Score Threshold	Sensitivity	Specificity	PPV	NPV	PLR	NLR
6	0.930	0.362	0.117	0.982	1.450	0.207
14	0.037	0.993	0.333	0.920	5.478	0.969

4.4.5.4: Score Performance in the Derivation vs Validation Datasets

When we compared the performance of the developed score in both datasets, the derivation and the validation datasets; using the same measures of performance assessment that we used for assessment of the score in each dataset (AUC of ROC curve for assessment of discriminative performance and Hosmer-Lemeshow statistic for assessment of goodness of fit), the score performance in the validation dataset was similar to that in the derivation dataset (Table 25).

The classification accuracy of the developed score in both derivation and validation datasets is presented in Table 26. It seems that the score performs better in the low risk category. It can be used to rule out obstructive CAD with high accuracy (> 98.5% in derivation dataset and >97.5% in the validation dataset).

4.4.5.6: Score Calibration in Derivation vs Validation Datasets

We assessed the score calibration between observed and predicted risk for each level in both derivation and validation datasets. At low and intermediate score levels, the score had a good calibration in both datasets, but decreased at high score levels due to small number of patients with high score (Figure 8).

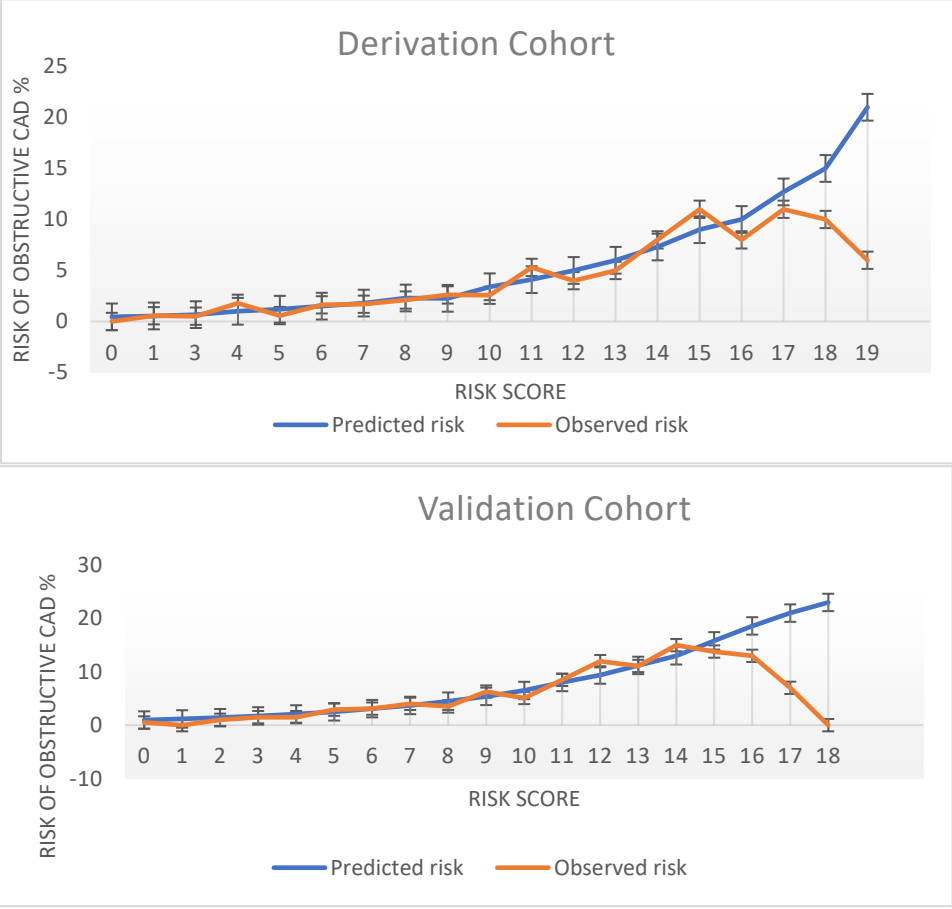
Table 25: Score performance in derivation and validation datasets

Performance Measure	Derivation Dataset	Validation Dataset
AUC of ROC curve	0.680 (0.630-0.730)	0.688 (CI:0.651-0.725)
Hosmer-Lemeshow	7.50 (0.38)	6.48 (0.59)

Table 26: Classification accuracy of the developed score in derivation and validation datasets

Risk Category	Derivation Cohort		Validation Cohort	
	Total Patients N (%)	Confirmed Obstructive CAD N (%)	Total patients N (%)	Confirmed Obstructive CAD N (%)
Low risk ≤ 6	2090 (43)	24 (1.15)	1,743 (41)	38 (2.18)
Intermediate risk 7-13	2699 (55)	75 (2.80)	2,403 (56)	140 (6)
High risk ≥ 14	114 (2)	13 (11.40)	144 (3)	21 (15)

Figure 8: Calibration plot of expected vs observed risk of the score in derivation and validation datasets



4.5: Results of Reduced Model Development (Univariate Screening Approach)

In this section we present the results of the model developed using the approach of univariate screening of all candidate predictors to select the predictors for inclusion in the multivariable model. This includes results of univariate screening through examining the association of each predictor with the outcome of obstructive CAD, multivariable model of predictors selected based on the univariate screening, the final multivariable model and model performance.

4.5.1: Univariate Screening

Univariate screening of all candidate predictors to select predictors be included in the multivariable model was based on investigation of the association of each candidate predictor with the outcome. Significant predictors were included in an initial multivariable model. To examine for association between each candidate predictor and the outcome, we used chi square test or Fisher exact test where appropriate for categorical variables and t-test for continuous variables.

As shown in Table 27, variables with significant association with the outcome of obstructive CAD in the univariate screening, as defined by maximum p-value of 0.1; were male sex, typical chest pain, diagnosis with hyperlipidemia, diagnosis with hypertension, treatment with aspirin, treatment with lipid lowering agents, probability of coronary artery disease, creatinine and baseline systolic blood pressure.

Table 28 shows the final list of significant predictors based on the results of univariate screening, their type, range of values and proportion of missing values for each variable. Unlike variables included in our full model which was developed through pre-specification of model predictors, some of the variables with statistical significance in the univariate screening have missing values. The proportion of missing values among these variables was noticeably small.

Table 27: Candidate predictors and association with outcome

Predictor	**No Obstructive CAD** **N (%)** **Mean/SD**	**Obstructive CAD** **N (%)** **Mean/SD**	**P- value**
Age	53(10.3)	53 (10)	0.6
Male	1993(42)	65(58)	0.0006
Family history	2194(46)	56(50)	0.39
Hypertension	1838(38)	54(48)	0.04
Hyperlipidemia	1970(41)	67(60)	0.0001
Current Smoking	650(14)	15(13)	0.9
Diabetes mellitus	432(9.02)	15(13)	0.132
Chest pain Typical Atypical/Non-cardiac	 370(8) 2628(55)	 22(20) 51(46)	 0.0002
Shortness of breath	2935(61)	61(54)	0.25
Palpitations	2766(58)	60(54)	0.49
Treatment with Aspirin	2010 (42)	63(56)	0.0035
Treatment with Beta blocker	1601(33)	47(42)	0.068
Treatment with Lipid lowering agents	1408(29)	49(44)	0.0016
Treatment with Ca channel blockers	472(10)	6(5)	0.14

Predictor	No Obstructive CAD N (%) Mean/SD	Obstructive CAD N (%) Mean/SD	P- value
Treatment with ACE inh	725 (15)	21(19)	0.28
BMI	29(6)	30(5)	0.50
Systolic BP (mmHg)	131(18)	134(21)	0.02
Diastolic BP	78(10)	79(9)	0.18
Heart Rate	68(12)	67(13)	0.25
Creatinine	76.5(19)	80(19)	0.08
Ischemic ECG changes	1672 (35)	33(29)	0.27
Probability of CAD	0.21(0.36)	0.32 (0.32)	0.003
ACE inh: Angiotensin converting enzyme inhibitors Ca channel blockers: Calcium channel blockers			

4.5.2: Multivariable Modelling (Reduced model)

After defining the final list of predictors to be included in the multivariable model based on univariate screening (reduced model), we conducted the multivariable modeling starting with multiple imputations of the missing values followed by step-wise selection of variables and finally fitting the final multivariable model.

4.5.2.1: Multiple Imputation

The missing values pertaining to the final list of predictors based on the univariate screening were completely at random (MCAR) based on the missing pattern we performed. We performed multiple imputations using chained equations (PROC MI in the statistical software package SAS 9.4© (SAS institute, Cary, NC, USA.)) to produce 10 imputed datasets. Trace plots of imputed datasets show that there are no specific trends in means of missing values which indicate randomness of means between different iterations and good convergence (Appendix 9).

Table 28: Final list of predictors to be included in a multivariable model based on univariate screening

Variable	Type	Range (Min, Max)	Missing%
Gender	Categorical	Male Female	0
Chest pain typicality	Categorical	Typical Atypical /Noncardiac No chest pain	0
Diagnosis with hyperlipidemia	Categorical	Yes No	0
Diagnosis with Hypertension	Categorical	Yes No	0
Treatment with aspirin	Categorical	Yes No	0
Treatment with lipid lowering agents	Categorical	Yes No	0
Probability of CAD	Continuous	0.3-100	0.1
Baseline systolic blood pressure	Continuous	78-227	0.2
Creatinine	Continuous	25-557	2.04

4.5.2.2: Variables Selection (Reduced-model approach)

For each imputed dataset, forward, backward and step-wise selection of variables in a multivariable logistic regression model to predict the presence of obstructive CAD is presented in Table 29. There was a consistency in the selection of the same set of variables between the different automated selection method across all imputed dataset. These variables were male sex, typical chest pain, diagnosis of hyperlipidemia and baseline systolic blood pressure. Other variables with statistical significance in the univariate screening, such as diagnosis of hypertension, treatment with aspirin, treatment with beta blockers, treatment with lipid lowering agents, probability of CAD and level of creatinine did not meet the statistical significance in the automated selection of variables and hence were not included in the final multivariable model for each of the imputed dataset.

Variable selection based on complete case analysis revealed selection of the same set of variables as for each imputed dataset.

4.5.2.3: Final Multivariable Prediction Model (Reduced model)

Table 30 shows the regression coefficients, standard errors, odds ratios and their confidence intervals of the final multiple regression model to predict presence of obstructive CAD using univariate screening. The regression coefficients represent the pooled estimates of regression coefficients of all imputation datasets. Typical chest pain variable was associated with the largest odds ratio of 3.036 (CI 1.76-5.22), while male sex and hyperlipidemia had odds ratio of 2 for each of them (CI 1.39-3.00). Baseline systolic blood pressure was significantly associated with the presence of obstructive CAD. We tested for interaction between gender and chest pain typicality in the multivariable model and there was no interaction.

4.5.3: Model Performance

Model performance was assessed in term of discrimination and goodness of fit. As discussed previously, ROC curve plots true positive rate versus false positive rate at different thresholds. The curve performs better as it curves toward the left upper corner of the plot.

Table 29: Variable selection for each imputed dataset by selection mode

Variables by forward selection	Variables by backward selection	Variables by step wise selection
1. Sex (Male) 2. Typical chest pain 3. Hyperlipidemia 4. Baseline systolic blood pressure	1. Sex (Male) 2. Typical chest pain 3. Hyperlipidemia 4. Baseline systolic blood pressure	1. Sex (Male) 2. Typical chest pain 3. Hyperlipidemia 4. Baseline systolic blood pressure

Table 30: Model parameters of the final multivariable model by univariate screening

Parameter	Regression Coefficient	Standard Error	P value	Odds Ratio	Confidence Interval (OR)
Intercept	-6.08	0.75	<0.0001	--	--
Male	0.72	0.20	0.0003	2.05	1.39 – 3.01
Typical chest pain	1.11	0.28	<0.0001	3.036	1.76-5.22
Hyperlipidemia	0.71	0.20	0.0003	2.04	1.39-3.00
Baseline systolic blood pressure	0.011	0.005	0.04	1.01	1.008-1.014

For our developed reduced model, the curve of the first imputed dataset shows an acceptable discriminative ability with AUC (c statistics) of 0.685 (Figure 7).

The Hosmer-Lemeshow statistic was used for assessment of the goodness of fit of the model in each of the imputation dataset. Hosmer-Lemeshow statistics in the first imputed dataset has a chi-Square of 6.82 and *p* value 0.56. This indicates an adequate fit of the model in the first imputed dataset. Table 31 presents a summary of AUC of ROC curve and Hosmer-Lemeshow statistic for each of the derivation imputed dataset.

4.6: Internal Validation of Reduced Model

We validated our reduced model (developed through univariate screening) internally since some of the variables in the final reduced model are not available in the external validation dataset used for validation of our developed full model.

4.6.1: Results of Variable Selection in Bootstraps

For each of the 500 samples obtained by bootstrapping technique, we fitted the multivariable logistic regression model with step-wise selection to evaluate for the consistency of predictors selected in each strap.

Table 32 presents the frequency of selection of each predictor by bootstraps. Predictors included in the final multivariable model derived by univariate screening were selected consistently in more than 50% of the bootstraps, with variable frequency. Variables of male sex and typical chest pain were selected in more than 90% of the bootstraps. The variable of diagnosis with hyperlipidemia was selected in 78% of the bootstraps, and baseline systolic blood pressure was selected only in 50% of the bootstraps. This indicates an underlying inherent instability of the automated selection procedure used for variable selection in modelling using univariate screening approach.

4.6.2: Optimism and Overfitting

The discriminative ability of the final multivariable model in the bootstrap samples was assessed using the c statistics. The average c statistic of the model in the bootstrap samples is 0.691 (CI: 0.689-0.693). The optimism in the c statistic of the developed model was calculated and found to be -0.006. This indicates that the developed model was not overfitted for the derivation dataset.

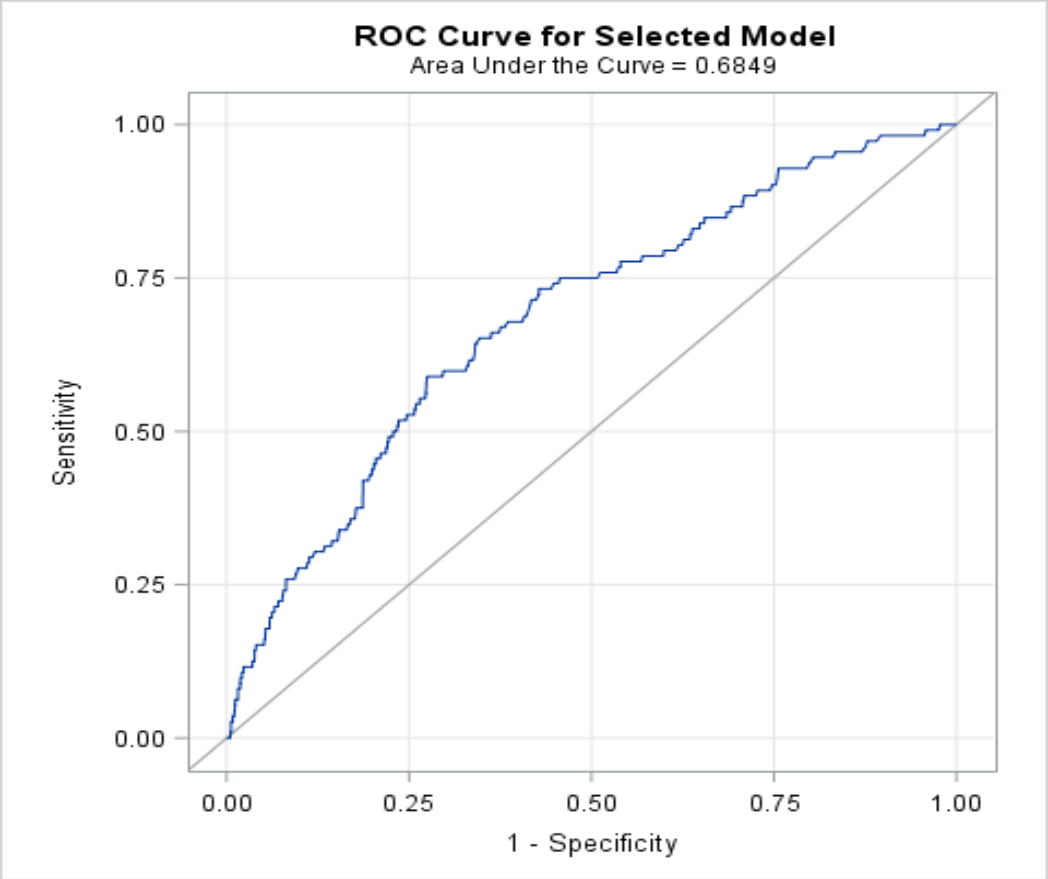
Table 31: AUC of ROC curve and Hosmer-Lemeshow statistic for the reduced model in each imputed derivation dataset

Imputation	AUC of ROC curve (c statistics)	Hosmer-Lemeshow statistic (Chi-Square (p value))
1	0.685	6.80 (0.56)
2	0.685	6.15 (0.63)
3	0.685	6.80 (0.56)
4	0.685	6.80 (0.56)
5	0.685	6.82 (0.56)
6	0.685	6.84 (0.55)
7	0.685	6.88 (0.55)
8	0.685	6.82 (0.56)
9	0.685	6.83 (0.55)
10	0.685	6.84 (0.55)

4.6.3: Pooled Estimates of Final Model in Bootstraps

We pooled the regression coefficients estimates of the final reduced model when fitted in the 500 bootstraps samples. Table 33 presents the pooled parameters estimates of the final model derived using the univariate approach when performed in 500 bootstrap samples. These pooled estimates were similar overall to those in the derived model in term of direction and significance, except for baseline systolic blood pressure which was not significant when the corresponding regression coefficients were pooled. This finding could be explained by the lower frequency at which this variable was selected by the stepwise approach in the bootstrap samples compared to the consistency at which the other variables in the model were selected at. Variables of typical chest pain, male sex and diagnosis of hyperlipidemia have pooled estimates of multivariable model from 500 bootstraps that are significantly associated with obstructive CAD.

Figure 9: ROC of the reduced model in the first imputed derivation dataset



4.7: Model Presentation (Reduced Model)

We derived a point score system from the final multivariable clinical prediction model developed using the univariate screening approach (reduced model), for an efficiency of use in clinical practice. The risk score was developed then the diagnostic characteristics for each threshold of the risk score was calculated to define an acceptable risk classification threshold of individual patients.

4.7.1: Risk Score Development

The risk score points were assigned based on the regression coefficients of the final multivariable following the same method used for deriving the risk score of the full model. Table 34 presents the risk score points assigned to each predictor and the total points. The score scale

range from 0 to 5. We dichotomized the baseline systolic blood pressure to high and normal blood pressure based on the definition of hypertension as per the definition by the 8th Joint National Commission (JNC8) 2014 report [134]. High baseline systolic blood pressure is defined as ≥ 140 mmHg.

Discriminative ability of the risk score was assessed in the derivation cohort. The AUC of ROC curve of the risk score is 0.68 (Figure 8), consistent with that of the developed model (reduced model approach).

Table 35 presents the risk score thresholds with their observed and predicted risk and the number of patients with each threshold.

4.7.2: Operating Characteristics of Score Thresholds

Sensitivity, specificity, positive prediction value (PPV), negative prediction value (NPV), positive likelihood ratio (PLR) and negative likelihood ratio (NLR) were calculated to each score threshold (Table 36).

Unlike the risk score developed from the clinical model derived using the model pre-specification method (the full model), the scale of our current risk score has limited number of thresholds which limits the score classification ability and hence its clinical use, however; we chose a score of ≤ 3 as a threshold to rule out presence of obstructive CAD in this population due to its diagnostic characteristics. This threshold has a specificity of 0.90 (CI:0.89-0.91), a negative predictive value of 0.981(CI:0.977-0.991) and negative likelihood ratio of 0.813 (CI: 0.716-0.910).

Table 32: Frequency of predictors selection in the bootstrap models

Variable	**Selected in the Final Model**	**Frequency Selected in Bootstrap Models N (%)**
Gender	✓	475 (95)
Chest pain typicality	✓	465 (93)
Diagnosis with hyperlipidemia	✓	392 (78)
Baseline systolic blood pressure	✓	251 (50.2)
Treatment with Aspirin		216 (43)
Treatment with beta blockers		146 (29)
Treatment with lipid lowering agents		76 (15)
Diagnosis with hypertension		32 (6)
Probability of CAD		16 (3)
Creatinine		16 (3)

Table 33: Pooled estimates of multivariable model in bootstrap samples vs derivation dataset

Parameter	Bootstrap samples- Pooled estimates			Derivation Cohort		
	Regression Coefficient	Standard Error	*P* value	Regression Coefficient	Standard Error	*P* value
Intercept	-6.091			-6.084	0.75	<0.0001
Gender	0.730	0.203	0.012	0.721	0.20	0.0003
Typical	1.101	0.281	0.011	1.112	0.28	<0.0001
Hyperlipidemia	0.721	0.200	0.010	0.711	0.20	0.0003
Baseline systolic blood pressure	0.012	0.005	0.141	0.011	0.005	0.04

Table 34: Risk score points derived from the final multivariable model using the univariate screening

Variable	Points
Sex - Male - Female	 1 0
Chest pain typicality - Typical - Atypical, non-cardiac or no chest pain	 2 0
Hyperlipidemia - Yes - No	 1 0
Baseline systolic blood pressure - ≥140 mmHg - < 140 mmHg	 1 0
Total	5

Figure 10: ROC for the risk score derived from the reduced model

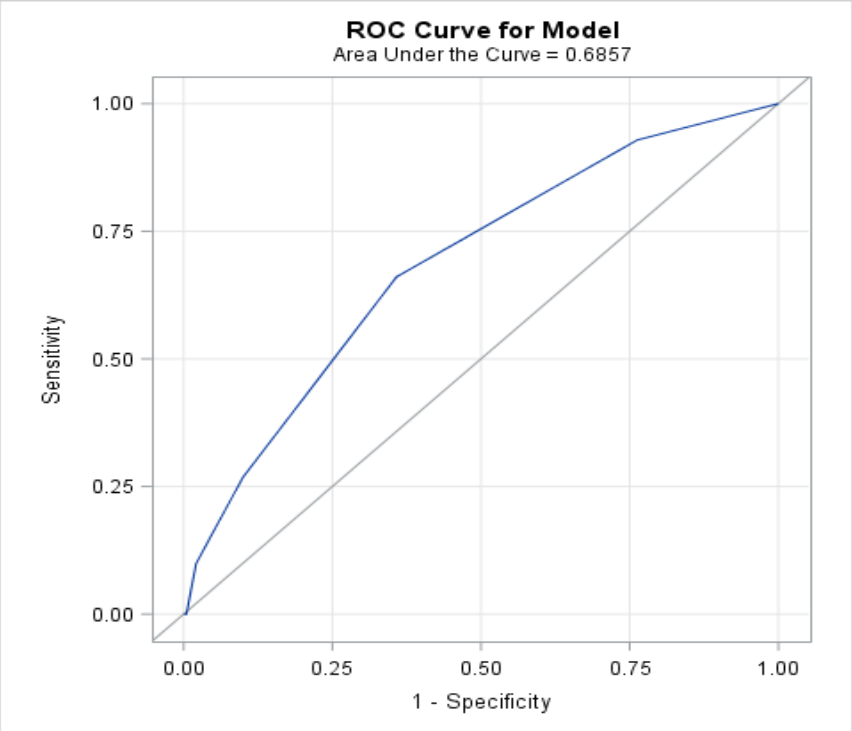


Table 35: Risk score thresholds with corresponding risk of obstructive CAD

Risk Threshold	Observed Risk %	Predicted Risk %	Number of Patients
0	0.70	0.90	1,147
1	1.53	1.62	1,967
2	3.4	2.92	1,282
3	4.79	5.23	397
4	12.5	9.20	88
5	0	15.65	22

Table 36: Operating characteristics of risk score thresholds derived from the model developed using univariate screening

Score	Sensitivity	Specificity	PPV	NPV	PLR	NLR
0	1.00 (0.968-1.00)	--	0.023 0.019-0.027	--	--	--
1	0.929 (0.864-0.969)	0.238 (0.226-0.250)	0.028 (0.022-0.034)	0.993 (0.986-0.997)	1.218 (1.149-1.288)	0.301 (0.087-0.514)
2	0.661 (0.565-0.747)	0.642 (0.628-0.656)	0.041 (0.033-0.052)	0.988 (0.983-0.991)	1.846 (1.575-2.116)	0.528 (0.383-0.673)
3	0.268 (0.189-0.360)	0.900 (0.892-0.910)	0.059 (0.040-0.083)	0.981 (0.977-0.985)	2.690 (1.783-3.600)	0.813 (0.716-0.910)
4	0.098 (0.051-0.169)	0.980 (0.974-0.983)	0.100 (0.051-0.172)	0.979 (0.975-0.982)	4.753 (1.757-7.754)	0.921 (0.861-0.961)
5	00	0.995 (0.993-0.997)	--	0.977 (0.973-0.983)	00	1.027 (0.825-1.229)

CHAPTER 5: DISCUSSION

5.1: Summary of Findings

In this thesis, we present two clinical prediction models to predict the presence of obstructive coronary artery disease in patients with calcium score of zero. We used two distinct approaches to develop our models, the first model was developed using pre-specification of the model predictors based on input from the clinical content experts and we called it "The Full Model"; and the second model was developed based on the univariate screening for statistically significant predictors in the univariable analysis and we called it "The Reduced Model". We used the same database for derivation of both models; however, we validated our developed full model externally using a non-overlapping dataset from CONFIRM (Coronary CT Angiography Evaluation of Clinical Outcomes: An International Multicenter). The reduced model was validated internally through bootstrap technique. We derived a point score to serve as a tool for clinical use from each of the developed models.

For the development of the model using the full model approach, a list of candidate predictors was identified based on a review of the literature, input from the clinical content expert and data availability. The final list of the variables of the multivariable model was pre-specified by surveying a group of practicing physicians who are expert in managing the clinical problem under investigation. They were asked to select a list of eight predictors from the list of all candidate predictors. The selected predictors were included in a multivariable logistic regression model. Continuous variables were kept continuous during modelling without any categorization. Data on variables selected to be included in the multivariable clinical model were complete with no imputation needed for deriving of the model using the full model approach. The final multivariable model using the full model approach included variables of age, sex, typical chest pain, family history of coronary artery disease, diagnosis with hyperlipidemia, diagnosis with hypertension and diagnosis with diabetes mellitus. Current smoking was selected by the surveyed physicians, however; it had a paradoxical association with the outcome of obstructive CAD. We evaluated the association between smoking in general, including current and past smoking; and the presence of obstructive CAD but the result was similar. The univariable analysis of the current smoking and outcome of obstructive CAD shoed that current smoking is associated with insignificant increase

of obstructive CAD. The same was observed in the univariable and multivariable analysis of the association between smoking and the outcome of obstructive CAD in patients with and without coronary calcification together. This may be explained by the tendency to investigate smoking patients with an over-all lower risk profile. The smoking status definition is challenging as the dose-response relationship is not well understood. Other possible explanation is the limited role of smoking as a risk factor in patients who have no coronary calcification. The discriminative performance of the developed model as assessed by AUC of the ROC was 0.68 (CI:0.63-0.73) which indicates a fair discriminative ability. The goodness of fit of the developed model as assessed by Hosmer-Lemeshow statistic revealed a chi-square statistic of 4.08 and a p value of 0.85 which indicates that the developed model fit well the derivation dataset as p value is > 0.05.

To present our model for the clinical use; we developed a clinical tool in form of risk score based on the regression coefficients of the variables in the final model with the resulting score ranged from zero to 20. The diagnostic operating characteristics for each risk threshold of the score was calculated. Based on the likelihood ratio for developing the outcome of obstructive CAD for each risk threshold on the score, the risk was categorized into low (score ≤6); intermediate (score 7-13) and high (≥14). The score seems to perform well at lower thresholds and would serve better to rule out obstructive CAD. The proportion of patients with obstructive CAD in low risk group is 1.15%.

The developed score using the full model approach was validated in an external dataset. The score discrimination performance in the validation dataset was acceptable with AUC of the ROC of 0.688 (0.651-0.725). The goodness of fit of the score in the validation dataset was assessed using Hosmer-Lemeshow statistic and showed that the score fit well the validation dataset (chi-square statistic of 6.48 and a p value of 0.59). The accuracy of the developed score when used to identify the low risk group (score ≤ 6) is > 97%.

In the model developed using the "Reduced Model Approach"; we used the same initial list of candidate predictors. We performed a univariate screening for selection of the predictors to be included in the initial multivariable model based on the statistical significance of the association between each predictor and the outcome in the univariate analysis as defined by a p value of <0.1. All predictors with significant association in the univariate analysis were included in an initial multivariable logistic regression model with a step-wise selection of the final multivariable clinical

model. The final multivariable clinical model using this approach included variables of male sex, typical chest pain, hyperlipidemia and systolic blood pressure. The systolic blood pressure was the only continuous variable in the final model using the reduced model approach and was modelled as a continuous variable. The discriminative ability of the model developed using the reduced model approach was assessed as for the model developed using the full model approach using AUC of ROC which was 0.685, which indicates an acceptable discriminative performance. The goodness of fit of the developed model was assessed using Hosmer-Lemeshow statistic with chi-square statistic value of 6.82 and p value of 0.56, which indicates that the developed model fit well the derivation dataset.

We validated the reduced model internally using bootstrapping. The variables selected in the final multivariable model through the step-wise automated selection were consistently selected by the 500 bootstraps with variable frequency. The model had a comparable performance in the bootstrap samples.

We derived a risk score from the developed model based on the regression coefficients of the variables in the final multivariable model. The score points ranged from zero to 5. The diagnostic characteristics for each risk threshold on the score were calculated. A score of ≥4 had a high specificity for diagnosis of obstructive CAD (0.98 CI: 0.974-0.983) which can be used to rule out obstructive CAD. The risk categories based on this risk score are low risk (≤3) and high risk (≥4). The accuracy of the score developed using the reduced model approach to identify patients with low risk of obstructive CAD is > 97%.

When the two models were compared, we found that variables of sex , typical chest pain and hyperlipidemia were selected in both models using pre-specification or univariate screening approaches, while variables of other risk factors like age, hypertension ; family history and diabetes mellitus did not meet the significance level in either the univariate analysis or the step-wise selection in reduced model approach despite been selected a priori by the clinical expert in the full model approach. The discriminative ability of the two models were comparable when assessed using the AUC of ROC. The two models fit well the derivation dataset when assessed by Hosmer-Lemeshow statistics with a chi-square statistic of 4.08 and 6.82 for the for the full and reduced model respectively and a p value of >0.05 for both models.

Table 37: Comparison of the developed models

Model Parameters	**Full Model**		**Reduced Model**	
	Odds Ratio (CI)	**P value**	**Odds Ratio (CI)**	**P value**
Age	1.009	0.3665	NA	
Male	2.23	0.0001	2.05	0.0003
Typical chest pain	2.87	0.0001	3.04	<0.0001
Hyperlipidemia	1.87	0.003	2.04	0.0003
Family history	1.26	0.237	NA	
Hypertension	1.23	0.298	NA	
Diabetes mellitus	1.17	0.589	NA	
Baseline systolic blood pressure	NA		1.01	0.04
Performance Measures	**Full Model**		**Reduced Model**	
AUC of ROC curve	0.680		0.685	
Hosmer-Lemeshow				
Chi-square	4.08		6.82	
DF	8		8	
P value	0.85		0.56	

Abbreviation:

CI, confidence interval; AUC, area under the curve; ROC, receiver operating characteristic

DF, degree of freedom

The developed scores from both models were different in the range; zero to 20 for the score derived from the full model; and zero to 5 for the score derived from the reduced model. The range of the score follows the number of the variables in the final model and their regression coefficients. Although the discriminative ability of both scores were similar, the score derived from the full model over the one derived from the reduced model has a wider range and the risk can be categorized into low, intermediate and high risk.

5.2: Interpretation of the Results in Context of Available Studies

The results of our developed models are interpreted in context of two streams of the available published evidence. The first is the available clinical prediction tools to predict the presence of obstructive CAD and the second is the diagnostic utility of calcium score of zero.

5.2.1: Available Tools to Predict Obstructive CAD

The variables included in our final full model (developed using the model pre-specification approach) were included as predictors in the available models with some. A summary of the published models for prediction of CAD is presented in Table 1.

For our reduced model (developed using univariate screening), the variable baseline systolic blood pressure as measured prior to performing coronary CTA was included as a predictor in our model for the direst time as a predictor for obstructive CAD. This variable is unique to the clinical situation of performance of coronary CTA and it was included as one of the candidate variables as a part of the clinical assessment variables available to us in the database. Elevated blood pressure is a well-known risk factor for coronary artery disease. In our database, info on risk factors for coronary artery disease were either self-reported by the patient or recorded based on information available in the medical records. Hypertension, as a variable in our database, was defined as a prior diagnosis of hypertension (blood pressure ≥140/90 mmHg) based on medical record or patient reporting; or treatment with blood pressure lowering medications. Although hypertension as defined in the database was not significant in the multivariable reduced model and hence was not included in the final model, the baseline systolic blood pressure could reflect a status of poorly controlled blood pressure in general and as a result it is significantly associated with the presence of obstructive CAD in the multivariable model.

When we compared the variables included in our two models with those in the available models, we observed that variables related to demographics such as age and sex as well as chest pain and chest pain typicality were explicitly identified as predictors in all published tools for prediction of coronary artery disease. Other predictors in our final full model, namely: family history, hyperlipidemia, hypertension and diabetes mellitus were included variably in the published models. Diabetes mellitus was included in six out of the nine published models, hyperlipidemia and hypertension were included in five models and family history was included in three models only. Three published models did not include any of the traditional risk factors for coronary artery disease, the model by Diamond-Forrester et. al. and its updated version and the model by Bosner et. al.[106,110,111]. The model by Diamond-Forrester was developed by reviewing the literature to identify its predictors using the Bayesian theorem and hence, the developing method could have led to a limited number of predictors identified[106]. In the updated version of the model by Diamond-Forrester, Genders et. al. did not include other predictors but rather extended the age intervals and modified the format of presentation to a classic format of a clinical tool[110]. In the model by Bosner et. al., data were obtained from a general practice setting with possible lower prevalence of the risk factors for CAD. History of prior diagnosis with CAD or myocardial infarction was included as a predictor in four of the published models to predict presence CAD. Two of these models were derived from databases of cardiac catherization laboratory by Pryor et. el., and two models were derived from general practice, the model by Gencer et. al. and the model by Bosner et. al. History of prior diagnosis with CAD including myocardial infarction is a well-known risk factor for recurrence of coronary artery disease events and hospitalization[1] . Although this variable is important in prediction models of CAD, but depending on the exact outcome being investigated, it is not relevant in our model as our goal is to derive a tool that predict presence of obstructive CAD in patient with no coronary calcification; and we excluded all patients with known prior diagnosis of CAD or revascularization. The coronary artery calcification was included as a predictor in one of the published models, which is the most recently published by Genders et. al., in contrast to our models which were developed in patients with zero calcification [113].

The outcome of the existing models varies significantly as presented in Table 1. The variation was not only confined to the definition but extended to the method of outcome assessment. In the model by Gencer et. al. and Bosner et. al., the outcome assessment was

performed in a follow up visit up to one year after initial assessment based on self-reporting and results of work up. The assessment of outcome in these two models is at risk of recall bias and misclassification. The outcome of our developed models was obstructive CAD as defined by the presence of ≥50% luminal stenosis of any segment of the major epicardial coronary arteries as assessed by coronary CTA. Although conventional coronary angiogram is the gold standard for the diagnosis of coronary artery disease, coronary CTA has been shown to have a high diagnostic accuracy for obstructive coronary artery disease especially in populations with low to intermediate pre-test probability[26,27]. However, this modality seems to have the best performance in ruling out the disease given the consistent high negative predictive value across multiple studies [30,124,125].

Our population is composed of patients with zero calcium score and a low prevalence of obstructive CAD. For derivation of our model, we used data from the University of Ottawa Heart Institute Cardiac CT Registry. This registry included all patients who were referred for coronary CTA, however, our cohort had a low prevalence of obstructive CAD. This is likely due to referral bias as the same institute has a well-established cardiac catheterization and nuclear perfusion imaging services. These two services are likely to have more referrals especially for patients with a higher probability for obstructive CAD. In two of the published models (the Duke clinical scores by Pryor et. al.), data were obtained from cardiac catherization database which is likely to have higher prevalence of obstructive CAD compared to our population. As the gold standard diagnostic modality, conventional angiogram is the last line of working up of chest pain and the probability of obstructive CAD may have increased based on results of non-invasive testing prior to coronary angiogram. In the model by Morise et. al. and the two models by Genders et. al., the data for model development were obtained from cohorts of mixed patients with high and low probability of obstructive CAD referred for work up of chest pain using non-invasive modalities or conventional coronary angiogram. In the models by Gencer et. al. and Bosner et. al., data used for the development of their models were obtained from clinical assessment of symptomatic patients in a general practice setting. This setting is different from all other settings, including ours, due to the heterogeneity of the population of such setting compared to specialist practice setting. The clinical setting of the developed tool is important for generalizability of the tool.

From methodological point of view, our developed tools were developed using a rigorous statistical method in derivation and validation phases. Most of the published models for prediction

of CAD have methodological issues of concern, in particular; related to dealing with missing values and modelling of continuous predictors.

Overall, our developed tools were developed uniquely to predict obstructive CAD in patients with zero calcium score using a rigorous statistical approach. These tools serve better as a ruling out tools due to their performance in lower risk groups.

5.2.2: Diagnostic Utility of Zero Calcium Score

Coronary artery calcification has important diagnostic and prognostic implications. CAC is a marker of atherosclerotic disease and is associated with future cardiovascular events and all-cause mortality [38,40,49,135–138]. When added to conventional risk factors, calcium score improves the performance of prediction models for cardiovascular events and improves the reclassification of individuals' risks. Adding the Agatston score to the Framingham risk score led to significant reclassification of individuals to higher or lower risk categories [60–62,139].

The Agatston score = 0 has been investigated in several studies of asymptomatic and symptomatic participants[38–40,49,50]. Raggi et. al. reported very low annual coronary event rates of 0.11% for Agatston score of zero compared to 4.8% for score of 400 or more in asymptomatic patients[49]. Among 3,409 patients with Agatston score of zero in the Multi-Ethnic Study of Atherosclerosis (MESA), only 0.4% developed any coronary event over the follow up period of 3 years; versus an event rate of 8% for those with Agatston score ≥ 300[40]. Despite the prognostic utility of zero calcium score as proven by the low cardiovascular events rates, the presence of obstructive CAD among these patients cannot be absolutely ruled out. Several earlier studies reported a prevalence of obstructive CAD in patients with zero calcium score that varied widely from 7% to 38%[57,140–143]. This is likely explained by the high-risk presentations of populations studied and technology used. Villines et. al. reported a prevalence of 3.5% of obstructive CAD among patients with an Agatston score of zero[58]. More recently, Mittal et. al. reported a lower prevalence rate of obstructive CAD in patients with zero calcium score of 1.4 % in a cohort of mostly asymptomatic patients and patients with atypical presentation [59].

The diagnostic uncertainty of an Agatston score = 0 has limited its clinical use to rule out obstructive CAD. Our proposed clinical risk score when combined with calcium score can improve the diagnostic utility of an Agatston score = 0 by allowing it to rule out obstructive CAD with a

negative predictive value of 99%. Based on the performance of this risk score, we propose a new management algorithm for work up of suspected CAD when CCTA is considered (Figure 4). For patients presenting for CCTA to rule out obstructive CAD, those with a low risk score (≤ 6), an Agatston score be performed. In those with an Agatston score = 0, the presence of obstructive CAD can be ruled out with high certainty. Theoretically, this approach will result in lower radiation exposure, eliminate the need for contrast media and reduce healthcare costs.

5.3: Methodological Aspects of Models Development

We developed two models to predict obstructive CAD among patients with zero calcium score using two distinct approaches. While we used model pre-specification approach to develop the first model (the full model), we used the univariate screening approach to develop the second (the reduced model). We will discuss some of the methodological aspects of the process of development and validation of the two models to predict obstructive CAD. In particular, we will discuss the models' specification, modelling the continuous predictors, dealing with missing data especially in the validation dataset, validation of the developed models and the derived scores.

5.3.1: Methods of Models' Specification

The main difference between the two approaches is the method of model predictors selection. In the model pre-specification approach, the list of predictors to be included in the final model were selected a priori by clinical experts who are considered the end-user of the developed tool. Model predictors in the univariate screening approach were based on the significance of association between each candidate predictor and the outcome in the univariable analysis to be included in an initial multivariable model then stepwise selection of predictors in the final multivariable model. The latter approach is statistically driven with no or minimal clinical consideration when select the predictors in the final model. Selection of predictors based on their statistical significance only is associated with few issues such as overfitting of the model, overestimated regression coefficients, biased estimate of the model performance and inclusion of predictors with less clinical utility and exclusion of more clinically relevant predictors. In our developed model, the main issue with the model developed using the univariate screening approach, which is a statistically driven selection of predictors, was the number of variables

included in the final multivariable model with only four predictors. Several predictors that are considered clinically relevant based on the available evidence and the selection of such predictors by the surveyed clinical experts for the full model were not included in the reduced model due to lack of statistical significance. The reduced model, which was developed using the univariate screening approach, included variables of sex, typical chest pain, hyperlipidemia and baseline systolic blood pressure; while the full model included variables of age, sex, typical chest pain, hyperlipidemia, family history, hypertension and diabetes mellitus. Variables such as risk factors including hyperlipidemia, hypertension, diabetes mellitus and family history are important for assessment of risk of developing coronary artery disease and would be expected to be included in tools for such purpose. Exclusion of these variables from the model developed by the univariate screening approach will limit its utility in clinical practice despite the acceptable discriminatory performance of this model.

Model pre-specification can be a challenging process depending on the nature of the clinical problem under investigation. Rare diseases (outcomes) or diseases not well defined may not be suitable for model pre-specification. Predictors exploration in such cases using the univariate screening is a preferred approach. The selection process for our full model was conducted by surveying practicing cardiologists who are probably more aware of the cardiovascular disease literature compared to general practitioners or specialists other than cardiologists. The list of predictors selected a priori was consistent with what has been published in most of previous models and known to be associated with the outcome being modelled.

5.3.2: Modelling Continuous Predictors

The number of continuous predictors in our models was limited. There was one continuous predictor in each model, the age in the full model, and the baseline systolic blood pressure in the reduced model. We modelled these two variables as continuous predictors without categorization as this approach is recommended by most of the experts. Dichotomization of continuous variable is associated with loss of information as discussed in chapter 2. However, to simplify the use of the score derived from our models. we categorized age into six groups based on the clinical relevance of these age groups and age categorization in some of the published models[114]. The baseline systolic blood pressure variable in the reduced model was dichotomized into normal

baseline systolic blood pressure and elevated baseline systolic blood pressure (≥140 mmHg) based on the clinical guidelines published by the 8th Joint National Commission (JNC8) 2014 report [134]. The categorization of the age in the full model and dichotomization of the baseline systolic blood presser did not affect the score performance compared to the model performance. Advanced functions for transformation of continuous predictors would be considered in more complex models with several continuous predictors, especially continuous predictors without a clinical basis for dichotomization or categorization.

5.3.3: Dealing with Missing Data

Missing data is frequently encountered in health research. The proportion of missing data in our derivation dataset was very small overall. We performed multiple imputation for the missing values of the variables included in the reduced model, but no imputation was needed for the development of the full model.

In the validation dataset, the proportion of missing values of some of the variables of our final model was large. In some of the centers, the missing proportion of typical chest pain or any of the criteria used to define typical chest pain reach up to 100%. This large missing proportion was considered as an indicator of a low-quality data collected at these centers. The decision to exclude these centers was reasonable. When complete case analysis was performed, the score performance had a comparable result with an AUC of 0.688 (Appendix 10). Imputation of all missing values without limiting to a certain proportion (≤20% was selected in this analysis) of the whole set would be a source of significant bias. We followed a strategy combining exclusion of centers that display large missing proportions as this indicates a lower quality of the collected data and performing multiple imputation of the missing values in the included center if the proportion of the missing data is 20% or less of the total number of observations included for the final validation process.

5.3.4: Validation of the Developed Models

Validation of the developed tools is important to assess how valid these tools when used in new subjects. We validated the score developed from the full model externally using an international multi-center registry (CONFIRM registry)[130]. The score performance in the

validation dataset was comparable to the score performance in the derivation dataset. We think that the score derived from the full model can be used with confidence among patients with Agatston score of zero to categorize patients in term of risk of obstructive CAD. The validation cohort was slightly different from the derivation cohort, however, both cohorts are from the same population with low prevalence of obstructive CAD which is the scope of our study.

The reduced model was validated internally using the bootstrap technique. It showed a comparable performance in term of the consistency of variable selection in the bootstrap samples and model performance in these samples which indicates the degree of reproducibility of the developed model.

5.3.5: Classification Thresholds of the Developed Scores

The choice of a risk threshold on a score consisting of multiple levels is a function of the threshold diagnostic ability i.e. ability to identify the true positive cases (those with the outcome) and the true negatives (those without the outcome). Sensitivity of score level is defined as the proportion of the patients who scored that level and have the outcome from all patients who scored that level (i.e. True Positives/ (True Positives + False Negatives)). Specificity of a score level is defined as the proportion of the patients who scored that level and don't have the outcome from all patients who scored that level (i.e. True Negatives/ (True Negatives +False Positives)). These diagnostic measures of the score levels do not consider the prevalence of the disease or outcome in the population being investigated. Thus, if a disease or outcome is frequent, the score level has to be sensitive to be useful. On the contrary, if a disease or outcome is rare, the score level has to be specific to be useful. Estimation of prevalence of the disease or outcome for each patient by determining the prior likelihood of the disease from the available literature and based on the demographic features.

Due to the limitation of sensitivity and specificity, we used the likelihood ratio to select the thresholds of risk classification of the score derived from the full model. The likelihood ratio of a score level incorporates the sensitivity and specificity of the level and provides an estimate of the change in the odds of having the outcome for at that level.

5.4: Limitations

A major limitation of our study is the definition of the outcome of obstructive CAD, which is defined as ≥50% luminal stenosis of any segment of coronary arteries by coronary CTA. The gold standard modality for the diagnosis of obstructive CAD is the invasive coronary angiography, however; coronary CTA has shown an overall good correlation with invasive coronary angiography as discussed previously (Section 3.3.3.1).

Another limitation is related to the low prevalence of obstructive CAD among patients referred for coronary CTA for work up of chest pain in our institution. This results in referral bias as the case in most of tertiary care centers where referring physicians have access to several modalities for work up of CAD such as exercise treadmill tests, nuclear perfusion scans, echocardiographic stress tests and invasive coronary angiograms. In clinical practice, coronary CTA is popular as a rule-out test and thus patients with low probability of obstructive CAD are more likely to be referred to coronary CTA for work up of chest pain or for screening for risk stratification.

The third limitation is the large proportion of missing values on predictors and outcome among patients enrolled at certain centers in the validation dataset. We excluded these centers which results in exclusion of about 50% of the total eligible patients for validation. The decision to exclude these centers assumed that the quality of the data collected at these centers is questionable since the variable of chest pain and its typicality is crucial in such registry.

Finally, in the full model, the variable of current smoking was selected a priori by the surveyed cardiologists, however, smoking (current and previous) was associated with a negative regression coefficient in the multivariable model which is paradoxical to the well-established association of smoking with adverse health outcomes including coronary artery disease. This is likely due to relying on patient reporting for collection of information on the risk factors, including smoking, in our database and the potential misregistration of the pertaining data.

5.5: Strengths

Our study is the first study to develop a clinical score to predict obstructive CAD in patients with Agatston score of zero. So far, no published clinical tool for such purpose. Although the

developed tools, in particular the score derived from the full model, can be used to predict obstructive CAD among this population; the utility of our scores seems to be at the best in ruling out the outcome of obstructive CAD. Ruling out obstructive CAD in patients with low risk (score≤6) would result in avoiding further testing with elimination of unnecessary radiation and reducing the costs of healthcare. Identifying patients with high risk of obstructive CAD using the developed score may lead clinicians to choose more appropriate alternative modalities such as nuclear perfusion testing.

In addition to the clinical importance of our developed scores, we used a database with high standards in term of collection, maintenance and auditing of the data. Our sample size was more than 4,900 patients with no missing values when derived the full model and a very small proportion of missing values when developed the reduced one. Furthermore, our score derived from the full model was validated externally using an international multi-center registry.

5.6: Future Research

In our study, we developed two clinical tools to predict obstructive CAD in patients with Agatston score of zero. We validated the score derived from the full model externally. Both derivation and validation were performed retrospectively. Future research for external validation of the score derived from the reduced model to assess its external validity prior to clinical implementation is planned. Another future research area is to assess the acceptance and utilization of the validated tools in clinical practice.

A well-designed prospective study to derive and validate a clinical tool to predict obstructive CAD among patients with Agatston score zero is an ideal approach to narrow the knowledge gap in this area.

Finally, developing an algorithm using machine learning to help clinicians making decision for such population is underway.

5.7: Conclusion

In this study, we developed two clinical tools to predict obstructive CAD among patients with zero calcium score (Agatston score=0). We validated the score derived from the full model

externally and we think it can be used in such population with confidence. Its clinical use seems to be at its best in identifying patients with low risk of obstructive CAD and can rule out the outcome in this group with high accuracy. The score derived from the reduced model was validated internally and once validated externally can be used based on its performance in the external validation cohort.

Further, we used two distinct approaches to develop our models and compared the methodological steps of these two approaches and the impact on the clinical utility of the final tools. Future research is going to focus on the external validation of the tool derived from the reduced model, assessment of the clinical utility and acceptance of the developed tools and development of a clinical tool with the same purpose from a prospectively collected data.

References

1. Benjamin EJ, Blaha MJ, Chiuve SE, et al. Heart Disease and Stroke Statistics-2017 Update: A Report From the American Heart Association. *Circulation*. 2017;135(10):e146-e603. doi:10.1161/CIR.0000000000000485.

2. WHO | The global burden of disease: 2004 update. *WHO*. 2014. http://www.who.int/healthinfo/global_burden_disease/2004_report_update/en/. Accessed January 20, 2018.

3. Fihn SD, Julius Gardin CM, Chair V, et al. 2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS Guideline for the Diagnosis and Management of Patients With Stable Ischemic Heart Disease. *J Am Coll Cardiol J Am Coll Cardiol Immed Past Chair*. 2012;60(24):44-164. doi:10.1016/j.jacc.2012.07.013.

4. Torres FS, Crean AM, Nguyen ET, Paul N. Strategies for radiation-dose reduction and image-quality optimization in multidetector computed tomographic coronary angiography. *Can Assoc Radiol J*. 2010;61(5):271-279. doi:10.1016/j.carj.2009.11.013.

5. STRONG JP, McGILL HC. The natural history of coronary atherosclerosis. *Am J Pathol*. 1962;40:37-49. http://www.ncbi.nlm.nih.gov/pubmed/13917861. Accessed January 17, 2018.

6. Fuster V. Lewis A. Conner Memorial Lecture. Mechanisms leading to myocardial infarction: insights from studies of vascular biology. *Circulation*. 1994;90(4):2126-2146. http://www.ncbi.nlm.nih.gov/pubmed/7718033. Accessed January 17, 2018.

7. Benjamin EJ, Blaha MJ, Chiuve SE, et al. Heart Disease and Stroke Statistics—2017 Update: A Report From the American Heart Association. *Circulation*. 2017;135(10):e146-e603. doi:10.1161/CIR.0000000000000485.

8. Greenland P, Knoll MD, Stamler J, et al. Major Risk Factors as Antecedents of Fatal and Nonfatal Coronary Heart Disease Events. *JAMA*. 2003;290(7):891. doi:10.1001/jama.290.7.891.

9. Yusuf S, Hawken S, Ounpuu S, et al. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-

control study. *Lancet (London, England)*. 2004;364(9438):937-952. doi:10.1016/S0140-6736(04)17018-9.

10. Balady GJ, Larson MG, Vasan RS, Leip EP, O'Donnell CJ, Levy D. Usefulness of exercise testing in the prediction of coronary disease risk among asymptomatic persons as a function of the Framingham risk score. *Circulation*. 2004;110(14):1920-1925. doi:10.1161/01.CIR.0000143226.40607.71.

11. Sesso HD, Lee IM, Gaziano JM, Rexrode KM, Glynn RJ, Buring JE. Maternal and paternal history of myocardial infarction and risk of cardiovascular disease in men and women. *Circulation*. 2001;104(4):393-398. doi:10.1161/HC2901.093115.

12. Munafò M, Clark T, Johnstone E, Murphy M, Walton R. The genetic basis for smoking behavior: A systematic review and meta-analysis. *Nicotine Tob Res*. 2004;6(4):583-598. doi:10.1080/14622200410001734030.

13. Coady SA, Jaquish CE, Fabsitz RR, Larson MG, Cupples LA, Myers RH. Genetic Variability of Adult Body Mass Index: A Longitudinal Assessment in Framingham Families. *Obes Res*. 2002;10(7):675-681. doi:10.1038/oby.2002.91.

14. Harrison TA, Hindorff LA, Kim H, et al. Family history of diabetes as a potential public health tool. *Am J Prev Med*. 2003;24(2):152-159. doi:10.1016/S0749-3797(02)00588-3.

15. Deschepper CF, Boutin-Ganache I, Zahabi A, Jiang Z. In Search of Cardiovascular Candidate Genes: Interactions Between Phenotypes and Genotypes. *Hypertension*. 2002;39(2):332-336. doi:10.1161/hy0202.102787.

16. Lloyd-Jones DM, Nam B-H, D'Agostino, Sr RB, et al. Parental Cardiovascular Disease as a Risk Factor for Cardiovascular Disease in Middle-aged Adults. *JAMA*. 2004;291(18):2204. doi:10.1001/jama.291.18.2204.

17. Marenberg ME, Risch N, Berkman LF, Floderus B, de Faire U. Genetic Susceptibility to Death from Coronary Heart Disease in a Study of Twins. *N Engl J Med*. 1994;330(15):1041-1046. doi:10.1056/NEJM199404143301503.

18. Murabito JM, Pencina MJ, Nam B-H, et al. Sibling Cardiovascular Disease as a Risk Factor for Cardiovascular Disease in Middle-aged Adults. *JAMA*. 2005;294(24):3117.

doi:10.1001/jama.294.24.3117.

19. Shiraishi J, Kohno Y, Yamaguchi S, et al. Acute Myocardial Infarction in Young Japanese Adults. *Circ J*. 2005;69(12):1454-1458. doi:10.1253/circj.69.1454.

20. Ciruzzi M, Schargrodsky H, Rozlosnik J, et al. Frequency of Family History of Acute Myocardial Infarction in Patients With Acute Myocardial Infarction. *Am J Cardiol*. 1997;80(2):122-127. doi:10.1016/S0002-9149(97)00304-4.

21. Li R, Bensen JT, Hutchinson RG, et al. Family risk score of coronary heart disease (CHD) as a predictor of CHD: the atherosclerosis risk in communities (ARIC) study and The NHLBI Family Heart Study. *Genet Epidemiol*. 2000;18(3):236-250. doi:10.1002/(SICI)1098-2272(200003)18:3<236::AID-GEPI4>3.0.CO;2-0.

22. Williams RR, Hunt SC, Heiss G, et al. Usefulness of cardiovascular family history data for population-based preventive medicine and medical research (The Health Family Tree Study and the NHLBI Family Heart Study). *Am J Cardiol*. 2001;87(2):129-135. doi:10.1016/S0002-9149(00)01303-5.

23. Hawe E, Talmud PJ, Miller GJ, Humphries SE. Family History is a Coronary Heart Disease Risk Factor in the Second Northwick Park Heart Study. *Ann Hum Genet*. 2003;67(2):97-106. doi:10.1046/j.1469-1809.2003.00017.x.

24. Scanlon PJ, Faxon DP, Audet AM, et al. ACC/AHA guidelines for coronary angiography: executive summary and recommendations. A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee on Coronary Angiography) developed in collaboration with the Society for Cardiac Angiography and Interventions. *Circulation*. 1999;99(17):2345-2357. doi:10.1161/01.CIR.99.17.2345.

25. Meijboom WB, Meijs MFL, Schuijf JD, et al. Diagnostic Accuracy of 64-Slice Computed Tomography Coronary Angiography: A Prospective, Multicenter, Multivendor Study. *J Am Coll Cardiol*. 2008;52(25):2135-2144. doi:10.1016/J.JACC.2008.08.058.

26. Budoff MJ, Dowe D, Jollis JG, et al. Diagnostic Performance of 64-Multidetector Row Coronary Computed Tomographic Angiography for Evaluation of Coronary Artery

Stenosis in Individuals Without Known Coronary Artery Disease: Results From the Prospective Multicenter ACCURACY (Assessment by Coro. *J Am Coll Cardiol*. 2008;52(21):1724-1732. doi:10.1016/J.JACC.2008.07.031.

27. Miller JM, Rochitte CE, Dewey M, et al. Diagnostic Performance of Coronary Angiography by 64-Row CT. *N Engl J Med*. 2008;359(22):2324-2336. doi:10.1056/NEJMoa0806576.

28. Vanhoenacker PK, Heijenbrok-Kal MH, Van Heste R, et al. Diagnostic Performance of Multidetector CT Angiography for Assessment of Coronary Artery Disease: Meta-analysis. *Radiology*. 2007;244(2):419-428. doi:10.1148/radiol.2442061218.

29. Achenbach S. Cardiac CT: state of the art for the detection of coronary arterial stenosis. *J Cardiovasc Comput Tomogr*. 2007;1(1):3-20. doi:10.1016/j.jcct.2007.04.007.

30. Abdulla J, Abildstrom SZ, Gotzsche O, Christensen E, Kober L, Torp-Pedersen C. 64-multislice detector computed tomography coronary angiography as potential alternative to conventional coronary angiography: a systematic review and meta-analysis. *Eur Heart J*. 2007;28(24):3042-3050. doi:10.1093/eurheartj/ehm466.

31. Taylor AJ, Cerqueira M, Hodgson JM, et al. ACCF/SCCT/ACR/AHA/ASE/ASNC/NASCI/SCAI/SCMR 2010 Appropriate Use Criteria for Cardiac Computed Tomography: A Report of the American College of Cardiology Foundation Appropriate Use Criteria Task Force, the Society of Cardiovascular Computed Tomography, the American College of Radiology, the American Heart Association, the American Society of Echocardiography, the American Society of Nuclear Cardiology, the North American Society for Cardiovascular Imaging, the Society for Cardiovascular Angiograp. *J Am Coll Cardiol*. 2010;56(22):1864-1894. doi:10.1016/J.JACC.2010.07.005.

32. Budoff MJ, Achenbach S, Blumenthal RS, et al. Assessment of coronary artery disease by cardiac computed tomography: a scientific statement from the American Heart Association Committee on Cardiovascular Imaging and Intervention, Council on Cardiovascular Radiology and Intervention, and Committee on Cardiac Imaging, Council on Clinical Cardiology. *Circulation*. 2006;114(16):1761-1791.

doi:10.1161/CIRCULATIONAHA.106.178458.

33. Alluri K, Joshi PH, Henry TS, Blumenthal RS, Nasir K, Blaha MJ. Scoring of coronary artery calcium scans: History, assumptions, current limitations, and future directions. *Atherosclerosis*. 2015;239(1):109-117. doi:10.1016/j.atherosclerosis.2014.12.040.

34. Margolis JR, Chen JT, Kong Y, Peter RH, Behar VS, Kisslo JA. The diagnostic and prognostic significance of coronary artery calcification. A report of 800 cases. *Radiology*. 1980;137(3):609-616. doi:10.1148/radiology.137.3.7444045.

35. Bartel AG, Chen JT, Peter RH, Behar VS, Kong Y, Lester RG. The significance of coronary calcification detected by fluoroscopy. A report of 360 patients. *Circulation*. 1974;49(6):1247-1253. doi:10.1161/01.CIR.49.6.1247.

36. Tomography C, Agatston AS, Janowitz WR, et al. Quantification of coronary artery calcium using ultrafast computed tomography. *JACC*. 1990;15(4):827-832. doi:10.1016/0735-1097(90)90282-T.

37. Budoff MJ, Georgiou D, Brody A, et al. Ultrafast computed tomography as a diagnostic modality in the detection of coronary artery disease: a multicenter study. *Circulation*. 1996;93(5):898-904. doi:10.1161/01.CIR.93.5.898.

38. LaMonte MJ, FitzGerald SJ, Church TS, et al. Coronary Artery Calcium Score and Coronary Heart Disease Events in a Large Cohort of Asymptomatic Men and Women. *Am J Epidemiol*. 2005;162(5):421-429. doi:10.1093/aje/kwi228.

39. Taylor AJ, Bindeman J, Feuerstein I, Cao F, Brazaitis M, O'Malley PG. Coronary calcium independently predicts incident premature coronary heart disease over measured cardiovascular risk factors: Mean three-year outcomes in the Prospective Army Coronary Calcium (PACC) project. *J Am Coll Cardiol*. 2005;46(5):807-814. doi:10.1016/j.jacc.2005.05.049.

40. Detrano R, Guerci AD, Carr JJ, et al. Coronary Calcium as a Predictor of Coronary Events in Four Racial or Ethnic Groups. *N Engl J Med*. 2008;358(13):1336-1345. doi:10.1056/NEJMoa072100.

41. Vliegenthart R. Coronary Calcification Improves Cardiovascular Risk Prediction in the

Elderly. *Circulation*. 2005;112(4):572-577. doi:10.1161/CIRCULATIONAHA.104.488916.

42. Kavousi M, Desai CS, Ayers C, et al. Prevalence and Prognostic Implications of Coronary Artery Calcification in Low-Risk Women. *JAMA*. 2016;316(20):2126. doi:10.1001/jama.2016.17020.

43. Yamamoto H, Kitagawa T, Kihara Y. Clinical implications of the coronary artery calcium score in Japanese patients. *J Atheroscler Thromb*. 2014;21(11):1101-1108. http://www.ncbi.nlm.nih.gov/pubmed/25263530. Accessed March 11, 2017.

44. McEvoy JW, Blaha MJ, DeFilippis AP, et al. Coronary Artery Calcium Progression: An Important Clinical Measurement? *J Am Coll Cardiol*. 2010;56(20):1613-1622. doi:10.1016/j.jacc.2010.06.038.

45. Budoff MJ, Young R, Lopez VA, et al. Progression of Coronary Calcium and Incident Coronary Heart Disease Events. *J Am Coll Cardiol*. 2013;61(12):1231-1239. doi:10.1016/j.jacc.2012.12.035.

46. Coronary artery calcium score prediction of all cause mortality and cardiovascular events in people with type 2 diabetes: systematic review and meta-analysis. *BMJ*. 2013;346. doi:10.1136/bmj.f1654.

47. Mallikethi-Reddy S, Rubenfire M, Jackson LA, Brook RD. Coronary artery calcium in hypertension: a review. *J Am Soc Hypertens*. 2015;9(12):993-1000. doi:10.1016/j.jash.2015.09.004.

48. Shemesh J, Motro M, Morag-Koren N, et al. Coronary Artery Calcification Predicts Long-Term Mortality in Hypertensive Adults. *Am J Hypertens*. 2011;24(6):681-686. doi:10.1038/ajh.2011.28.

49. Raggi P, Cooil B, Callister TQ. Use of electron beam tomography data to develop models for prediction of hard coronary events. *Am Heart J*. 2001;141(3):375-382. doi:10.1067/mhj.2001.113220.

50. Arad Y, Goodman KJ, Roth M, Newstein D, Guerci AD. Coronary Calcification, Coronary Disease Risk Factors, C-Reactive Protein, and Atherosclerotic Cardiovascular

Disease Events. *J Am Coll Cardiol*. 2005;46(1):158-165. doi:10.1016/j.jacc.2005.02.088.

51. Georgiou D, Budoff MJ, Kaufer E, Kennedy JM, Lu B, Brundage BH. Screening patients with chest pain in the emergency department using electron beam tomography: a follow-up study. *J Am Coll Cardiol*. 2001;38(1):105-110. doi:10.1016/S0735-1097(01)01364-X.

52. Laudon DA, Vukov LF, Breen JF, Rumberger JA, Wollan PC, Sheedy PF. Use of Electron-Beam Computed Tomography in the Evaluation of Chest Pain Patients in the Emergency Department. *Ann Emerg Med*. 1999;33(1):15-21. doi:10.1016/S0196-0644(99)70412-9.

53. Anand D V., Lim E, Hopkins D, et al. Risk stratification in uncomplicated type 2 diabetes: prospective evaluation of the combined use of coronary artery calcium imaging and selective myocardial perfusion scintigraphy. *Eur Heart J*. 2006;27(6):713-721. doi:10.1093/eurheartj/ehi808.

54. Becker A, Knez A, Becker C, et al. Prädiktion kardiovaskulärer Ereignisse durch Koronarkalkbestimmung mit der Mehrzeilen-Computertomographie. *DMW - Dtsch Medizinische Wochenschrift*. 2005;130(43):2433-2438. doi:10.1055/s-2005-918586.

55. Valenti V, Hartaigh BÓ, Cho I, et al. Absence of Coronary Artery Calcium Identifies Asymptomatic Diabetic Individuals at Low Near-Term But Not Long-Term Risk of Mortality: A 15-Year Follow-Up Study of 9715 Patients. *Circ Cardiovasc Imaging*. 2016;9(2):e003528. doi:10.1161/CIRCIMAGING.115.003528.

56. Chaikriangkrai K, Palamaner Subash Shantha G, Jhun HY, et al. Prognostic Value of Coronary Artery Calcium Score in Acute Chest Pain Patients Without Known Coronary Artery Disease: Systematic Review and Meta-analysis. *Ann Emerg Med*. 2016;68(6):659-670. doi:10.1016/j.annemergmed.2016.07.020.

57. Gottlieb I, Miller JM, Arbab-Zadeh A, et al. The Absence of Coronary Calcification Does Not Exclude Obstructive Coronary Artery Disease or the Need for Revascularization in Patients Referred for Conventional Coronary Angiography. *J Am Coll Cardiol*. 2010;55(7):627-634. doi:10.1016/j.jacc.2009.07.072.

58. Villines TCTC, Hulten EAEA, Shaw LJLJ, et al. Prevalence and severity of coronary

artery disease and adverse events among symptomatic patients with coronary artery calcification scores of zero undergoing coronary computed tomography angiography: Results from the CONFIRM (Coronary CT Angiography Evalu. *J Am Coll Cardiol*. 2011;58(24):2533-2540. doi:10.1016/j.jacc.2011.10.851.

59. Mittal TK, Pottle A, Nicol E, et al. Prevalence of obstructive coronary artery disease and prognosis in patients with stable symptoms and a zero-coronary calcium score. *Eur Hear J - Cardiovasc Imaging*. 2017;2:404-411. doi:10.1093/ehjci/jex037.

60. Elias-Smale SE, Proença RV, Koller MT, et al. Coronary Calcium Score Improves Classification of Coronary Heart Disease Risk in the Elderly. *J Am Coll Cardiol*. 2010;56(17):1407-1414. doi:10.1016/j.jacc.2010.06.029.

61. Erbel R, Möhlenkamp S, Moebus S, et al. Coronary Risk Stratification, Discrimination, and Reclassification Improvement Based on Quantification of Subclinical Coronary Atherosclerosis. *J Am Coll Cardiol*. 2010;56(17):1397-1406. doi:10.1016/j.jacc.2010.06.030.

62. Polonsky TS, McClelland RL, Jorgensen NW, et al. Coronary Artery Calcium Score and Risk Classification for Coronary Heart Disease Prediction. *JAMA*. 2010;303(16):1610. doi:10.1001/jama.2010.461.

63. Kelkar AA, Schultz WM, Khosa F, et al. Long-Term Prognosis After Coronary Artery Calcium Scoring Among Low-Intermediate Risk Women and MenCLINICAL PERSPECTIVE. *Circ Cardiovasc Imaging*. 2016;9(4):e003742. doi:10.1161/CIRCIMAGING.115.003742.

64. Toll DB, Janssen KJM, Vergouwe Y, Moons KGM, Wragg A, Robson J. Validation, updating and impact of clinical prediction rules: A review. *J Clin Epidemiol*. 2008;61(11):1085-1094. doi:10.1016/j.jclinepi.2008.04.008.

65. Wasson JH, Sox HC, Neff RK, Goldman L. Clinical Prediction Rules. *N Engl J Med*. 1985;313(13):793-799. doi:10.1056/NEJM198509263131306.

66. Adams ST, Leveson SH. Clinical prediction rules. *BMJ*. 2012;344(jan16 1):d8312-d8312. doi:10.1136/bmj.d8312.

67. Draper D. Assessment and Propagation of Model Uncertainty. *J R Stat Soc Ser B*. 1995;57:45-97. doi:10.2307/2346087.

68. Chatfield C. Model Uncertainty, Data Mining and Statistical Inference. *J R Stat Soc Ser A (Statistics Soc*. 1995;158(3):419. doi:10.2307/2983440.

69. Gail MH, Brinton LA, Byar DP, et al. Projecting Individualized Probabilities of Developing Breast Cancer for White Females Who Are Being Examined Annually. *JNCI J Natl Cancer Inst*. 1989;81(24):1879-1886. doi:10.1093/jnci/81.24.1879.

70. Steyerberg EW. *Clinical Prediction Models : A Practical Approach to Development, Validation, and Updating*. Springer; 2009.

71. Lee Y-H, Bang H, Kim DJ. How to Establish Clinical Prediction Models. *org Endocrinol Metab*. 2016;31:38-44. doi:10.3803/EnM.2016.31.1.38.

72. HARRELL FE, LEE KL, MARK DB. MULTIVARIABLE PROGNOSTIC MODELS: ISSUES IN DEVELOPING MODELS, EVALUATING ASSUMPTIONS AND ADEQUACY, AND MEASURING AND REDUCING ERRORS. *Stat Med*. 1996;15(4):361-387. doi:10.1002/(SICI)1097-0258(19960229)15:4<361::AID-SIM168>3.0.CO;2-4.

73. Harrell FE. Multivariable Modeling Strategies. In: ; 2001:53-85. doi:10.1007/978-1-4757-3462-1_4.

74. Rutjes A, Reitsma J, Coomarasamy A, Khan K, Bossuyt P. Evaluation of diagnostic tests when there is no gold standard. A review of methods. *Health Technol Assess (Rockv)*. 2007;11(50). doi:10.3310/hta11500.

75. Wilson PWF, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of Coronary Heart Disease Using Risk Factor Categories. *Circulation*. 1998;97(18):1837-1847. doi:10.1161/01.CIR.97.18.1837.

76. Campbell MJ. Frontmatter. In: *Statistics at Square Two*. Oxford, UK: Blackwell Publishing Ltd; :i-x. doi:10.1002/9780470755839.fmatter.

77. Sevčíková H, Raftery A. Estimation of multinomial logit model using the Begg &

Gray approximation. 2012. https://cran.r-project.org/web/packages/mlogitBMA/vignettes/conversion.pdf. Accessed January 18, 2018.

78. Cox DR. Regression Models and Life-Tables. *J R Stat Soc Ser B*. 1972;34(2):187-220. http://links.jstor.org/sici?sici=0035-9246%281972%2934%3A2%3C187%3ARMAL%3E2.0.CO%3B2-6. Accessed January 18, 2018.

79. Balmaña J, Stockwell DH, Steyerberg EW, et al. Prediction of MLH1 and MSH2 Mutations in Lynch Syndrome. *JAMA*. 2006;296(12):1469. doi:10.1001/jama.296.12.1469.

80. Harrell FE, Lee KL, Califf RM, Pryor DB, Rosati RA. REGRESSION MODELLING STRATEGIES FOR IMPROVED PROGNOSTIC PREDICTION. *Stat Med*. 1984;3:143-152. https://s3.amazonaws.com/academia.edu.documents/45700180/sim.478003020720160517-9190-6akgcd.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1516321785&Signature=EIeZ72u1EpnF4IcUphLWk5j9A6o%3D&response-content-disposition=inline%3B filename%3DRegression_mo. Accessed January 18, 2018.

81. Steyerberg EW, Eijkemans MJ, Habbema JD. Stepwise selection in small data sets: a simulation study of bias in logistic regression analysis. *J Clin Epidemiol*. 1999;52(10):935-942. doi:10.1016/S0895-4356(99)00103-1.

82. Copas JB, Long T. Estimating the Residual Variance in Orthogonal Regression with Variable Selection. *Stat*. 1991;40(1):51. doi:10.2307/2348223.

83. Van Houwelingen JC, Le Cessie S. Predictive value of statistical models. *Stat Med*. 1990;9(11):1303-1325. http://www.ncbi.nlm.nih.gov/pubmed/2277880. Accessed January 18, 2018.

84. Babyak MA. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. *Psychosom Med*. 66(3):411-421.

http://www.ncbi.nlm.nih.gov/pubmed/15184705. Accessed January 18, 2018.

85. Bouwmeester W, Zuithoff NPA, Mallett S, et al. Reporting and Methods in Clinical Prediction Research: A Systematic Review. Macleod MR, ed. *PLoS Med*. 2012;9(5):e1001221. doi:10.1371/journal.pmed.1001221.

86. Moons KGM, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. *BMJ*. 2009;338(jun04 2):b606-b606. doi:10.1136/bmj.b606.

87. Peduzzi P, Concato J, Feinstein AR, Holford TR. Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates. *J Clin Epidemiol*. 1995;48(12):1503-1510. doi:10.1016/0895-4356(95)00048-8.

88. Pavlou M, Ambler G, Seaman SR, et al. How to develop a more accurate risk prediction model when there are few events. *bmj BMJ*. 2015;101383008. doi:10.1136/bmj.h3868.

89. Greenland S, Finkle WD. A Critical Look at Methods for Handling Missing Covariates in Epidemiologic Regression Analyses. *Am J Epidemiol*. 1995;142(12):1255-1264. doi:10.1093/oxfordjournals.aje.a117592.

90. Little RJA, Rubin DB. *Statistical Analysis with Missing Data*. https://www.wiley.com/en-us/Statistical+Analysis+with+Missing+Data%2C+2nd+Edition-p-9780471183860. Accessed January 18, 2018.

91. Habbema JD, Hilden J, Bjerregaard B. The measurement of performance in probabilistic diagnosis. I. The problem, descriptive tools, and measures based on classification matrices. *Methods Inf Med*. 1978;17(4):217-226. http://www.ncbi.nlm.nih.gov/pubmed/366334. Accessed January 18, 2018.

92. Donders ART, van der Heijden GJMG, Stijnen T, Moons KGM. Review: a gentle introduction to imputation of missing values. *J Clin Epidemiol*. 2006;59(10):1087-1091. doi:10.1016/j.jclinepi.2006.01.014.

93. DELPRIORE G, ZANDIEH P, LEE M. Treatment of continuous data as categoric variables in. *Obstet Gynecol*. 1997;89(3):351-354. doi:10.1016/S0029-7844(96)00504-2.

94. Lagakos SW. Effects of mismodelling and mismeasuring explanatory variables on tests of their association with a response variable. *Stat Med*. 7(1-2):257-274. http://www.ncbi.nlm.nih.gov/pubmed/3353607. Accessed January 18, 2018.

95. Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat Med*. 2006;25(1):127-141. doi:10.1002/sim.2331.

96. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Ann Intern Med*. 1999;130(6):515-524. http://www.ncbi.nlm.nih.gov/pubmed/10075620. Accessed January 18, 2018.

97. Altman DG, Royston P. What do we mean by validating a prognostic model? *Stat Med*. 2000;19(4):453-473. http://www.ncbi.nlm.nih.gov/pubmed/10694730. Accessed January 18, 2018.

98. Harrell FE. *Regression Modeling Strategies : With Applications to Linear Models, Logistic Regression, and Survival Analysis*. Springer New York; 2001.

99. Chernick MR, González-Manteiga W, Crujeiras RM, Barrios EB. Bootstrap Methods. In: *International Encyclopedia of Statistical Science*. Berlin, Heidelberg: Springer Berlin Heidelberg; 2011:169-174. doi:10.1007/978-3-642-04898-2_150.

100. Steyerberg EW, Bleeker SE, Moll HA, Grobbee DE, Moons KGM. Internal and external validation of predictive models: a simulation study of bias and precision in small samples. *J Clin Epidemiol*. 2003;56(5):441-447. http://www.ncbi.nlm.nih.gov/pubmed/12812818. Accessed January 18, 2018.

101. Wicki J, Perneger T V, Junod AF, Bounameaux H, Perrier A. Assessing clinical probability of pulmonary embolism in the emergency ward: a simple score. *Arch Intern Med*. 2001;161(1):92-97. http://www.ncbi.nlm.nih.gov/pubmed/11146703. Accessed January 18, 2018.

102. Sullivan LM, Massaro JM, D RB, Sr A. TUTORIAL IN BIOSTATISTICS Presentation of multivariate data for clinical use: The Framingham Study risk score functions. *Stat Med Stat Med*. 2004;23:1631-1660. doi:10.1002/sim.1742.

103. Wessler BS, Lai YH L, Kramer W, et al. Clinical Prediction Models for Cardiovascular

Disease. *Circ Cardiovasc Qual Outcomes*. 2015. http://circoutcomes.ahajournals.org/content/early/2015/07/07/CIRCOUTCOMES.115.001693. Accessed April 1, 2017.

104. Salisbury AC, Spertus JA. Realizing the Potential of Clinical Risk Prediction Models. *Circ Cardiovasc Qual Outcomes*. 2015;8(4). http://circoutcomes.ahajournals.org/content/8/4/332.full. Accessed May 9, 2017.

105. A G Damen JA, Hooft L, Schuit E, et al. Prediction models for cardiovascular disease risk in the general population: systematic review. *bmj BMJ BMJ*. 2016;353353. doi:10.1136/bmj.i2416.

106. Diamond GA, Forrester JS. Analysis of Probability as an Aid in the Clinical Diagnosis of Coronary-Artery Disease. *N Engl J Med*. 1979;300(24):1350-1358. doi:10.1056/NEJM197906143002402.

107. Pryor DB, Harrell FE, Lee KL, Califf RM, Rosati RA. Estimating the likelihood of significant coronary artery disease. *Am J Med*. 1983;75(5):771-780. doi:10.1016/0002-9343(83)90406-0.

108. Pryor DB, Shaw L, Harrell FE, et al. Estimating the likelihood of severe coronary artery disease. *Am J Med*. 1991;90(5):553-562. doi:10.1016/S0002-9343(05)80005-1.

109. Morise AP, Haddad WJ, Beckner D. Development and Validation of a Clinical Score to Estimate the Probability of Coronary Artery Disease in Men and Women Presenting with Suspected Coronary Disease. *Am J Med*. 1997;102(4):350-356. doi:10.1016/S0002-9343(97)00086-7.

110. Genders TSS, Steyerberg EW, Alkadhi H, et al. A clinical prediction rule for the diagnosis of coronary artery disease: validation, updating, and extension. *Eur Heart J*. 2011;32(11):1316-1330. doi:10.1093/eurheartj/ehr014.

111. Bösner S, Haasenritter J, Becker A, et al. Ruling out coronary artery disease in primary care: development and validation of a simple prediction rule. *CMAJ*. 2010;182(12):1295-1300. doi:10.1503/cmaj.100212.

112. Gencer B, Vaucher P, Herzig L, et al. Ruling out coronary heart disease in primary care

patients with chest pain: a clinical prediction score. *BMC Med*. 2010;8(1):9. doi:10.1186/1741-7015-8-9.

113. Genders TSS, Steyerberg EW, Hunink MGM, et al. Prediction model to estimate presence of coronary artery disease: retrospective pooled analysis of existing cohorts. *BMJ*. 2012;344. http://www.bmj.com.proxy.bib.uottawa.ca/content/344/bmj.e3485.long. Accessed May 9, 2017.

114. Yang Y, Chen L, Yam Y, et al. A Clinical Model to Identify Patients With High-Risk Coronary Artery Disease. *JACC Cardiovasc Imaging*. 2015;8(4):427-434. doi:10.1016/j.jcmg.2014.11.015.

115. Pryor DB, Shaw L, McCants CB, et al. Value of the History and Physical in Identifying Patients at Increased Risk for Coronary Artery Disease. *Ann Intern Med*. 1993;118(2):81. doi:10.7326/0003-4819-118-2-199301150-00001.

116. Eslick GD. Chest pain: a historical perspective. *Int J Cardiol*. 2001;77(1):5-11. doi:10.1016/S0167-5273(00)00395-8.

117. Kroenke K, Mangelsdorff AD. Common symptoms in ambulatory care: Incidence, evaluation, therapy, and outcome. *Am J Med*. 1989;86(3):262-266. doi:10.1016/0002-9343(89)90293-3.

118. Moons KGM, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration. *Ann Intern Med*. 2015;162(1):W1. doi:10.7326/M14-0698.

119. Goff DC, Lloyd-Jones DM, Bennett G, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *J Am Coll Cardiol*. 2014;63(25 Pt B):2935-2959. doi:10.1016/j.jacc.2013.11.005.

120. Piepoli MF, Hoes AW, Agewall S, et al. 2016 European Guidelines on cardiovascular disease prevention in clinical practice. *Eur Heart J*. 2016;37(29):2315-2381. doi:10.1093/eurheartj/ehw106.

121. Steyerberg EW, Moons KGM, van der Windt DA, et al. Prognosis Research Strategy

(PROGRESS) 3: Prognostic Model Research. *PLoS Med*. 2013;10(2):e1001381. doi:10.1371/journal.pmed.1001381.

122. Collins GS, Reitsma JB, Altman DG, Moons K. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. *BMC Med*. 2015;13(1):1. doi:10.1186/s12916-014-0241-z.

123. Leipsic J, Co-Chair F, Abbara S, et al. SCCT Guidelines SCCT guidelines for the interpretation and reporting of coronary CT angiography: A report of the Society of Cardiovascular Computed Tomography Guidelines Committee. *J Cardiovasc Comput Tomogr*. 2014;8:342-358. doi:10.1016/j.jcct.2014.07.003.

124. Menke J, Kowalski J. Diagnostic accuracy and utility of coronary CT angiography with consideration of unevaluable results: A systematic review and multivariate Bayesian random-effects meta-analysis with intention to diagnose. *Eur Radiol*. 2016;26(2):451-458. doi:10.1007/s00330-015-3831-z.

125. Salavati A, Radmanesh F, Heidari K, Dwamena BA, Kelly AM, Cronin P. Dual-source computed tomography angiography for diagnosis and assessment of coronary artery disease: Systematic review and meta-analysis. *J Cardiovasc Comput Tomogr*. 2012;6(2):78-90. doi:10.1016/J.JCCT.2011.10.018.

126. Mowatt G, Cook JA, Hillis GS, et al. 64-Slice computed tomography angiography in the diagnosis and assessment of coronary artery disease: systematic review and meta-analysis. *Heart*. 2008;94(11):1386-1393. doi:10.1136/hrt.2008.145292.

127. Austin PC, Tu J V. Automated variable selection methods for logistic regression produced unstable models for predicting acute myocardial infarction mortality. *J Clin Epidemiol*. 2004;57(11):1138-1146. doi:10.1016/j.jclinepi.2004.04.003.

128. Austin PC. Using the bootstrap to improve estimation and confidence intervals for regression coefficients selected using backwards variable elimination. *Stat Med*. 2008;27(17):3286-3300. doi:10.1002/sim.3104.

129. Altman DG, Andersen PK. Bootstrap investigation of the stability of a Cox regression model. *Stat Med*. 1989;8(7):771-783. http://www.ncbi.nlm.nih.gov/pubmed/2672226.

Accessed January 22, 2018.

130. Min JK, Dunning A, Lin FY, et al. Rationale and design of the CONFIRM (COronary CT Angiography EvaluatioN For Clinical Outcomes: An InteRnational Multicenter) Registry. *J Cardiovasc Comput Tomogr*. 2011;5(2):84-92. doi:10.1016/j.jcct.2011.01.007.

131. Hadamitzky M, Achenbach S, Al-Mallah M, et al. Optimized Prognostic Score for Coronary Computed Tomographic Angiography: Results From the CONFIRM Registry (COronary CT Angiography EvaluatioN For Clinical Outcomes: An InteRnational Multicenter Registry). *J Am Coll Cardiol*. 2013;62(5):468-476. doi:10.1016/J.JACC.2013.04.064.

132. Debray TPA, Vergouwe Y, Koffijberg H, Nieboer D, Steyerberg EW, Moons KGM. A new framework to enhance the interpretation of external validation studies of clinical prediction models. *J Clin Epidemiol*. 2015;68(3):279-289. doi:10.1016/j.jclinepi.2014.06.018.

133. Allison PD. *Missing Data*. Sage Publications; 2002. https://us.sagepub.com/en-us/nam/missing-data/book9419. Accessed February 27, 2018.

134. James PA, Oparil S, Carter BL, et al. 2014 Evidence-Based Guideline for the Management of High Blood Pressure in Adults. *JAMA*. 2014;311(5):507. doi:10.1001/jama.2013.284427.

135. Taylor AJ, Bindeman J, Feuerstein I, Cao F, Brazaitis M, O'Malley PG. Coronary Calcium Independently Predicts Incident Premature Coronary Heart Disease Over Measured Cardiovascular Risk Factors. *J Am Coll Cardiol*. 2005;46(5):807-814. doi:10.1016/j.jacc.2005.05.049.

136. Pletcher MJ, Tice JA, Pignone M, Browner WS, TJ O, WB K. Using the Coronary Artery Calcium Score to Predict Coronary Heart Disease Events. *Arch Intern Med*. 2004;164(12):1285. doi:10.1001/archinte.164.12.1285.

137. Al Rifai M, McEvoy JW, Nasir K, et al. Traditional cardiovascular disease risk factors associated with one-year all-cause mortality among those with coronary artery calcium scores ≥400. *Atherosclerosis*. 2015;241(2):495-497.

doi:10.1016/j.atherosclerosis.2015.06.002.

138. Budoff MJ, Hokanson JE, Nasir K, et al. Progression of Coronary Artery Calcium Predicts All-Cause Mortality. *JACC Cardiovasc Imaging*. 2010;3(12):1229-1236. doi:10.1016/j.jcmg.2010.08.018.

139. McClelland RL, Jorgensen NW, Budoff M, et al. 10-Year Coronary Heart Disease Risk Prediction Using Coronary Artery Calcium and Traditional Risk Factors: Derivation in the MESA (Multi-Ethnic Study of Atherosclerosis) With Validation in the HNR (Heinz Nixdorf Recall) Study and the DHS (Dallas Heart Study). *J Am Coll Cardiol*. 2015;66(15):1643-1653. doi:10.1016/J.JACC.2015.08.035.

140. Rubinshtein R, Gaspar T, Halon DA, Goldstein J, Peled N, Lewis BS. Prevalence and Extent of Obstructive Coronary Artery Disease in Patients With Zero or Low Calcium Score Undergoing 64-Slice Cardiac Multidetector Computed Tomography for Evaluation of a Chest Pain Syndrome. *Am J Cardiol*. 2007;99(4):472-475. doi:10.1016/j.amjcard.2006.08.060.

141. Haberl R, Tittus J, Böhme E, et al. Multislice spiral computed tomographic angiography of coronary arteries in patients with suspected coronary artery disease: An effective filter before catheter angiography? *Am Heart J*. 2005;149(6):1112-1119. doi:10.1016/j.ahj.2005.02.048.

142. Cademartiri F, Maffei E, Palumbo A, et al. Diagnostic accuracy of computed tomography coronary angiography in patients with a zero calcium score. *Eur Radiol*. 2010;20(1):81-87. doi:10.1007/s00330-009-1529-9.

143. Akram K, O'Donnell RE, King S, Superko HR, Agatston A, Voros S. Influence of symptomatic status on the prevalence of obstructive coronary artery disease in patients with zero calcium score. *Atherosclerosis*. 2009;203(2):533-537. doi:10.1016/j.atherosclerosis.2008.07.008.

Appendices

Appendix 1: TRIPOD Statement Checklist

TRIPOD Checklist: Prediction Model Development and Validation

Section/Topic	Item		Checklist Item	Page
Title and abstract				
Title	1	D;V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	
Introduction				
Background and objectives	3a	D;V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model or both.	
Methods				
Source of data	4a	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	
	4b	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	
Participants	5a	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	
	5b	D;V	Describe eligibility criteria for participants.	
	5c	D;V	Give details of treatments received, if relevant.	
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.	
Predictors	7a	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.	
Sample size	8	D;V	Explain how the study size was arrived at.	
Missing data	9	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	
	10c	V	For validation, describe how the predictions were calculated.	
	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	
Risk groups	11	D;V	Provide details on how risk groups were created, if done.	
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	
Results				
Participants	13a	D;V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	
	13b	D;V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	
	15b	D	Explain how to the use the prediction model.	
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.	
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	
Discussion				
Limitations	18	D;V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	
	19b	D;V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.	
Other information				
Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	
Funding	22	D;V	Give the source of funding and the role of the funders for the present study.	

*Items relevant only to the development of a prediction model are denoted by D, items relating solely to a validation of a prediction model are denoted by V, and items relating to both are denoted D;V. We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.

Appendix 2: Research Ethics Board Approval

Ottawa Hospital Research Institute
Institut de recherche
u Ottawa
UNIVERSITY OF OTTAWA HEART INSTITUTE
INSTITUT DE CARDIOLOGIE DE L'UNIVERSITÉ D'OTTAWA

Ottawa Health Science Network Research Ethics Board/Conseil d'éthique de la recherche du Réseau de science de la santé d'Ottawa
Civic Box 411 725 Parkdale Avenue, Ottawa, Ontario K1Y 4E9 613-798-5555 ext. 14902 Fax : 613-761-4311
http://www.ohri.ca/ohsn-reb

September 20, 2016

Dr. Ben Chow
Room H-1220-A, H1 PET Centre
Division of Cardiac Imaging
University of Ottawa Heart Institute
40 Ruskin Street
Ottawa, ON K1Y 4W7

Dear Dr. Chow:

Re: Protocol # 20160684-01H **Establishing a clinical model to improve management of patients suspected of coronary artery disease who have a coronary artery calcium score of 0**

Protocol approval valid until - September 18, 2017

I am pleased to inform you that this protocol underwent expedited review by the Ottawa Health Science Network Research Ethics Board (OHSN-REB) and is approved. No changes, amendments or addenda may be made to the protocol without the OHSN-REB's review and approval.

REB approval is for the REB application and Protocol Version 1, dated August 8, 2016.

If the study is to continue beyond the expiry date noted above, a Renewal Form should be submitted to the REB approximately six weeks prior to the current expiry date. If the study has been completed by this date, a Termination Report should be submitted.

OHSN-REB complies with the membership requirements and operates in compliance with the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans; the International Conference on Harmonization - Good Clinical Practice: Consolidated Guideline; the provisions of the Personal Health Information Protection Act 2004.

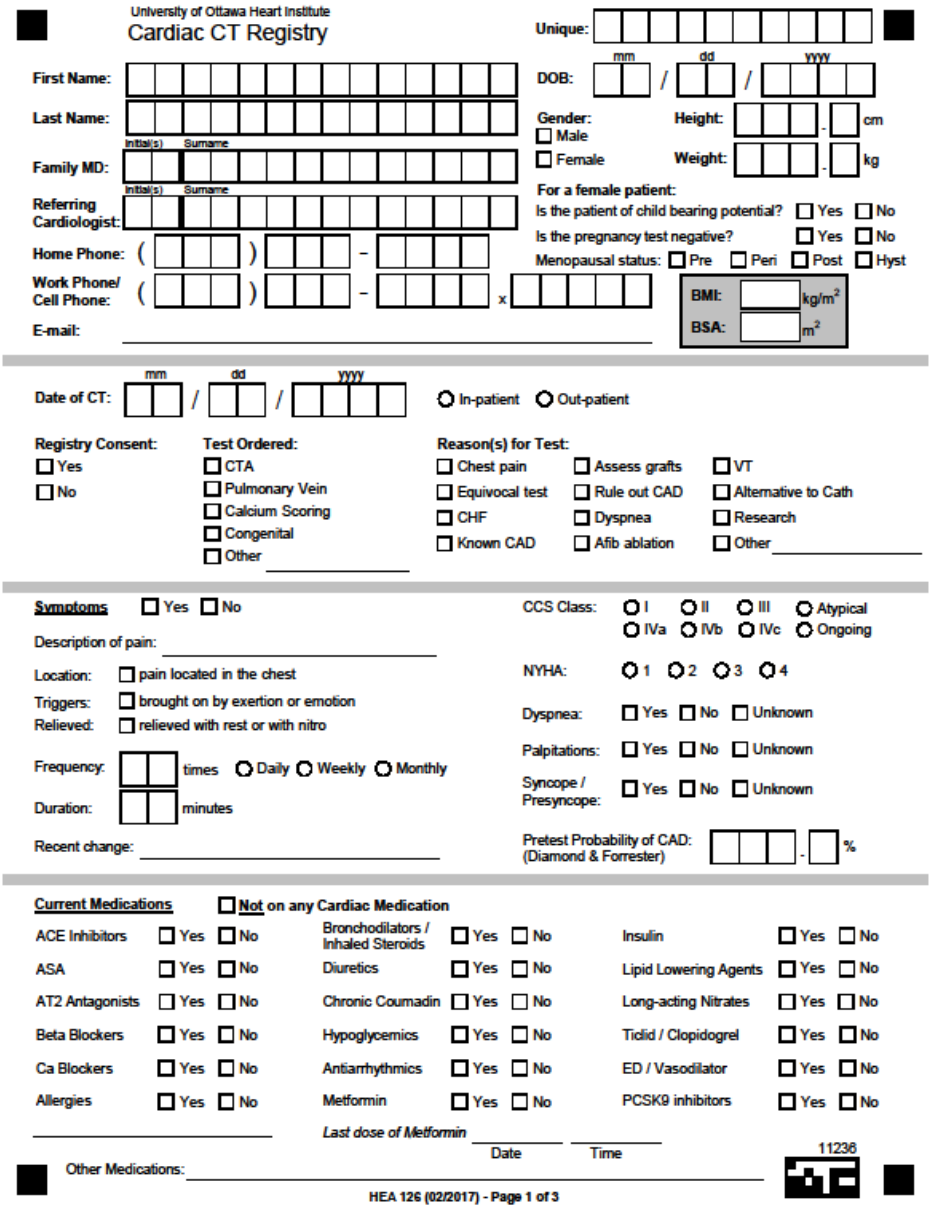
Yours sincerely,

[redacted]

Jim Rabbee, M.D. FCP
Vice-Chairperson
Ottawa Health Science Network Research Ethics Board

JR/dw

Appendix 3: Registry form for UOHI Cardiac CT database



University of Ottawa Heart Institute
Cardiac CT Registry

Unique:

First Name:
Last Name:
Family MD: Initial(s) Surname
Referring Cardiologist: Initial(s) Surname
Home Phone: ( ) -
Work Phone/ Cell Phone: ( ) - x
E-mail:

DOB: mm / dd / yyyy
Gender: Male Female
Height: . cm
Weight: . kg
For a female patient:
Is the patient of child bearing potential? Yes No
Is the pregnancy test negative? Yes No
Menopausal status: Pre Peri Post Hyst
BMI: kg/m²
BSA: m²

Date of CT: mm / dd / yyyy
In-patient Out-patient

Registry Consent: Yes No

Test Ordered: CTA; Pulmonary Vein; Calcium Scoring; Congenital; Other

Reason(s) for Test: Chest pain; Equivocal test; CHF; Known CAD; Assess grafts; Rule out CAD; Dyspnea; Afib ablation; VT; Alternative to Cath; Research; Other

Symptoms: Yes No
Description of pain:
Location: pain located in the chest
Triggers: brought on by exertion or emotion
Relieved: relieved with rest or with nitro
Frequency: times Daily Weekly Monthly
Duration: minutes
Recent change:

CCS Class: I II III Atypical; IVa IVb IVc Ongoing
NYHA: 1 2 3 4
Dyspnea: Yes No Unknown
Palpitations: Yes No Unknown
Syncope / Presyncope: Yes No Unknown
Pretest Probability of CAD: (Diamond & Forrester) . %

Current Medications: Not on any Cardiac Medication

ACE Inhibitors	Yes	No	Bronchodilators / Inhaled Steroids	Yes	No	Insulin	Yes	No
ASA	Yes	No	Diuretics	Yes	No	Lipid Lowering Agents	Yes	No
AT2 Antagonists	Yes	No	Chronic Coumadin	Yes	No	Long-acting Nitrates	Yes	No
Beta Blockers	Yes	No	Hypoglycemics	Yes	No	Ticlid / Clopidogrel	Yes	No
Ca Blockers	Yes	No	Antiarrhythmics	Yes	No	ED / Vasodilator	Yes	No
Allergies	Yes	No	Metformin	Yes	No	PCSK9 inhibitors	Yes	No

Last dose of Metformin Date Time

Other Medications:

11236

HEA 126 (02/2017) - Page 1 of 3

Unique:

Cardiac History

	Yes	No	Unknown
CAD	☐	☐	☐
Prior MI	☐	☐	☐
mm / dd / yyyy			
Prior Cath	☐	☐	☐
mm / dd / yyyy			
Prior PTCA	☐	☐	☐
mm / dd / yyyy			
Prior CABG	☐	☐	☐
mm / dd / yyyy			
Valvular Heart Disease	☐	☐	☐
Valve Repair / Replacement	☐	☐	☐
Congenital Heart Disease	☐	☐	☐
Transplant	☐	☐	☐
Congestive Heart Failure	☐	☐	☐

AICD / Pacemaker ☐ PPM ☐ AICD ☐ No

Other History: ☐ cardiac ☐ non-cardiac

_____

Risk Factors

	Yes	No	Unknown
Family Hx of CAD	☐	☐	☐
Hypertension	☐	☐	☐
Hyperlipidemia	☐	☐	☐
Diabetes Mellitus - Type I	☐	☐	☐
- Type II	☐	☐	☐

Smoking ☐ Current ☐ Ex > 1 yr ☐ Never ☐ Unknown

pk yrs Year Quit:

Other History

	Yes	No	Unknown
Peripheral Vascular Disease	☐	☐	☐
Cerebrovascular Disease	☐	☐	☐
Renal Insufficiency	☐	☐	☐
Renal Protection	☐	☐	☐

Creatinine Level umol/L eGFR

Date: mm / dd / yyyy

Lipid Profile

TC: . mmol/L TG: . mmol/L

HDL: . mmol/L LDL: . mmol/L

Date: mm / dd / yyyy

Previous tests (within last year)

	Yes	No	Unknown	mm / dd / yyyy
Stress Test	☐	☐	☐	/ /
Myocardial Perfusion Imaging	☐	☐	☐	/ /
Stress Echocardiogram	☐	☐	☐	/ /
Viability	☐	☐	☐	/ /
RNA	☐	☐	☐	/ /
Cardiac MRI	☐	☐	☐	/ /

LV Function Assessment

EF: %

Calculated with:

○ Nuclear ○ Angio ○ Echo

LV Class:

○ 1 ○ 2 ○ 3 ○ 4

Comments:

11236

Unique:

Procedure Data

RN: (Initials)

Meds given:

Baseline	HR:	BP:	/	
Pre CT	HR:	BP:	/	
PO Metoprolol: ☐ Yes ☐ No	Dose:	mg		
IV Metoprolol: ☐ Yes ☐ No	Dose:	mg		
PO Diltiazem: ☐ Yes ☐ No	Dose:	mg		
IV Diltiazem: ☐ Yes ☐ No	Dose:	mg		
PO Other: ☐ Yes ☐ No	Dose:	mg		
IV Other: ☐ Yes ☐ No	Dose:	mg		
☐ **Nitro given @** _____				
Imaging	HR:	BP:	/	
Post CT	HR:	BP:	/	

	HR	BP	Time
1			@
2			@
3			@
4			@
5			@
6			@
7			@
8			@
9			@
10			@
D/C			@

Scanner
☐ GE VCT ☐ Siemens Flash

Protocol
☐ Prospective ☐ High Pitch ☐ Retrospective

Tech: (Initials)

Slice Thickness . mm

Reconstruction
☐ Diastole ☐ Systole

Kernel
1
2
3

Contrast	cc	timing bolus	*Rate:*	cc/s
	cc	cardiac CT	*Rate:*	cc/s
Contrast / Saline Mix	cc	% contrast	*Rate:*	cc/s
Saline	cc		*Rate:*	cc/s

☐ Bolus Tracking

Type of Contrast
○ Visipaque 270
○ Visipaque 320
○ Omnipaque 300
○ Omnipaque 350
○ Other

MA/mAs:		**DLP Non-Contrast:**	
kVp:		**Timing bolus:**	
Gantry rotation:	mSec	**Cardiac:**	

Total DLP:

ECG Rhythm
☐ sinus rhythm
☐ atrial fibrillation
☐ AV block
☐ other

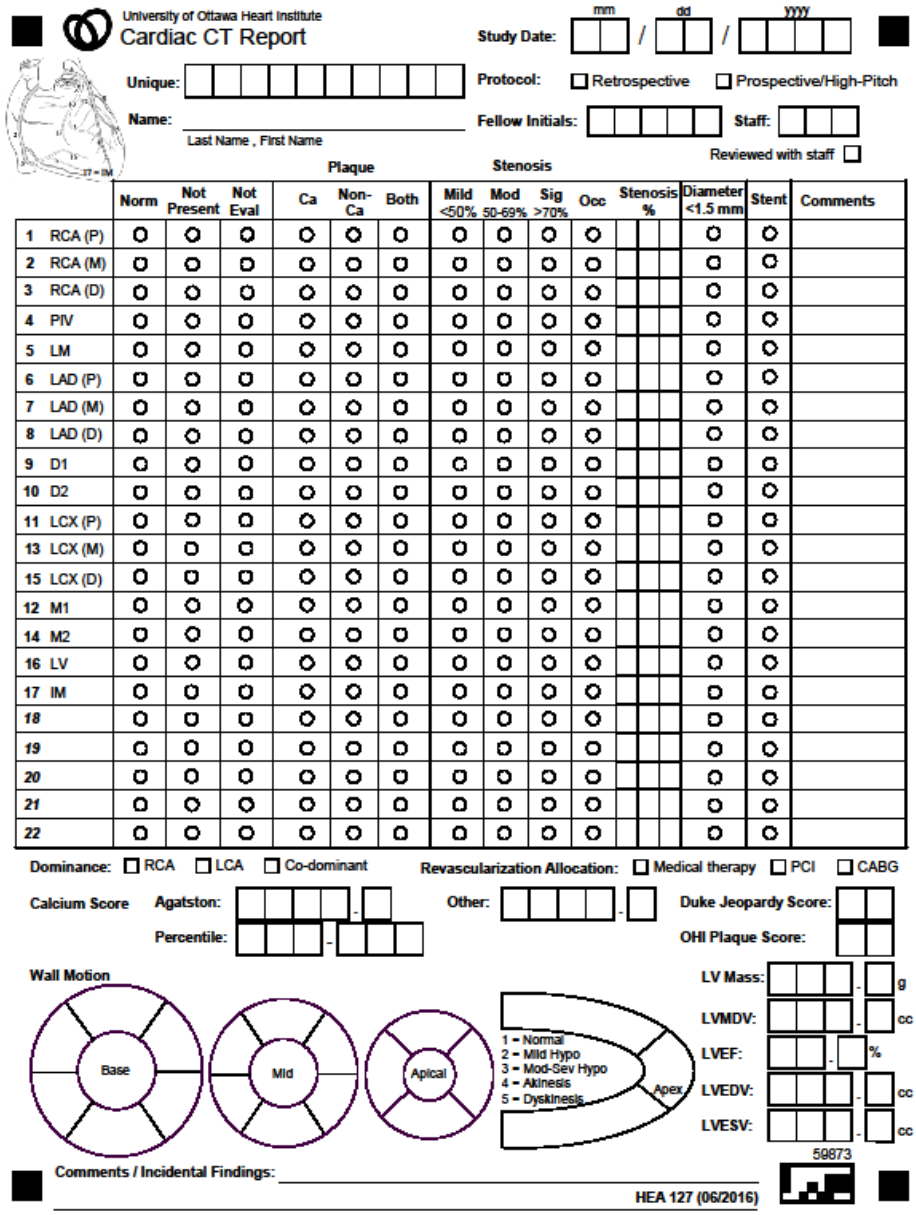
Arrhythmias
☐ none
☐ PVCs [ ] rare [ ] frequent
☐ PACs [ ] rare [ ] frequent
☐ other

ECG Interpretation
☐ normal
☐ 1st degree heart block
☐ 2nd degree AV block
☐ permanent pacer
☐ left axis deviation
☐ intraventricular delay
☐ LBBB
☐ RBBB
☐ inferior MI
☐ possible old inferior MI
☐ anterior MI
☐ possible old anterior MI
☐ non-specific ST-T wave changes
☐ LVH
☐ digoxin
☐ other

TEST CANCELLATION:
☐ no IV access
☐ atrial fibrillation
☐ ++ Ca
☐ High HR
☐ other _____

11236

Appendix 4: Report Form of UOHI Cardiac CT Registry



University of Ottawa Heart Institute
Cardiac CT Report

Study Date: mm / dd / yyyy

Unique:

Protocol: ☐ Retrospective ☐ Prospective/High-Pitch

Name: _____ (Last Name , First Name)

Fellow Initials: Staff:

Reviewed with staff ☐

	Plaque						Stenosis									
	Norm	Not Present	Not Eval	Ca	Non-Ca	Both	Mild <50%	Mod 50-69%	Sig >70%	Occ	Stenosis %			Diameter <1.5 mm	Stent	Comments
1 RCA (P)	O	O	O	O	O	O	O	O	O	O				O	O	
2 RCA (M)	O	O	O	O	O	O	O	O	O	O				O	O	
3 RCA (D)	O	O	O	O	O	O	O	O	O	O				O	O	
4 PIV	O	O	O	O	O	O	O	O	O	O				O	O	
5 LM	O	O	O	O	O	O	O	O	O	O				O	O	
6 LAD (P)	O	O	O	O	O	O	O	O	O	O				O	O	
7 LAD (M)	O	O	O	O	O	O	O	O	O	O				O	O	
8 LAD (D)	O	O	O	O	O	O	O	O	O	O				O	O	
9 D1	O	O	O	O	O	O	O	O	O	O				O	O	
10 D2	O	O	O	O	O	O	O	O	O	O				O	O	
11 LCX (P)	O	O	O	O	O	O	O	O	O	O				O	O	
13 LCX (M)	O	O	O	O	O	O	O	O	O	O				O	O	
15 LCX (D)	O	O	O	O	O	O	O	O	O	O				O	O	
12 M1	O	O	O	O	O	O	O	O	O	O				O	O	
14 M2	O	O	O	O	O	O	O	O	O	O				O	O	
16 LV	O	O	O	O	O	O	O	O	O	O				O	O	
17 IM	O	O	O	O	O	O	O	O	O	O				O	O	
18	O	O	O	O	O	O	O	O	O	O				O	O	
19	O	O	O	O	O	O	O	O	O	O				O	O	
20	O	O	O	O	O	O	O	O	O	O				O	O	
21	O	O	O	O	O	O	O	O	O	O				O	O	
22	O	O	O	O	O	O	O	O	O	O				O	O	

Dominance: ☐ RCA ☐ LCA ☐ Co-dominant

Revascularization Allocation: ☐ Medical therapy ☐ PCI ☐ CABG

Calcium Score Agatston: . Other: . Duke Jeopardy Score:

Percentile: - OHI Plaque Score:

Wall Motion

LV Mass: . g

LVMDV: . cc

LVEF: . %

LVEDV: . cc

LVESV: . cc

59873

Comments / Incidental Findings: _____

HEA 127 (06/2016)

Appendix 5: Survey for Model Predictors Pre-specification

Prediction of Obstructive CAD in Patients with Zero Calcium Score

In this study we plan to derive a clinical tool that can identify patients with obstructive coronary artery disease among patients with zero calcium score who were sent for coronary computed topographic angiography (CTA).

What are the clinical variables that you think are associated with increased probability of an underlying obstructive coronary artery disease in these patients?

Please select 10 variables from the following (you can add other variables if needed)

1. Age
2. Sex
3. Body mass index
4. Indication of the test

   RISK FACTORS
5. Family history of coronary artery disease
6. Known diagnosis with hyperlipidemia
7. Known diagnosis with hypertension
8. Known diagnosis with diabetes
9. Current smoking

   SYMPTOMS
10. Typical chest pain
11. Atypical chest pain
12. Shortness of breath
13. Palpitations

    MEDICATIONS PROFILE
14. Treatment with Aspirin
15. Treatment with beta blockers
16. Treatment with calcium channel blockers
17. Treatment with Angiotensin converting enzyme inhibitors
18. Treatment with lipid lowering agents

    PHYSICAL ASSESSMENT and BEDSIDE TESTS
19. Baseline systolic blood pressure
20. Baseline diastolic blood pressure
21. Baseline heart rate
22. Ischemic ECG findings
23. Probability of CAD

OTHERS (List below)

Appendix 6: Coronary CT Angiography Evaluation for Clinical Outcomes: An International Multicenter (CONFIRM) Registry Centers

All Centers

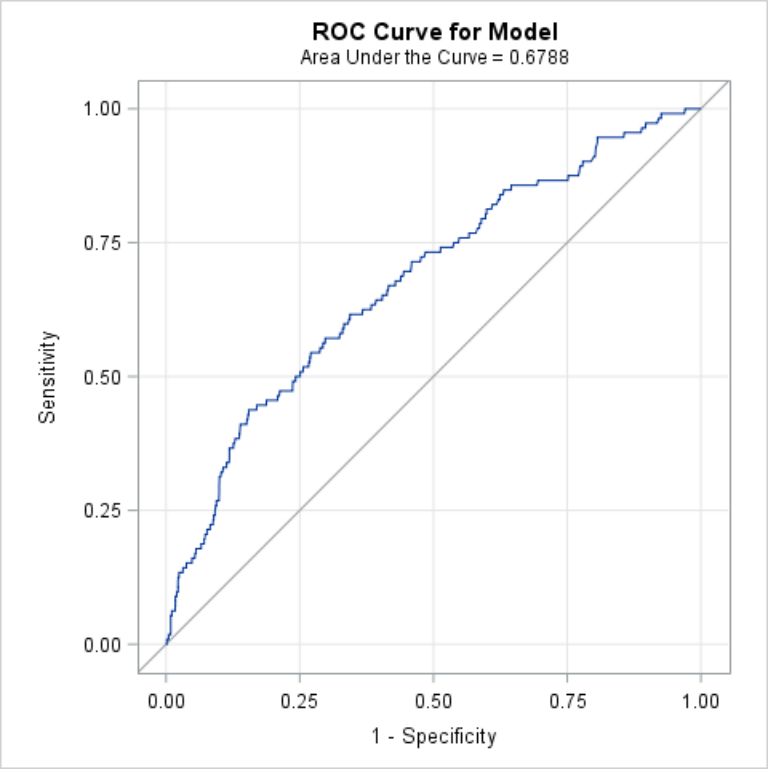
Center Number	Site	Location
1	Tennessee Heart & Cardiovascular Institute, Hendersonville, TN	North America
2*	Capital Cardiology Associates, Albany, NY	North America
3	Cedar-Sinai Medical Center, Los Angeles, CA	North America
4	University of Munich, Munich, Germany	Europe
5*	University of Ottawa Heart Institute, Ottawa, ON	North America
6*	Harbor UCLA Medical Center, Los Angeles, CA	North America
7	Hospital University Zurich, Zurich, Switzerland	Europe
8*	Severance Cardiovascular Center, Seoul, South Korea	Asia
9	Wayne State University, Henry Ford Hospital, Detroit, MI	North America
10	Walter Reed Army Medical Center, Washington, DC	North America
11	University Hospital of Parma, Parma, Italy	Europe
12	William Beaumont Hospital, Royal Oak MI	North America

*****Excluded centers

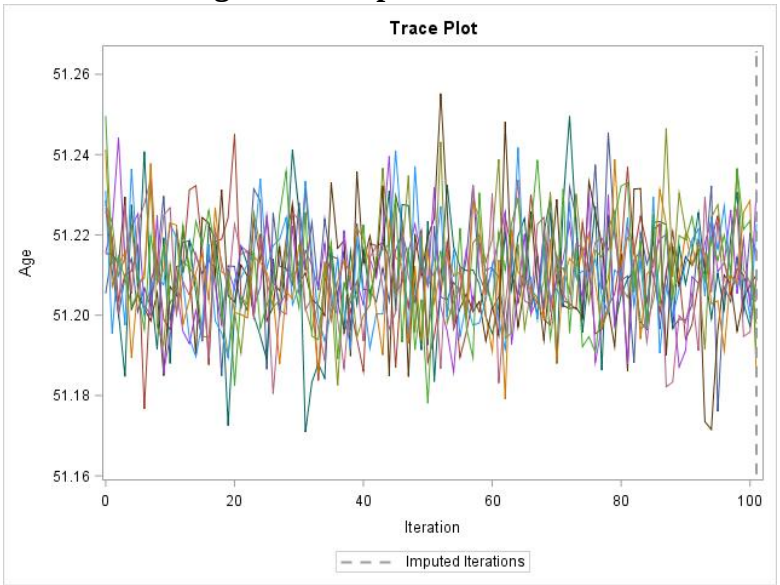
Appendix 7: Multivariable Model with Interaction Term (Full Model)

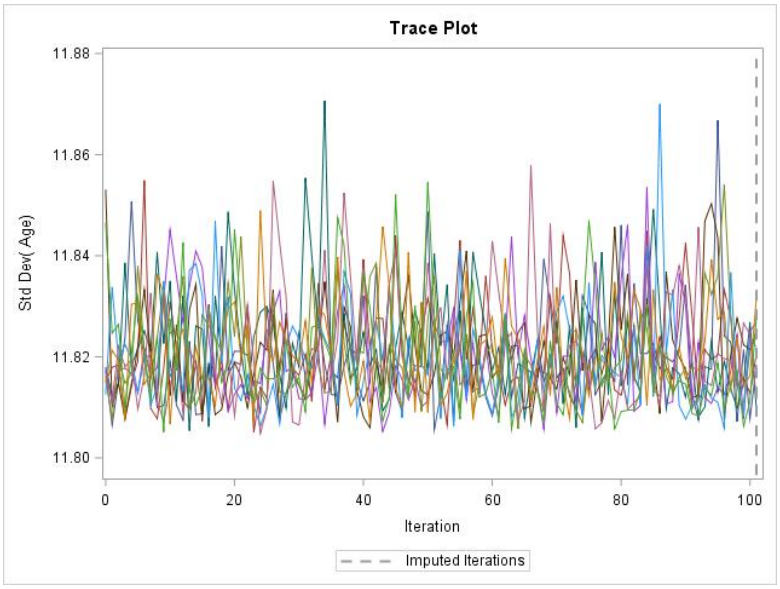
Parameter	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	-5.4373	0.6830	63.3828	<.0001
Age	0.00900	0.0104	0.7543	0.3851
Male	1.0097	0.3671	7.5640	0.0060
Typical chest pain	1.3461	0.4254	10.0122	0.0016
Hyperlipidemia	0.6313	0.2076	9.2478	0.0024
Hypertension	0.2203	0.2013	1.1976	0.2738
Family History	0.2240	0.1973	1.2889	0.2562
Diabetes	0.1585	0.2953	0.2879	0.5916
Male*typical chest pain	-0.5163	0.5734	0.8106	0.3679

AUC of ROC for the Multivariable Model



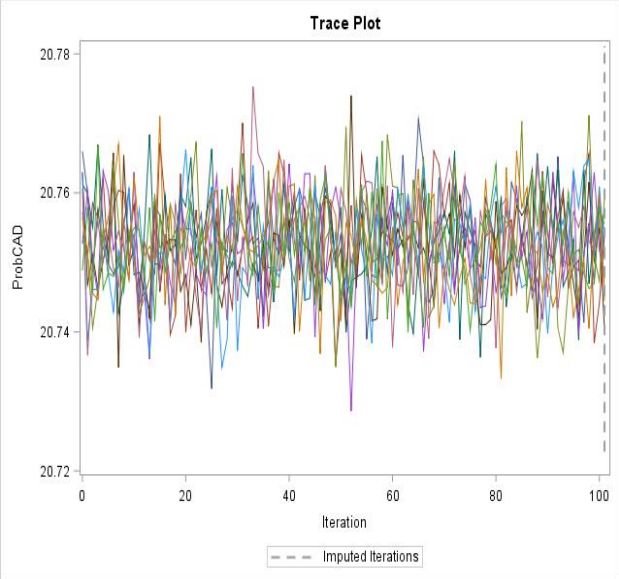
Appendix 8: Trace Plots of the Age in the Imputed Final Validation Datasets



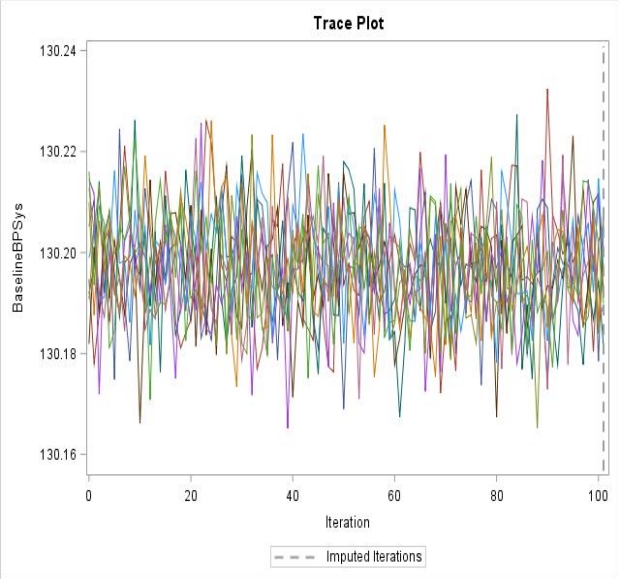


Appendix 9: Trace Plots of Continuous Variables in the Imputed Derivation Datasets (reduced model)

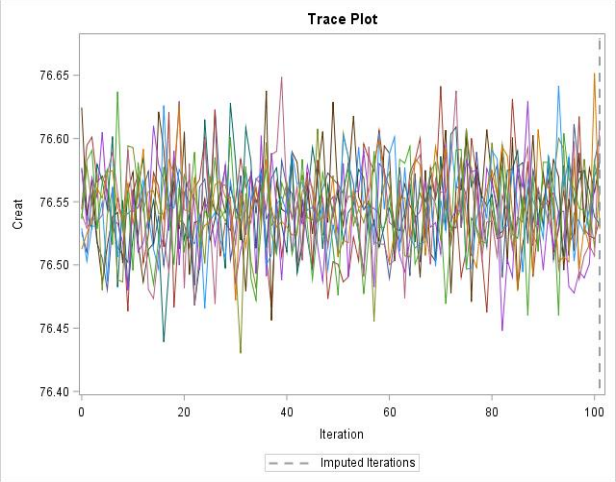
Probability of CAD



Baseline systolic blood pressure



Creatinine



Appendix 10: ROC Curve of the Developed Score from the Full Model in the Preliminary Vlidation Dataset (complete case analysis)