MONASH University

# Predicting Cardiac Surgery Outcome in an Australian Patient Cohort


**Md Nazmul Karim**

Bachelor of Medicine, Bachelor of Surgery
Master of Clinical Epidemiology
Master of Public Health


A thesis submitted for the degree of Doctor of Philosophy
March, 2017


Department of Epidemiology and Preventive Medicine
School of Public Health and Preventive Medicine
Faculty of Medicine, Nursing and Health sciences
Monash University, Melbourne, Australia

# Table of Content

# Copyright notice

# Abstract


**Background**

Cardiac surgery carries more risk of adverse outcome than many other surgical interventions. Estimation of the risk involved with surgery allows both surgeon and patient to participate in an informed manner in the decision-making process. Risk prediction models investigate surgical outcomes in relation to perioperative patient and disease characteristics to estimate coefficients for each risk factor, which are translated to risk scores. Then, the scores assigned to each risk factor are added to calculate the overall risk score for a patient and to construct clinical risk groups.

Cardiac surgery is among the most dynamic fields of medicine. Advances in surgical and postoperative care approaches allow more patients with co-morbidities, previous operation history and of extreme age are to be eligible for surgery. This change in the risk profile of patient populations may restrict general applicability or optimal performance of currently available models. To cope with the contemporary clinical practice, identification of the pattern and predictors of cardiac surgery outcomes in patient with altered risk profile is indispensable. Current research aimed at studying the aspects of development of risk prediction models to improve cardiac surgery outcome assessment.

**Research approach**

The first approach taken to pursue the aim was to identify knowledge gap and research need in cardiac surgery risk prediction modelling through review of existing models. Findings of the review were then used to guide subsequent research. The research used information of 84,233 patient, from the ANZSCTS (Australia and New Zealand Society of Cardiac and Thoracic Surgeons Registry) database, who underwent cardiac surgery between 2001 and 2014. The research investigated the impact of variable misclassification, missing value and different variable selection methods on the performance of the risk prediction models. After studying all these gaps identified by the review, the knowledge gained were applied to develop novel

models for predicting long-term survival following Coronary Artery Bypass Grafting (CABG) surgery.

**Key findings**

The systematic review identified wide variation in the development methodology of the risk prediction models. Ambiguous predictors and outcome definition, sub-optimum sample size, inappropriate handling of missing data and inefficient predictor selection technique were the key issues prevalent among the contemporary models. Misclassification of patients to 'urgent' category of surgery in the ANZSCTS database was high and this misclassification results in overestimation of mortality risk. This study proposes a new definition of 'urgent' clinical status to prevent future misclassification. Investigation of missing values shows that, multiple imputation of missing values during model development increases the precision and performance of the risk prediction models. Clinical suitability in terms of parsimony and prediction performance can best be achieved using bootstrap bagging technique for the development of risk prediction models. A set of novel risk prediction models for predicting long-term survival at four distinct time intervals (31-90 days, 91-365 days, 1-3 years and > 3 years) following CABG surgery was developed.

**Conclusion**

This research has provided new knowledge about the existing practice in the risk prediction modelling for cardiac surgery patients and provided a range of evidence based suggestion regarding model development practices for improving outcome assessment following cardiac surgery. The research also provided a set of novel risk prediction models for predicting long-term survival at four distinct time intervals following CABG surgery. These models along with the existing short-term mortality model will provide surgeons and patients greater confidence in surgical decision making.

# General declaration

**Declaration for the thesis based or partially based on conjointly published or unpublished work**

In accordance with Monash University Doctorate Regulation 17.2 Doctor of philosophy and Research Master's regulations the following declaration are made:

I hereby declare that this thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

This thesis includes three published, one accepted and one submitted peer reviewed journals publications. The core theme of the thesis is 'Predicting Cardiac Surgery Outcome in Australian Patient Cohort'.

The ideas, development and writing up of all the papers in the thesis were the principal responsibility of myself, the student, working within the 'Department of Epidemiology and Preventive Medicine' under the supervision of Dr Baki Billah, Professor Christopher Reid, A/Prof Andrew Cochrane and Dr Lavinia Tran.

The inclusion of co-authors reflects the fact that the work came from active collaboration between researchers and acknowledges input into team-based research.

In the case of chapters 3, 4, 5, 6 and 7 my contribution to the work involved the following:

| Thesis Chapter | Publication Title | Status | Nature and % of student contribution | Co-author name(s) Nature and % of Co-author's contribution* | Co-author, Monash student Y/N* |
|---|---|---|---|---|---|
| 3 | Mortality Risk Prediction Models for Coronary Artery Bypass Graft: Current scenario and future direction | Accepted | 70%, Study concept, data analysis, Interpretation of result and manuscript writing | 1. Christopher M Reid, Supervision & Input in to manuscript - 7%<br>2. Andrew Cochran Supervision & Input in to manuscript - 5%<br>3. Lavinia Tran Supervision - 3%<br>4. Md Alramadan Data extraction - 3%<br>5. Md N Hossain Data extraction - 2%<br>6. Baki Bilhah Supervision & Input in to manuscript - 10% | No<br>No<br>No<br>Yes<br>Yes<br>No |
| 4 | When is 'Urgent' Really Urgent and Does It Matter? Misclassification of Procedural Status and Implications for Risk Assessment in Cardiac Surgery. | Published | 75 %, Study concept, data analysis, Interpretation of result and manuscript writing | 1. Christopher M Reid, Supervision & Input in to manuscript - 7%<br>2. Andrew Cochran Supervision & Input in to manuscript - 5%<br>3. Lavinia Tran Supervision & Input in to manuscript - 3%<br>4. Baki Bilhah Supervision & Input in to manuscript - 10% | No<br>No<br>No<br>No |
| 5 | Missing Value imputation improves mortality risk prediction following cardiac | Published | 75%, Study concept, data analysis, Interpretation | 1.Christopher M Reid, Supervision & Input in to manuscript - 7%<br>2. Andrew Cochran | No<br>No |

| Thesis Chapter | Publication Title | Status | Nature and % of student contribution | Co-author name(s) Nature and % of Co-author's contribution* | Co-author, Monash student Y/N* |
|---|---|---|---|---|---|
| | surgery: an investigation of an Australian patient cohort | | of result and manuscript writing | Supervision & Input in to manuscript - 5%<br><br>3. Lavinia Tran<br>Supervision - 3%<br><br>4. Baki Bilhah<br>Supervision & Input in to manuscript - 10% | <br><br><br>No<br><br>No |
| 6 | Variable selection methods for multiple regressions influence the parsimony of risk prediction models for cardiac surgery | Published (online ahead of print) | 75%, Study concept, data analysis, Interpretation of result and manuscript writing | 1.Christopher M Reid, Supervision & Input in to manuscript - 7%<br><br>2. Andrew Cochran<br>Supervision & Input in to manuscript - 5%<br><br>3. Lavinia Tran<br>Supervision - 3%<br><br>4. Baki Bilhah<br>Supervision & Input in to manuscript - 10% | No<br><br>No<br><br><br><br>No<br><br>No |
| 7 | Predicting long-term survival after coronary artery bypass grafting surgery | Submitted under editorial review | 70%, Study concept, data analysis, Interpretation of result and manuscript writing | Christopher Reid, Supervision & Input in to manuscript - 7%<br><br>Andrew Cochran<br>Supervision & Input in to manuscript - 5%<br><br>Lavinia Tran<br>Supervision - 3%<br><br>Samuel L Brilleman, Data analysis and Input in to manuscript - 5%<br><br>Baki Bilhah<br>Supervision & Input in to manuscript - 10% | No<br><br>No<br><br><br><br>No<br><br>Yes<br><br><br><br>No |

I have not renumbered sections of submitted or published papers in order to generate a consistent presentation within the thesis.

Student signature: [redacted] Date: 23 /03/2017

The undersigned hereby certify that the above declaration correctly reflects the nature and extent of the student's and co-authors' contributions to this work.

Main Supervisor signature: [redacted] Date: 23 /03/2017

# Acknowledgements

It's my pleasure to thank the many kind people around me without the support of whom writing this thesis would not have been possible, to only some of whom it is possible to give particular mention here.

I wish to firstly acknowledge the immeasurable support and encouragement from my family, friends and colleagues. I am forever grateful to my wife Lisa, and our two children, Rydha and Rahid, for their love, ongoing support and patience especially during the difficult times.

I would like to thank my amazing supervisors, Baki Billah, Christopher Reid, Andrew Cochrane and Lavinia Tran. Throughout my PhD candidature, supervisors have provided encouragement, sound advice and a lot of good ideas. Their guidance, patience and enthusiasm never ceases to amaze me and their speed of returning my draft and answering my never-ending questions are truly incredible. I am forever indebted to them for all they have done to help me achieve this goal. Baki especially has not only been an incredible primary supervisor, he stayed with me through my personal ups and downs, always setting me back on course. My co-authors have been fantastic and none of the words on the following pages could have been completed without their directive and assistance.

I would like to acknowledge the financial, academic and technical support of the Monash university, Australia and its staff, particularly in the award of International Postgraduate Research Scholarship and Monash Graduate Scholarship that provided necessary financial support for the thesis.

I would like to thank Professor Andrew Forbes, Professor Rory Wolfe and A/ Professor Stephane Heritier for their statistical insight and time. I would like to thank Kathlyn Ronaldson, Sam Brilleman and Rakibul Islam for helping proof-read my thesis and keeping my formatting in check. Thank you to Sam Brilleman for all his help with exceptional skill with the statistical software.

I must of course thank the ANZSCTS Cardiac Surgery Database program and the team for providing me the access to the enormous database.

My special thanks to Elizabeth Douglas, the doctoral coordinator at the School of Public Health and Preventive Medicine, who has been great source of support and inspiration for me.

# Dedication

*For my Parents*

*In loving memory*

# List of abbreviations and acronyms

| | |
|---|---|
| AIC | Akaike information criterion |
| ANN | Artificial neural network |
| ANZCTS | The Australian & New Zealand Society of Cardiac & Thoracic Surgeons |
| AUC | Area under ROC curve |
| BIC | Bayesian information criterion |
| BMI | Body mass index |
| BSA | Body surface area |
| CABG | Coronary artery bypass grafting |
| CBVD | Cerebrovascular disease |
| CCRE-T | Cardiovascular research and education in therapeutics |
| CCS | Canadian cardiovascular society classification |
| CHARMS | Checklist for Critical Appraisal for Reviews of Prediction Modelling Studies |
| CHF | Congestive heart failure |
| DOSA | Elective day of surgery admit |
| EF | Ejection fraction |
| eGFR | Estimated glomerular filtration rate |
| EPV | Event per variable |
| EuroSCORE | European system for cardiac operative risk evaluation |
| FHCAD | Family history of coronary artery disease |
| ICE | Imputation by chained equations |
| ICU | Intensive care unit |
| IE | Infective endocarditis |
| MAR | Missing at random |
| MCAR | Missing completely at random |
| MedCalc | MedCalc statistical software |
| MI | Myocardial infarction |
| MI | Multiple Imputation |
| MICE | Multiple imputation by chained equations |

| | |
|---|---|
| MNAR | Missing not at random |
| MORT30 | Mortality within 30 days of surgery |
| MUHREC | Monash university human research ethics committee |
| NDI | National Death Index |
| NYHA | New York Heart Association |
| PRECR | Preoperative creatinine level |
| PVD | Peripheral vascular disease |
| R | The R project for statistical computing |
| RMS | Regression modelling strategies package |
| ROC curve | Receiver operating characteristic curve |
| SCERH | Standing committee on ethics in research involving humans |
| SCTS | The society for cardiothoracic surgery of Great Britain and Ireland |
| Stata | Stata: data analysis and statistical software |
| STS | The society of thoracic surgeons |
| PBN | Prognostic Bayesian Networks |
| MCRS | Mayo Clinic Risk Score |
| PCI | Percutaneous Coronary intervention |
| THIRST | Texas Heart Institute Risk Scoring Technique |
| ACC | American College of Cardiology |
| AHA | American Heart Association |
| RAMR | Risk Adjusted Mortality Rate |
| MSE | Mean standard error |
| H-L test | Hosmer-Lemeshow test |

# Chapter 1: Introduction

## 1.1 Risk prediction in cardiac surgery

Cardiac surgery carries a higher degree of risk of adverse outcome than many other surgical interventions. Predicting outcomes in adult cardiac surgery is critical for decision-making purposes, particularly when there are different treatments options available (1, 2). Risk prediction allows trade-off between risks and benefits and facilitates evidenced based surgical decision making (3).

Risk prediction models in cardiac surgery investigate surgical outcomes in relation to peri-operative patient and disease characteristics to estimate coefficients for each risk factor, which are translated to risk scores. Then, the scores assigned to each risk factor are added to calculate the overall risk score for a patient and to construct clinical risk groups.

Over the past decades, the field of cardiac surgery has made considerable progress in the development of risk prediction models to enable outcome prediction and clinical quality monitoring. Risk prediction models for postoperative outcomes have become an integral part of cardiac surgical risk assessment. Health authorities, hospitals, medical practitioners are increasingly placing importance in risk prediction models, to obtain objective risk-adjusted prediction of mortality after cardiac surgery (4). National cardiac surgical registries have been established in many countries and many have developed risk prediction models suitable for local populations (5, 6).

### 1.1.1 Application of risk prediction models

Risk prediction models are primarily used as a decision support tool for clinicians. Estimation of the risk involved assists them in patient selection and in the choice of treatment strategy (7) and allows both surgeon and patient to participate in an informed manner in this process (2, 8). Surgeons can decide the most appropriate treatment plan for a specific patient by considering predicted risk in addition to their clinical assessment (3, 9). Prediction can be used to educate and counsel patients about the risk associated with the surgery (8) and thus to facilitate evidence based informed consent.

Risk prediction models are used as a bed-side tool for estimating risk of individual patients prior to any surgical procedure and can be applied in the assessment of the relative impact of specific risk

factors on surgical outcomes (10). Some of the models are available for use in the form of online calculators that makes them accessible to wide range of setting and population.

Quality assessment is an important component of evidence based approach to patient care. Risk prediction is essential for surgical quality assessment and improvement of surgical outcomes. Risk prediction models can be used for hospital and physician benchmarking which can facilitate comparison of provider performance. Results of individual physicians or hospitals can be compared with results from others to provide a point of reference. Risk-adjusted outcomes can be used as the basis for monitoring of performance.

The allocation of health care resources is another application of these models. Prediction of postoperative complication, length of intensive care unit (ICU) stay and length of hospital stay may allow efficient allocation of resources (11). Prediction from these models could serve as a basis for planning the optimal schedule for cardiac surgery (12, 13). Scores generated from the risk prediction models can be used for academic research involving estimation of the effect of risk factors or therapies on patient.

### 1.1.2 Currently used models

Several studies mostly by anaesthesiologists in the early 1980s have attempted to identify predictive factors for mortality following cardiac surgery (14-16). Paiement et al (17) first reported a scoring system for cardiac surgery patients based on the presence of risk factors associated with adverse outcome. The scoring system—a simple method of classifying patients before surgery—was considered a reliable method of identifying patients at increased risk of perioperative mortality (18) and was used routinely by the anaesthetists. The Parsonnet system (19) developed in 1989, was the first widely accepted model. The original score was later modified in 1994 and is known as the 'modified Parsonnet score' (20).

Over the past couple of decades, risk prediction models have been devised, around the world, for predicting risk of adverse outcomes following cardiac surgery. The Society of Cardiothoracic Surgeons of Great Britain and Ireland (SCTS) developed the UK society score (21, 22) for coronary artery bypass grafting (CABG) surgery patients in the United Kingdom (1). They later developed a 'complex Bayes model' with 9 risk factors and a 'simple Bayes model' with 5 risk factors. The 'European System for Cardiac Operative Risk Evaluation (EuroSCORE)' was developed in 1995 for predicting early mortality in cardiac surgical patients (6). The scoring system has widely been used for prediction of immediate

death after adult cardiac surgery. In 2012 an updated version of the model—EuroSCORE II— was released (23). The Society of Thoracic Surgeons (STS) risk models were developed in 1999 and have undergone periodic revisions (5). Predictive performance of the STS algorithms is in general comparable with other systems and remains the most widely used model in the United States. Both STS and EuroSCORE II developed online risk calculators.

In Australia, the EuroSCORE model was most widely used until Yap and colleagues showed that it performs poorly for the Australian cohort (24). Several prediction models were developed for Australian population using the Australian and New Zealand Society of Cardiac and Thoracic Surgeons (ANZSCTS) registry data (25-28). The Australian System for Cardiac Operative Risk Evaluation (AusSCORE) was developed in 2009 for predicting 30-day mortality risk following isolated coronary artery bypass graft (CABG) surgery (25), and its updated version— AusSCORE II— was published in 2014 (26). A global model for predicting 30-day mortality was developed in 2010 (27) and a similar model was developed for patient undergoing Aortic valve replacement surgery in 2011 (28). A complete list of models developed from ANZSCTS registry data are presented in Appendix 1.

## 1.2 Emerging issues in risk prediction modelling in cardiac surgery

### 1.2.1 Changing patient population

The landscape in which surgeons perform cardiac surgery is changing drastically. This evolution has been driven by advances in interventional cardiology and minimally invasive cardiac procedures. Newer technologies are meant to increase the efficiency of surgery and quality of care. Therefore, larger number of elderly patients, those with greater burden of comorbid illness, with concomitant valvular disease and with history of previous surgery are pooled for surgery (11). Thus, the risk profile of patients presenting for cardiac surgery have significantly changed now-a-days (29).

### 1.2.2 Ambiguous predictor variable definition

While developing risk prediction models, strict standardization of definitions for predictor variables and for endpoints must be ensured (3). Inaccurate or ambiguous definition makes the predictor variables prone to misclassification (30). A misclassified variable may create opportunity for gaming as benchmarking is particularly more sensitive to misclassification of key predictors (31). Due to misclassification, the concerning predictor will lose its ability to predict a patient's risk precisely, which, in turn, will decrease model performance (31).

There has been widespread heterogeneity among the currently used models in defining the predictors. Diverse definitions and measurement methods of predictors are a potential source of misclassification. For example, misclassification of an important predictor–clinical urgency—was found in the ANZSCTS database (32). Definition of renal dysfunction was found to affect EuroSCORE performance. Renal impairment was defined as Creatinine >200 mmol/L in the EuroSCORE. When the model was recalibrated redefining the renal dysfunction with creatinine as a continuous variable or glomerular filtration rate as a categorical, the predictive accuracy of the EuroSCORE model for hospital mortality increased significantly (33).

Misclassification irrespective of intent should be addressed appropriately. The best method to manage misclassification is obviously to avoid it. However, it's not possible to eradicate. Accuracy of definitions should always be assessed and updated accordingly (6, 34).

### 1.2.3 Choice between preoperative and intraoperative predictors

Preoperative predictors are popularly being used for predicting risk of adverse outcome following cardiac surgery. Although inclusion of intraoperative and/or postoperative characteristics in the model might improve the prediction, their inclusion should be judged based on aim of the model. As intraoperative data are not available prior to surgery, surgeons have no alternative of resorting to preoperative characteristics to foresee prognosis for patient counselling and surgical decision making. Inclusion of intraoperative and/or postoperative characteristics would be essential in a model that intends to compare surgical performances between surgeons or hospitals (35).

### 1.2.4 Varying endpoint of interest for the prediction

#### 1.2.4.1 Variation in outcome definition

Mortality as an endpoint is widely used in the cardiac surgery risk prediction models. Morbidity, resource utilization, costs and patient satisfaction are among the other endpoints being used. The advantages of mortality as an endpoint include: as a hard event, it poses little ambiguity; and data can validly be obtained from a range of sources. However, there is still room for uncertainty about the timing of death. Commonly used mortality endpoints focus primarily on short-term mortality which include operative death, in-hospital mortality or 30-day mortality (patients who die within 30 days of the surgery).

Small difference in reported deaths due to different timing of death incorporated in definitions may affect model prediction (36). It is likely that the number of deaths will differ between 'in-hospital mortality' and '30-day mortality'. Many patients may die after 30 days of surgery because of late complication or comorbidity. Their death will not be captured if '30-day mortality' is used as outcome. These deaths may be counted as the 'in-hospital mortality' only if the patient is still hospitalized. In this way, some models (5, 12) are likely to overestimate operative mortality by using 'in-hospital mortality' as outcome. 'In-hospital mortality' data are relatively easier to collect; however, duration of hospital stay may sometimes be due to institutional habits concerning postoperative patient care. If a hospital discharges the patients earlier an underestimation of hospital mortality is likely.

#### 1.2.4.2 Long term mortality

Although popularly being used, there are concerns that short-term mortality is probably not by itself an adequate indicator of cardiac surgery performance (8). Short-term mortality is predominantly used to evaluate the early risk of surgical procedures (37). Short-term mortality does not capture patient satisfaction, quality of care and length of survival following the surgery (38). Advancement in surgical techniques and post-operative patient management has increasingly delayed the death among the postoperative patients even among those with critical condition (39). Deaths among such patients may be delayed substantially after the surgery. Delayed postoperative deaths may cause an underestimation of the operative death if the definition is based on 30-day mortality (39).

Further, due to advancements in surgical technologies and perioperative care, operative and 30-day mortality rates have declined over the last few decades. Hence more attention is now required towards improving long-term survival following cardiac surgery to get a complete prognosis (40). Prediction of long-term survival can be used to determine the most appropriate post-discharge care strategies. This may essentially help patients and their doctors to implement behavioural and therapeutic modifications to optimize benefit from surgery (40).

#### 1.2.4.3 Paradigm shift to morbidity

Some argue that emphasis on mortality as the only endpoint, may engender negative behaviours such as high-risk case avoidance by surgeons and institutions. Which in turn may reduce access to surgery for people with elevated risk of mortality who might benefit the most (41, 42). For similar reasons, such models may encourage gaming of the reporting system when used for benchmarking and

comparison of performance among surgeons or institutions (43). Morbidity rates are increasingly becoming an important way to describe procedural outcome (12). Some consider morbidity as a more suitable endpoint for analysis because a patient is more likely to have a deterioration in health or recover unusually slowly from surgery than to die.

Further, morbidity may correlate better than mortality with admission to the ICU, length of hospital stays, return to work, quality of life, and most importantly costs. However, morbidity data are difficult to collect, and there is problem with standardization of morbidity definitions. For example, renal dysfunction may be defined as anuria, the need for dialysis, or elevation in serum creatinine levels above a preoperative baseline value. Therefore, the development of models to predict morbidity in the field is yet to succeed comprehensively (44-46). Because of the heterogeneity of morbidity events, future scoring systems should probably develop models for mortality and major morbidity events (eg. stroke, myocardial infraction, renal impairment, arrhythmia etc.) separately.

### 1.2.5 Choice between the procedure specific model and an all-procedures model

Many models serve as a general cardiac surgery risk prediction model (19, 23, 27, 47-49), one model-fit-for-all procedure type. Some other models (5, 25, 26) are intended to be used for specific cardiac surgery (eg. CABG, valve surgery etc.). While within the cardiac surgery population, there is a wide variety of procedures with different determinants of mortality. Gameren and colleagues (50) showed that dedicated risk models for specific surgery type may be useful to provide more valid estimates of mortality after surgery.

Further exploration is needed to clarify how specific should be the procedure type that requires a separate model. The question persists whether a separate model is required for specific type or subtype of surgery. For example, in case of valve surgery, should there be a separate model for each of valve types (e.g. aortic valve surgery, mitral valve surgery etc.)

### 1.2.6 Advances in model development technique

#### 1.2.6.1 Missing value imputation

Risk prediction models are usually developed using data routinely collected in hospitals or general practices or by registries. These data contain information on predictors based on patient characteristics. Irrespective of the design and diligence of those involved in the data collection

process, missing data is common in these databases (51). A common approach to dealing with missing data in model development is to drop cases with incomplete data from analysis. In multivariable analysis—commonly used in model development—case deletion often results in a large portion of the data being discarded and can result in substantially smaller sample sizes (52, 53). Along with wastage of valuable information, collected with cost and effort, this analysis with only the complete cases may lead to generation of biased estimates of parameters in the prediction model (53). Multiple imputation is a statistical technique for analysing such data and has become popular because of recent software development (54). Despite recommendation for imputing missing values prior to risk prediction modelling (55), currently used risk prediction models do not seem to adequately address the impact of missing data.

#### 1.2.6.2 Adjustment for surgeon and hospital factor

Factors with the potential to affect patient outcome, are not necessarily restricted to preoperative patient characteristics. Some of exogenous factors (not related to patients) may also affect the outcome of surgery. These Include variables related to the skill and experience of the surgeon (surgeon factor) and postoperative care teams, which in turn influence various aspects of the intra-operative and immediate postoperative period (hospital factor) (9). Most widely used models are developed using large multicentre registry data. Along with differences in institutional practice patterns, hospitals also may vary the profiles of their patients in relation to socioeconomic status, education, compliance, diet and even severity of illness. Therefore, it is not appropriate to assess the quality of care by measuring crude procedural mortality alone. It emphasizes that comparisons of operative mortality rates among centres are meaningless without risk adjustments derived from case-mix (3). All these variations have the potential to bias the estimate unless they are being adjusted for using robust statistical techniques (9). Statistical techniques like multilevel modelling (56) or structural equation modelling (57) can be used to adjust the issues while developing the model.

#### 1.2.6.3 Robust variable selection methods

The trade-off between parsimony and performance is a major challenge in risk prediction modelling (9). A parsimonious model is computationally simpler with relatively smaller number of predictors for the clinician to implement in day-to-day practice (58). Limiting the number of predictors in the model is an important way to achieve parsimony (59). On the contrary, omitting important prognostic factors has the potential to result in inaccurate prediction (59).

Many earlier models used clinical acumen and/or univariate association with outcome to choose predictors for the model. Statistical variable selection techniques are now-a-days popularly used for choosing most suitable predictors (60, 61). Bayesian-algorithm (62), Machine learning algorithms (63, 64) and multivariable regression (65) are among the most commonly used techniques. These methods excel in different tasks and have their inherent limitations. There is no consensus about the most suitable method for model development.

## 1.3 Rationale

Although the existing practice of risk prediction modelling is the result of decades of research, there is still room for improvement and updates are required to reflect current clinical practice. Cardiac surgery is among the most dynamic fields of medicine. Newer technologies are constantly emerging. Incessant advances resulting in innovative approaches and improved outcomes. More patients with co-morbidities, previous operation history and of extreme age are made eligible for surgery (66). The changes in patient's risk profile have the potential to affect the predictive accuracy and applicability of currently available models (67). Therefore, models should be recalibrated and updated to adapt the changes in the patient population and their risk profile. Further, risk models are nowadays expected to accurately predict administrative outcomes like length of hospital or ICU stay, cost of care and hospital resource needs in relation to the optimal schedule for cardiac surgery (12). Risk modelling requires expansion to incorporate all these diverse aims and to adapt with contemporary clinical practice. Emergence of robust statistical methods for model development (multiple imputation of missing value, multilevel modelling, bootstrap bagging, neural networks etc.) also necessitate the refurbishment of models to optimize prediction. A review by Nilsson (68) showed that the predictive performance of older models is usually poorer compared with more recent ones. Risk modelling requires refurbishment and upgradation to maximize prediction performances (35).

## 1.4 Aim of the research

### 1.4.1 Overall aim

Overall aim of the research was to study aspects of the development of risk prediction models for short and long-term mortality to improve cardiac surgery outcome assessment

### 1.4.2 Specific objective

1. To critically appraise the methods used by existing risk prediction models for patients undergoing coronary artery bypass grafting surgery
2. To study impact of procedural status misclassification on the performance of the risk prediction model
3. To study impact of missing values on the performance of the risk prediction model
4. To study impact of variable selection methods on parsimony of the risk prediction model
5. To develop a model for predicting long-term survival following coronary artery bypass grafting surgery

## 1.5 Overview of the thesis

The PhD research is presented as a thesis by publication consisting eight chapters. Chapter two provides the description of the data source for the research and an overview of the research methodology and theoretical frame works applied in the research. A description of the ANZSCTS national cardiac surgery database includes the structure and the management process of the database registry and definitions of the key variables relevant to this thesis. Detailed methodology is described in the respective chapters.

Chapter three includes a paper on systematic review of the existing risk prediction models, for CABG surgery patients, currently being used around the globe. The paper presented an overview of the methodology used in the development of the risk prediction models and outlined the gaps in knowledge and practice in contemporary prediction modelling.

Chapter four, five and six include three papers on three methodological issues identified by the systematic review in chapter three. Chapter four includes a paper that studied registry data quality

aspect of risk prediction modelling and assessed impact of clinical status misclassification on predicted mortality risk following cardiac surgery. Chapter five includes a paper that assessed impact of missing values on risk prediction models' performance. Chapter six includes a paper that compared the variable selection methods for multiple regression to assess its influence on the parsimony of risk prediction models for cardiac surgery.

Chapter seven includes a paper presenting a set of novel risk prediction models for predicting long-term survival following CABG surgery. The research presented in chapter seven addressed all the issues identified in contemporary prediction modelling and incorporated all the knowledge gathered in the researches in the previous chapters.

Chapter eight concludes the key findings of the thesis, their relevance and implication in the field of risk prediction modelling in cardiac surgery and future direction. The chapter also presented the strength and limitation of the thesis.

# Chapter 2: General methodology

## 2.1 The Australian and New Zealand Society of Cardiac and Thoracic Surgery (ANZSCTS) database program

The Australian and New Zealand Society of Cardiac and Thoracic Surgeons (ANZSCTS) database program (registry) was established in 2001 with the aim of reporting risk-adjusted clinical outcomes for patients undergoing cardiac surgery. Fundamental to the process of establishing the registry was the development and agreement of a standard dataset and definitions to be used by all hospitals participating in the program (69). The identification of key performance indicators and subsequent generation of local standards afforded the ability to benchmark individual and unit performance across Australia (70). The program is funded by the Department of Health, Victoria, the Health Administration Corporation (GMCT) and the Clinical Excellence Commission (CEC), NSW, and funding from the individual participating hospitals.

The objective of the program was to provide:

- A common dataset with identical definitions of all data points
- Collation of reliable data for research, risk assessment and outcome prediction
- A system for evaluation of individual, unit, hospital, state and national performance
- Improvement of the quality of patient care through an effective peer review mechanism
- A core dataset for in-house mortality and morbidity review at Institutional level
- Appropriate information for external research purposes

The database contains information on all patients who have had cardiac surgery at one of the participating centres, since 1 July 2001 (or during the period of the participating hospital's involvement if this began later than 1 July 2001). Currently, 23 public hospitals out of 24 and eight of the private hospitals around Australia are contributing data on surgical procedures into the registry. Geographical distribution of the centres across Australia is presented at the Appendix 2.1

The registry captures all adult cardiac surgical procedures, performed in participating hospitals, including coronary artery bypass grafting (CABG) and valve procedures. The database consists of more than 300 preoperative, intra-operative and post-operative variables. Data elements were defined and adapted from internationally standardized data definitions (71). Preoperative variables included patient demography, risk factors, preoperative cardiac status, history of previous intervention and

preoperative hemodynamic state of the patients. Intra-operative variables included the procedure, cardiopulmonary bypass and support, and procedure specific information. Post-operative variables included information on post-operative support, complication, readmission and outcome. Standard data is collected on the paper form. The ANZSCTS data collection form is presented at the Appendix 2.2

Data are entered and transmitted through a secured online web-based system. Data management, analysis, and database development are maintained by Centre of Cardiovascular Research and Education in Therapeutics (CCRE-T), in the School of Public Health and Preventive Medicine, Monash University. Each hospital has a designated data manager who is responsible for the completeness of the data collection. All data are verified on receipt. The data are subject to both local validation and an external data quality audit program, which is performed on site to evaluate the completeness and accuracy of the data (69).

Outcome indicators of the database were mortality (in-hospital or 30-day post-surgery), complications including cardiac, neurological, renal, gastrointestinal, infections, return to theatre, readmissions within 30 days post-surgery. The index outcome variable for current thesis was 30-day-mortality, defined as death within 30-days post-procedure, was collected by the hospital data managers by contacting patients, family members or medical practitioners by follow-up visits or via telephone as part of routine clinical care. The Database also includes re-admission data. Mortality information is further validated through linkage to National Death Index (NDI) data. Mortality data outside 30-day of surgery were collected through linkage with the NDI database.

The database program publishes comprehensive annual reports describing the activities and outcomes of participating sites in a comparative de-identified format. The registry has developed several risk prediction models which was used for benchmarking surgical performance at a national and international level (25-28).

## 2.2 Systematic review of risk prediction model

The systematic review of risk prediction models involved articles those presented models for predicting short-term mortality following CABG. The review aim, search strategy and study selection process have been framed based on **Ch**ecklist for critical **A**ppraisal and data extraction for systematic **R**eviews of prediction **M**odelling **S**tudies, alias the CHARMS checklist (72).

Medline via Ovid was searched for peer reviewed articles published between 1946 and 2016 and EMBASE via Ovid for articles published between 1974 and 2016 to identify short-term mortality risk prediction models for patients undergoing CABG surgery. Search strategies included medical subject heading (MeSH) terms and keywords. The CHARMS checklist for review of prediction modelling studies was used for appraisal and data extraction (72) (Appendic 3.3). Clinical aim, as well as the methods of these models were critically appraised. Analyses of the extracted data focused on summarizing information on methodological characteristics of these models. Descriptive statistics was generated about model characteristics, detailed methodology, model performance and selected predictors across models. Association of a-priori defined individual methodological characteristics were sought with the discrimination capacity in validation data via Kruskal-Wallis one-way analysis of variance and Mann-Whitney-U test.

## 2.3 Definition of the key variables

The ANZSCTS registry collects preoperative, intraoperative and post-operative data from each patient undergoing cardiac surgery. Risk prediction models consider only pre-operative variables because these models are used for patients' pre-operative risk assessments. Intra-operative and post-operative variables are not available before the surgery. Preoperative variables collected by the ANZCTS registry include variables related to administrative, patient demographics, risk factors, cardiac status, hemodynamic status and previous interventions. After excluding administrative variables (for example name, address, contact details, date of birth, Medicare number, and patient identifiers) and sub-headings or supporting information for other variables (for example the variable related to arrhythmia includes seven additional variables related to subtypes of arrhythmia), 52 variables were identified as potential candidates for inclusion in risk prediction models. Based on an extensive literature review of existing cardiac surgery risk prediction models and clinical judgement (through discussion with cardiac surgeons and cardiologists) a total of 47 variables were identified as the final set of variables we would consider as potential predictors. Detailed definitions of the relevant variables are presented at the Appendix 2.3

## 2.4 Overview of model development

Detailed description of statistical analysis is provided in the respective section. In this chapter, an overview of model development method is provided; hence some duplication may be encountered. Generally, a risk prediction model refers to the function which relates the occurrence of the outcome

of interest to a set of predictors. Predictors may range from demographic characteristics, anthropometry, physical or haematological state, comorbidities any diagnostic test result etc. (73).

In cardiac surgery, these models predict operative mortality risk preoperatively, although outcomes would preferably be those that matter to individuals or patients. These could include both mortality and morbidity. In the current thesis mortality following coronary artery bypass grafting surgery among adults was modelled using ANZCTS registry data. Strategies for model development include imputation of missing values, variable selection, model development, final model estimation and validation (64, 73).

### 2.4.1 Treatment of missing data

A total of 15.8% patients had one or more predictors missing. 'Reduced ejection fraction' had the highest missing data (11.3%) followed by the New York Heart Association classification (4.5%). The remaining predictors had <1% missing observations. The pattern and extent of missing-ness in the dataset was assessed, through generating descriptive statistics and a missing indicator variable, to check the assumption for multiple imputation (MI). Patients with missing information in one or more predictors were categorized as missing in the indicator variable. The association of each independent predictor with a missing indicator variable was evaluated using the chi-square test. Majority of the variables (Age, urgency of procedure, body mass index (BMI), inotropic medication use, peripheral vascular disease, type of procedure, NYHA classification, and 30-day mortality) were found to be associated with missing indicator variable, suggesting that the data are not missing completely at random (MCAR).

Multiple imputation of missing values was done using the Imputation by Chained Equations (ICE) method in Stata version 14. Imputation was performed in three distinct steps. In step one, 10 multiply imputed datasets were generated. In step two, each of the multiply imputed datasets were analysed separately and in step three, estimates from each multiply imputed dataset were combined to generate the aggregated estimates (74). For variable selection, bootstrap bagging process was run separately in each of the imputed dataset and subsequently all predictors those were selected in any of these datasets were used for final model development (75). For model development missing data were imputed using ICE method along with multivariable Cox regression. Model estimates were generated separately on each of the 10 imputed datasets. The estimates of these 10 imputed datasets were then combined into generate MI estimates.

### 2.4.2 Variable selection and model development

Among the available variables in the database only preoperative variables were considered. Plausible variables were identified through a variety of methods, including literature review, clinical acumen, or their use in other models developed using the same database. Univariable associations between preoperative patient characteristics and mortality were assessed. Bootstrap bagging technique along with multiple regression were used to select final set of predictors from the plausible variables for the multivariable models.

A bootstrap sample of the same size of the original sample was drawn from each of the imputed dataset. The plausible variables were entered into the multivariable regression and were applied to the bootstrap samples to test the significance of the variables. For short-term mortality in chapter four, five and six logistic regression and for long-term survival model in chapter 7 Cox regression was used. A variable with a p-value of less than or equal to 0.05 was considered as significant. The process was repeated 1000 times and the percentage of times that each variable appeared as significant in 1000 bootstraps (bootstrap coverage) in the imputed datasets was recorded. Bootstrap coverage of each predictor in imputed datasets was averaged to generate an overall coverage of individual predictors. The predictors were then ranked depending on the average bootstraps coverage (75). Plausible models were developed from variables that were significant in at least 50% of the bootstrap samples (76, 77). The first model (model 1) comprises predictors which appeared as significant in 100% of bootstrap samples. Subsequent models were generated through adding one variable at a time of decreasing rank per the bootstrap coverage. The area under ROC curve (AUC), Akaike information criterion (AIC) and BIC (Bayesian information criterion) values were calculated for these models. Variables in the model with the highest AUC value were chosen for final model estimation.

### 2.4.3 Final model coefficient estimation

Final multivariable model coefficients were estimated entering the set of variables chosen through bootstrap bagging technique. For the final model, non-linearity of continuous predictors was considered by fitting fractional polynomials in the multivariable model (logistic/ Cox regression) (78). The first order interaction effects between clinically relevant risk factors were investigated. To account for hospital-level clustering mixed effect logistic regression was used for short-term model and a hospital-level random effect (shared frailty) was included in the Cox regression for long-term survival model (79).

### 2.4.4 Model performance and validation (Discrimination and calibration)

Discrimination of the short-term mortality model was assessed by calculating the AUC in the validation sample and then again using multi-fold (K = 100) cross-validation in combined datasets. The calibration was evaluated using decile-decile plot of the observed and predicted 30-day mortality. To calculate the calibration intercept and slope parameters, a linear regression model was fitted with the deciles of observed outcome as the dependent variable and the deciles of predicted outcome as the independent variable. Calibration of the survival model was assessed using the Regression Modelling Strategies (RMS) package in the R statistical software. Locally weighted scatter-plot smoother (LOESS) calibration curves were generated for each of the four time intervals plotting these probabilities against corresponding Kaplan-Meier survival estimates, stratifying on intervals of predicted survival.

### 2.4.5 Presentation of the model

The final prediction model was presented as the original regression model equation, that is, regression coefficients. Such a model can also be made available as an online calculator or as a nomogram (61). A model can also be presented as a simplified but approximate model or scoring rule when the original regression coefficients are converted and rounded to numbers that are easy to add, which are then related to absolute outcome probabilities (80). An online calculator will be developed in future.

## 2.5 Statistical software

Statistical software packages Stata (version 14), Medcalc (version 17.2) and R (version 3.3.2) were used for the analyses.

## 2.6 Ethical approval

The Institutional Review Board of each participating hospital had approved the use of their data for research purposes (Alfred HREC: 262/09). The ANZCTS registry has approved collection of patient data using an 'opt-out consent approach' (MUHREC: CF08/0322 - 2008000065). The current study received ethical approval from the Monash University Standing Committee on Ethics in Research Involving Humans (SCERH) (MUHREC: CF14/1117 – 2014000476). Relevant ethics approval documents are presented in Appendix 2.4-2.7.

# Chapter 3: Mortality Risk Prediction Models for Coronary Artery Bypass Graft

## 3.1 Introduction

Preoperative risk prediction models are increasingly being used in contemporary cardiac surgical practices. Over the past three decades, many of such models were developed for predicting short-term mortality following Coronary Artery Bypass Graft (CABG) surgery. Many of them were developed decades ago. Over the past decades, the demography and risk profile of cardiac surgery patients have changed, and newer and more robust modelling techniques emerged. Hence, appraisal of the methodology and performance of these models is required to assess their applicability in current practice setting as well as for the necessity of upgradation.

The chapter reports the findings of systematic review of the existing risk prediction models, for Coronary Artery Bypass Graft (CABG) surgery patients, currently being used around the globe. The review critically appraised the development process of these models and outlined the gaps in knowledge and practice in contemporary prediction modelling.

The manuscript for the review 'Mortality risk prediction models for coronary artery bypass graft: current scenario and future direction' has been accepted for publication in The Journal of Cardiovascular Surgery.

## 3.2 Manuscript

Declaration for thesis chapter 3

**Manuscript:** Mortality Risk Prediction Models for Coronary Artery Bypass Graft: Current scenario and future direction (Accepted for publication in Journal of Cardiovascular Surgery)

**Declaration by candidate**

In the case of Chapter 3, the nature and extent of my contribution to the work was the following:

| Nature of contribution | Extent of contribution (%) |
|---|---|
| Study concept and design, literature search, data analysis and interpretation, manuscript development and preparation | 70% |

The following co-authors contributed to the work. Where co-authors are students at Monash University, the extent of their contribution in percentage terms is stated:

| Name | Nature of contribution | Extent of contribution (%) for student co-authors only |
|---|---|---|
| Billah, B | Study concept, data analysis and manuscript editing | |
| Reid, C | Study concept, study design and manuscript editing | |
| Cochrane, A | Study concept, Interpretation of result and manuscript editing | |
| Tran, L | Interpretation of result and manuscript editing | |
| AL Ramadan M | Data extraction and manuscript editing | 3% |
| Hossain N | Data extraction and manuscript editing | 2% |

The undersigned hereby certify that the above declaration correctly reflects the nature and extend o the candidate's and co-author' contributions to this work

| | | |
|---|---|---|
| **Candidate's Signature** | [redacted] | **Date 10/03/2017** |
| **Main Supervisor's Signature** | [redacted] | **Date 10/03/2017** |

**Mortality risk prediction models for coronary artery bypass graft: current scenario and future direction**

**Authors**

Md Nazmul Karim[1]

Christopher M Reid [1,2]

Lavinia Tran[1]

Andrew Cochrane[3]

Mohammed Alramadan[1]

Md N Hossain[1]

Baki Billah[1*]

[1]School of Public Health and Preventive Medicine, Monash University, Melbourne, Victoria, Australia

[2]School of Public Health, Curtin University, Perth, WA, Australia

[3]Dept. of Cardiothoracic Surgery and Dept. of Surgery, Monash Medical Centre, Clayton, Victoria, Australia

**Corresponding Author**

**Baki Billah, PhD**

Senior Lecturer of Biostatistics

School of Public Health and Preventive Medicine

Faculty of Medicine Nursing and Health Science, Monash University

The Alfred Centre, (level 6), 99 Commercial Road, Melbourne, VIC 3004, Australia

Email : [redacted]

Ph : + [redacted] Fax: + 61 3 9903 0828

**Mortality risk prediction models for coronary artery bypass graft: current scenario and future direction**

**ABSTRACT**

INTRODUCTION:

Many risk prediction models are currently in use for predicting short-term mortality following Coronary Artery Bypass Graft (CABG) surgery. This review critically appraised the methods that were used for developing these models to assess their applicability in current practice setting as well as for the necessity of upgradation.

EVIDENCE ACQUISITION:

Medline via Ovid was searched for articles published between 1946 and 2016 and EMBASE via Ovid between 1974 and 2016 to identify risk prediction models for CABG. Article selection and data extraction was conducted using the CHARMS checklist for review of prediction model studies. Association between model development methods and model's discrimination was assessed using Kruskal-Wallis one-way analysis of variance and Mann-Whitney-U test.

EVIDENCE SYNTHESIS:

A total of 53 risk-prediction models for short-term mortality following CABG were identified. The review found a wide variation in development methodology of risk prediction models in the field. Ambiguous predictor and outcome definition, sub-optimum sample size, inappropriate handling of missing data and inefficient predictor selection technique are major issues identified in the review. Quantitative synthesis in the review showed 'missing value imputation' and 'adopting machine learning algorithms' may result in better discriminative power of the models.

CONCLUSIONS:

There are aspects in current risk modelling, where there is room for improvement to reflect current clinical practice. Future risk modelling needs to adopt a standardised approach to defining both outcome and predictor variables, rational treatment of missing data and robust statistical techniques to enhance performance of the mortality risk prediction.

**Key words:** Coronary artery bypass surgery, Cardiac surgical procedures, Risk prediction model, coronary revascularization, operative mortality, short-term mortality, risk stratification, clinical prediction rule.

**BACKGROUND**

Understanding the operative risk, prior to coronary artery bypass grafting (CABG) surgery allows both surgeons and patients to participate effectively in deciding on choices for treatment (1, 2). Surgeons can decide the most appropriate treatment plan for a specific patient by considering predicted risk score generated by a prediction model in addition to their clinical assessment (3, 4). These scores can be used to counsel patients and thus to facilitate better informed consent. Prediction models are essential for benchmarking of physician and institution performance (5) and for the appropriate allocation of healthcare resources (6).

Paiement and colleagues (7), in 1983, proposed a scoring system for cardiac surgery patients. The Parsonnet system (8) developed in 1989, was the first to get widespread acceptance despite being criticised for including subjective variables in the model. Over the past decades, numerous risk prediction models have been proposed for predicting operative mortality following cardiac surgery and many are used in daily practice (9).

Newer concepts and technologies have emerged in the field of statistics which have the potential to improve prediction further. Many models' performance has been affected by a changing patient's demography. Further, some models lack in clarity on several key factors associated with the development process including; how data was managed, the variable selection process used, statistical techniques employed, and how the validation was performed. For a risk prediction model to be used routinely in practice, the modelling methodology should be clearly described and the proposed model should be easy to implement and clinically relevant. Appraisal of the methodology and performance of these models is required to assess their applicability in current practice setting as well as for the necessity of upgradation.

Several review articles were published on risk prediction models for cardiac surgery, most of them focussed either on comparison of different models for cardiac surgery (3, 4, 10-14) or validation of a model. None of them explicitly focused on appraising methodology of models predicting short-term mortality following CABG surgery. Nilsson (9) compared 19 models, including both morbidity and mortality predictions. A systematic review by Head (15) investigated only the risk factors for adverse event following cardiac surgery. The aim of the current review was to critically appraise the methodology used in developing short-term mortality risk prediction models for patients undergoing CABG.

**METHODS**

**Study selection and data extraction**

The review aim, search strategy and study selection process have been framed based on seven key items in CHARMS (16) checklist (Table 1).

Table 1: The review aim, search strategy, and study inclusion and exclusion criteria

| Sl. | Items and responses |
|---|---|
| 1 | Prognostic versus diagnostic prediction model |
| | Prognostic prediction model: The aim is to review models to predict future events. |
| 2 | Intended scope of the review |
| | The models intended to inform physicians' therapeutic decision making |
| 3 | Type of prediction modelling studies |
| | Prediction model development without external validation in independent data |
| 4 | Target population to whom the prediction model applies |
| | Patients undergone isolated coronary artery bypass grafting surgery |
| 5 | Outcome to be predicted |
| | Mortality following isolated coronary artery bypass grafting surgery |
| 6 | Time span of prediction |
| | Event within 30 days post operatively |
| 7 | Intended moment of using the model |
| | Models to be used preoperatively to predict the risk of postoperative complications |

Search strategy

Medline via Ovid was searched for peer reviewed articles published between 1946 and 2016 and EMBASE via Ovid (1974 -2016) to identify short-term mortality risk prediction models for patients undergoing CABG surgery. Search strategies included medical subject heading (MeSH) terms and keywords. For CABG, search terms included subject key words 'cardiac surgery', 'cardiothoracic surgery', 'cardiac surgical procedures', 'myocardial revascularization' and 'coronary artery bypass grafting'. For surgery outcome search terms included key words '30-day mortality' 'hospital mortality, 'in-hospital mortality, 'operative mortality', short-term mortality', 'post-operative death' and 'operative death. For risk prediction model, in addition to 'Ingui filter' for searching prognostic models (17), search terms included key words 'risk' in combination with different permutation of 'model', 'prediction', 'assessment', 'stratification, 'algorithm', 'score', 'index', 'rule', and 'tool'. Detailed search strategy and history is presented in supplementary Table 1 and 2. Our electronic search returned 1123 articles, with an additional 129 articles retrieved through google scholar and a hand search of the citations listed in publications from other risk prediction model for cardiac surgery review articles. After removing the duplicates, 818 papers were available for screening. Detailed search strategies are presented in Appendix 3.1 and 3.2

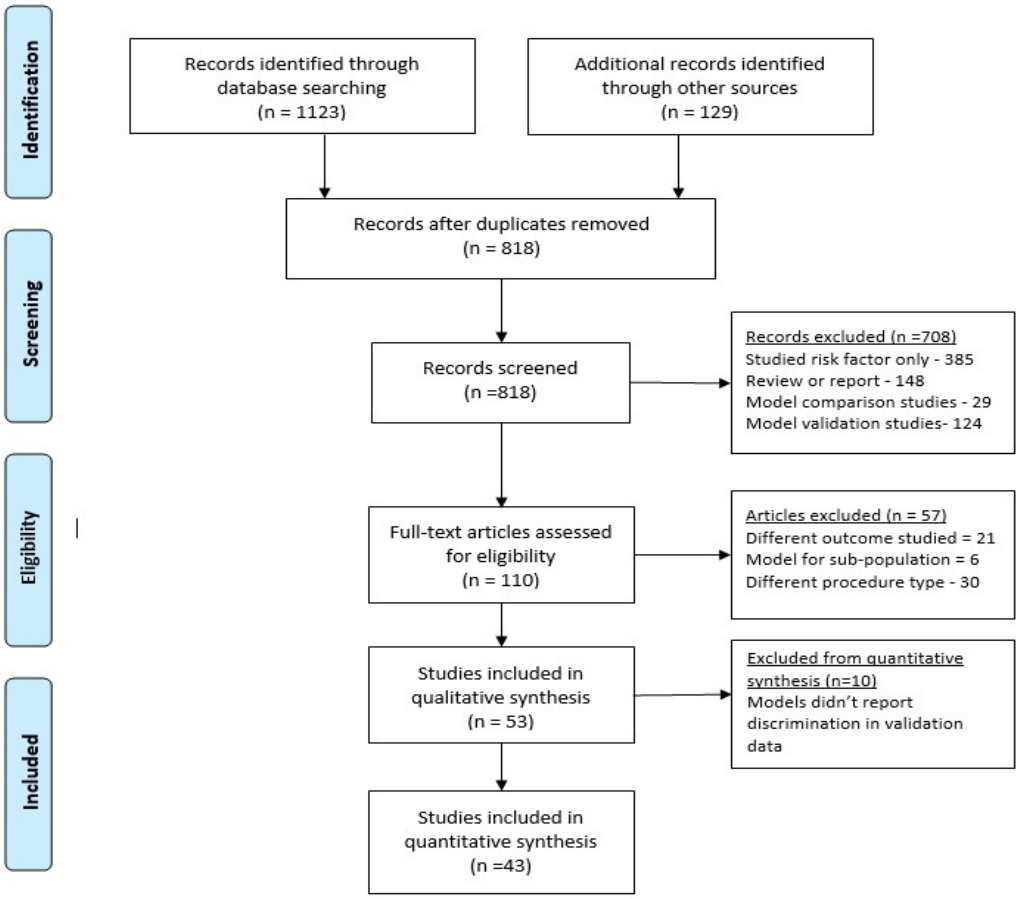
Study inclusion Criteria

An article was considered eligible, if it focused on development of a risk model to predict short-term mortality following specifically CABG, as well as presented a model. Articles which focused on updating previously developed models were considered eligible if an updated version of the model is presented. Articles illustrating the comparison of a locally developed model to an existing model were considered eligible only if they presented a validated model with estimated coefficients. Articles focused on development of a risk model for case-mix were considered eligible only if case-mix included CABG surgery.

Study exclusion Criteria

Articles which studied the effect of individual risk factor or comorbidity or preoperative cardiac status on short-term mortality were excluded. Articles which reviewed risk prediction models were excluded. Furthermore, articles comparing performance or applicability of models were considered ineligible if their aim was to validate previously developed models in a local population, or to evaluate the performance of models developed by others.

**Figure 1:** PRISMA flow diagram illustrating study selection.



**RESULTS**

A total of 53 (7, 8, 21-71) models were identified. Among them 31 are intended to predict short-term mortality following CABG only and 22 are all procedure models which include CABG (Table 2).

Table 2: Short-term mortality risk prediction models for CABG surgery

| Model reference | Model name | Population | Procedure type | Data collection | | | EPV* |
|---|---|---|---|---|---|---|---|
| | | | | Source | Period | Centres | |
| Aktuerk 2016 | HES model | UK | CABG only | Admin | 2008-2011 | Multiple | 54 |
| Antunes 2007 | Portugal model | Portugal | CABG only | Study | 1992-2001 | Single | 09 |
| Berg 2011 | Norway model | Norway | Case-mix | Study | 2000-2007 | Single | 17 |
| Bernstein 2000 | Bernstein-Parsonate | USA | Case-mix | Study | 1994-1995 | Multiple | 33 |
| Billah 2010 | Global Model | Australia | Case-mix | Registry | 2001-2008 | Multiple | 61 |
| Billah 2014 | Ausscore II | Australia | CABG only | Registry | 2001-2011 | Multiple | 166 |
| Bridgewater 1998 | The UK society score | UK | CABG only | Registry | 1995-1996 | Multiple | 03 |
| Carosella 2009 | Latin american model | Argentina | Case-mix | Study | 1994-2001 | Single | 21 |
| Cheng 2015 | Spanish model | Spain | Case-mix | Study | 2001-2014 | Single | 25 |
| Chong 2003 | ANN CABG model | Taiwan | CABG only | Study | 1997-2002 | Single | 04 |
| Chung 2015 | ACS NSQIP model | USA | CABG only | Registry | 2005-2010 | Single | 08 |
| D'Errigo 2007 | Italian CABG model | Italy | CABG only | Study | 2002-2004 | Multiple | 75 |
| Eagle 1999 | ACC/AHA | USA | CABG only | Registry | 1996-1998 | Multiple | 27 |
| Edward 1989 | CASS model | USA | CABG only | Registry | 1994-1998 | Multiple | 22 |
| Fortescue 2001 | QMMI score | USA | CABG only | Study | 1993-1995 | Multiple | 14 |
| Gabrielle 1997 | Modified parsonnet | France | Case-mix | Admin | 1992-1993 | Multiple | 10 |
| Grover 1993 | Vattern affairs | USA | CABG only | Study | 1987-1990 | Multiple | 58 |
| Ham'meister 1994 | VA Quality of care | USA | CABG only | Study | 1990-1992 | Multiple | - |
| Hannan 2006 | NYS Score 2000 | USA | CABG only | Registry | 2002-2003 | Multiple | 44 |
| Hannan 2013 | New York Risk score | USA | CABG only | Admin | 2009-2010 | Single | 26 |
| Higgins 1992 | Cleaveland score | USA | CABG only | Admin | 1986-1990 | Single | 21 |
| Huijskes 2003 | Amphiscore | Netherlands | Case-mix | Admin | 1997-2001 | Single | 22 |
| Keogh 2003 | SCTS Comples | UK | CABG only | Registry | 1999-2000 | Multiple | - |
| Keogh 2003 | SCTS Simple | UK | CABG only | Registry | 1999-2000 | Multiple | - |
| Kotting 2014 | German CABG Score | Germany | CABG only | Registry | 2004-2008 | Multiple | 157 |
| Lipperman 1997 | Neural network | USA | CABG only | Registry | 1993-1993 | Multiple | 139 |
| Magovern 1996 | Magovern | USA | CABG only | Admin | 1991-1994 | Single | 5 |
| Marshall 1994 | Bayesian-Logit model | USA | CABG only | Registry | 1987-1990 | Multiple | 83 |
| Mejia 2013 | InsCor | Brazil | Case-mix | Study | 2007-2009 | Single | 27 |
| Miyata 2015 | JACVD risk model | Japan | Case-mix | Registry | 2005-2009 | Multiple | 51 |
| Motumura 2008 | JACVSD model | Japan | CABG only | Registry | 2001-2005 | Multiple | 11 |
| Mozes 1998 | Israel model | Israel | CABG only | Study | 1994-1994 | Multiple | 19 |
| Nashef 1999 | EuroSCORE | Europe | Case-mix | Registry | 1995-1995 | Multiple | 41 |
| Nashef 2012 | EuroSCORE II | Global | Case-mix | Study | 2010-2010 | Multiple | 49 |
| Nilsson 2006 | ANN Global model | Europe | Case-mix | Registry | 1995-1995 | Multiple | 26 |
| O'Connor 1992 | NNE model | USA | CABG only | Study | 1987-1989 | Multiple | 17 |
| Paiement 1983 | Montreal heart model | Canada | Case-mix | Study | 1993-1997 | Single | 6 |
| Parsonnet 1989 | Parsonnet score | USA | Case-mix | Admin | 1982-1987 | Multiple | 31 |
| Pitkanen 2000 | Finland model | Finland | Case-mix | Admin | 1992-1996 | Single | 12 |
| Pons 1997 | Pons score | Spain | Case-mix | Study | 1994-1994 | Multiple | 10 |
| Ranucci 2009 | ACEF | Italy | Case-mix | Admin | 2001-2007 | Single | 85 |
| Reid 2009 | AusSCORE | Australia | CABG only | Registry | 2001-2005 | Multiple | 17 |
| Roques 1995 | French Score | France | Case-mix | Study | 1993-1993 | Multiple | - |
| Sanon 2013 | THIRST | USA | Case-mix | Admin | 1995-2007 | Single | 72 |
| Shahian 2009 | STS CABG model | USA | CABG only | Registry | 2002-2006 | Multiple | 557 |
| Shroyer 1998 | The 1995 CABG model | USA | CABG only | Registry | 1990-1994 | Multiple | 137 |
| Shroyer 1999 | The 1996 CABG model | USA | CABG only | Registry | 1990-1996 | Multiple | 157 |
| Sing 2008 | MCRS | USA | Case-mix | Admin | 2004-2006 | Multiple | - |
| Tu 1995 | Ontario score | Canada | Case-mix | Registry | 1991-1993 | Multiple | 66 |
| Verduijin 2007 | PBN | Netherlands | Case-mix | Admin | 1998-2004 | Single | 28 |
| Wong 1999 | NCRS | Canada | CABG only | Study | 1995-1995 | Single | 7 |
| Wouters 2002 | CORRAD score | Netherlands | CABG only | Admin | 1998-2000 | Single | 9 |
| Zheng 2013 | SinoScore | China | CABG only | Registry | 2007-2008 | Multiple | 22 |

*Event per variable

**Model development data**

Thirteen models used administrative data, 18 models used data from studies dedicated for model development, 4 of them were prospective studies. Only 22 of them were developed from a registry database. Among the models, 35 were based on multi-centre data, with the number of centres used for development of these models ranging between 2 and 819. Only 20 models used data from more than 10 centres. The number of patients in derivation cohorts was as low as 423. Only 16 of the models had a derivation sample larger than 10,000. Event per variable (EPV) of the models ranges from as low as 3 to 557. Out of 53 models, 38 have an EPV greater than the recommended 10 (Table 2). Most models (75.5%) did not report the frequency or type of missing data.

**Outcome of interest**

Mortality alone was the outcome in 48 models and 5 used 'adverse event' to predict both mortality and morbidity by use of the same model. Short-term mortality is defined as 'in-hospital mortality' by 21 models, as '30-day mortality' by 16 models and as 'operative death' by 5 models. Seven models used mortality within 30 days from operation or later if the patient is still hospitalised (Table 3). EuroSCORE (54) considered up to 90 days along with in-hospital mortality. The CORRAD score (62) considered early mortality as 180 days after surgery. Cleveland score (38) considered in-hospital and 30 days after discharge.

**Predictors of short-term mortality**

Predictor selection methods used by the models varied highly. Furthermore, a wide variation was seen across models regarding the definition of predictor variables**.** The median number of predictors reported in the included models was 10. Number of predictors in the models ranges from 3 in ACEF model (53) to as high as 42 in the Modified-Parsonnet model (34). Most frequently included demographic predictors were age (n=51) and gender (n=34). Among the pre-operative cardiac state variables priority of operation (n=42), ejection fraction (n=45) previous surgery (n=23), previous myocardial infarction (n=22), reoperation (n=22), NYHA-class (n=16), cardiogenic shock (n=15) were notable. Among the co-morbid conditions renal problem (n=39), peripheral vascular disease (n=28), respiratory diseases (n=26) and diabetes mellitus (n=20) were quite frequently included in the models.

Table 3: Model development process (n=53)

| Model characteristics | Frequency | Percent |
|---|---|---|
| A. Outcome | | |
| Mortality | 48 | 90.6 |
| Adverse event (include mortality) | 5 | 9.4 |
| B. Definition of mortality | | |
| In-hospital mortality | 21 | 39.6 |
| 30-day mortality | 16 | 30.2 |
| In-hospital or 30-day mortality | 7 | 13.2 |
| Operative death (not defined) | 5 | 9.4 |
| Others | 4 | 7.6 |
| C. Handling missing value | | |
| Missing data issue not addressed | 40 | 75.5 |
| Variables with missing data removed | 3 | 5.7 |
| Imputation | 10 | 18.9 |
| D. Model selection | | |
| Univariable association/clinical acumen (5) | 5 | 9.4 |
| Multiple regression analysis (34) | | |
| Full model | 11 | 20.8 |
| Stepwise | 23 | 43.4 |
| Machine learning algorithm (8) | | |
| Bootstrap bagging | 5 | 9.4 |
| Neural Network | 3 | 5.7 |
| Bayesian algorithm (6) | 6 | 11.3 |
| E. Other statistical issues* | | |
| Univariable screening of predictors | 33 | 62.3 |
| Reported model assumption testing | 3 | 5.7 |
| Interaction term included in model | 5 | 9.4 |
| Non-linear association addressed | 15 | 28.3 |
| Linear predictor categorized | 45 | 84.9 |
| Centre variance addressed | 2 | 3.8 |

* Total percentage does not equal 100 due to multiple responses

**Model development**

Table 3 summarises the development process adopted by the risk prediction models. Univariable screening was performed in 33 models to select predictors for the regression model. Most models (n= 45) converted one or more continuous variables (for example, serum creatinine value) to binary or categorical variables using (arbitrary) cut-off points. Eight models employed single imputation and 2 conducted multiple imputation of missing data prior to modelling. Only few reported model assumption testing (n=3) and addressed centre variance (n=2). Six of the models were developed using Bayesian-algorithm. Machine learning algorithms were used by 8 models, of these 3 used artificial neural networks (ANN), while 5 used bootstrap bagging technique. Multiple logistic regression (n=34) is the most commonly used technique for model selection, 11 used a full-model and 23 used stepwise technique. Five models were developed based on univariate relation of predictors with outcome or solely based on clinical acumen. Only 5 models included interaction terms in the model and 15 models

addressed non-linear association of predictors with outcome. In 25 models integer risk scores were derived for ease of use in the clinical setting. Only 2 models have online calculator.

Table 4: Model performance and validation (n=53)

| Performance measures | Frequency | Percent |
|---|---|---|
| A. Validation type based on data source | | |
| None reported (6) | 6 | 11.3 |
| Internal validation only (30) | | |
| Validation data; random split of original data | 17 | 32.1 |
| Re-sampling or multi-fold cross validation | 4 | 7.5 |
| Combined | 7 | 13.2 |
| Derivation data | 2 | 3.8 |
| External validation only (12) | | |
| Same population different time (Temporal) | 7 | 13.2 |
| Different population | 5 | 9.4 |
| Both internal and external (5) | | |
| Both internal and external validation | 5 | 9.4 |
| B. Calibration | | |
| Calibration type based on data source | | |
| None reported | 4 | 7.5 |
| Internal: derivation data | 9 | 17.0 |
| Internal: validation data | 28 | 52.8 |
| External data | 6 | 11.3 |
| Both internal and external data | 6 | 11.3 |
| Measures of calibration* | | |
| None reported | 4 | 7.5 |
| H-L GOF | 30 | 56.6 |
| Observed to predicted ratio (OPR) | 30 | 56.6 |
| Calibration plot | 16 | 18.0 |
| Risk Adjusted Mortality Rate (RMAR) | 3 | 5.7 |
| Mean Standard Error (MSE) | 2 | 3.8 |
| Others (kappa, shrinkage coefficient etc.) | 4 | 7.5 |
| C. Discrimination | | |
| Discrimination type based on data source | | |
| None reported | 6 | 11.3 |
| Internal: derivation data only | 9 | 17.0 |
| Internal: validation data only | 26 | 49.1 |
| External data only | 5 | 9.4 |
| Both internal and external data | 7 | 13.2 |

* Total percentage does not equal 100 due to multiple responses

**Model performance and validation**

Table 4 summarises the performance measures and validation processes of the risk prediction models. Information about validation was available for 47 models, only 5 of them reported both internal and external validation, 30 of them reported internal validation only and 12 reported external validation only. Internal validation includes validation dataset derived by random split of original sample (n=17) and resampling of derivation sample (bootstrap, cross validation etc.) (n=4). Seven models used both resampling and random split of derivation sample for validation. External validation included

validation on temporal data (n=7) and data from different setting or population (n=5) (Table 4). Calibration of the model was reported by 49 models. The Hosmer-Lemeshow (H-L) goodness-of-fit was used in 30 models, 30 models compared observed-to-predicted event rate and 16 models reported calibration-plot. Risk adjusted mortality rate (RAMR) was used by two models and mean standard error (MSE) was used by 2 models. Kappa-statistics and correlation coefficient was used by one each. AUC in validation set was reported as a measure of discrimination by 42 models. Median reported AUC was 0.776, ranges from 0.699 to 0.886.

**Association of methodological characteristics and model discrimination**

Figure 2 demonstrates the association between discriminative power (AUC) of the models in validation sample and individual methodological characteristics. Discrimination was higher in models with missing value being imputed (p=0.038). Among the predictor selection methods machine learning algorithm—including neural network and bootstrap bagging—were found to have higher discriminative power (p=0.034) over Bayesian algorithm and logistic regression.

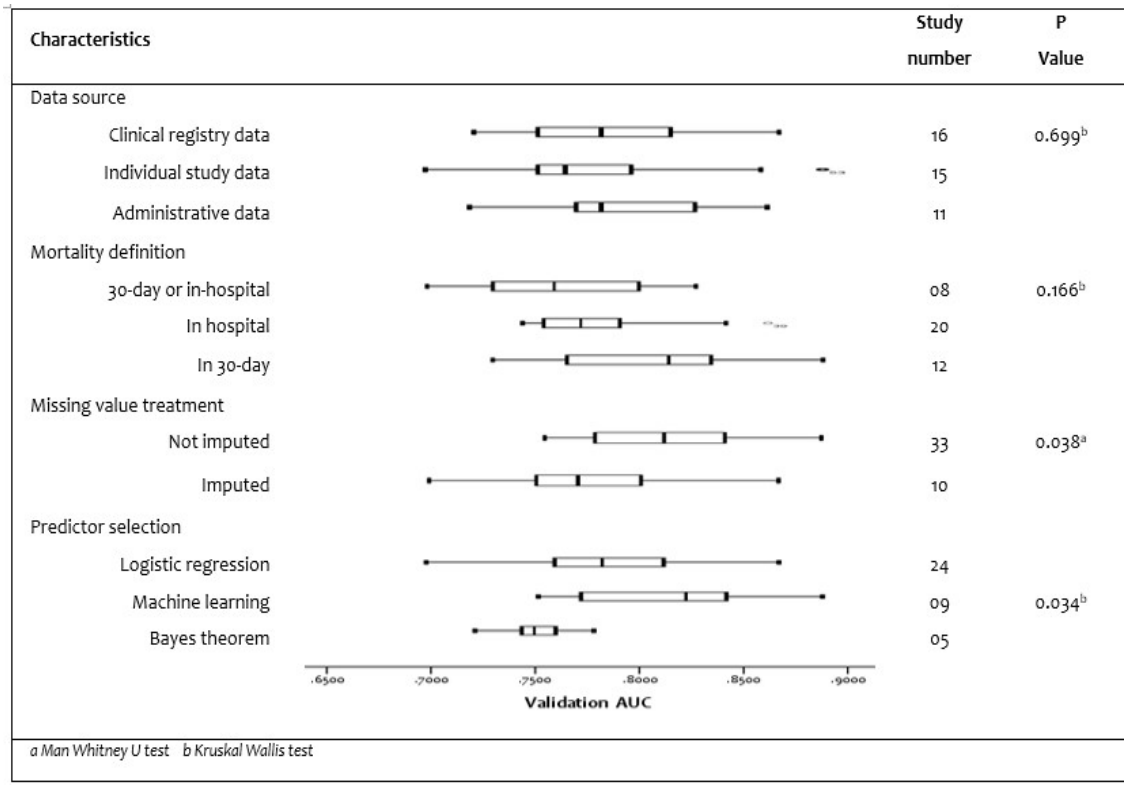



Figure 2: Box-plot illustrating association of methodological characteristics with model discrimination in validation data (n=43).

## DISCUSSION

This review included 53 models predicting risk of short-term mortality following CABG surgery from a variety of settings and populations and identified areas in which there is room for improvement, this so that we may reflect upon the implications it has on current clinical practice in context of a changing patient demography. Current reviewed used CHARMS checklist, designed specifically for primary prediction modelling studies, for critical **a**ppraisal and data extraction for reviews of prediction **m**odelling **s**tudies (16).

### Outcome of interest for the model

Post-operative mortality is a widely-accepted outcome used by prediction models despite its limitations (4, 8). It had been defined quite diversely in the included models. It is likely that the number of deaths will be less in a time span of 30 post-operative days than in a time span of 30 postoperative days or longer if the patient is still hospitalized. In this way, the EuroSCORE (47) and STS-CABG score (57) are likely to overestimate operative-mortality. A small difference in reported deaths due to differing definitions may affect prediction performance of the model (72). In-hospital mortality data, is relatively easier to collect; however, duration of hospital stay may sometimes be the function of some other factors unrelated to current CABG surgery. On the contrary after discharge from hospital factors unrelated to the current operation may alter the risk of death. However, evidence suggests a definite time-period is statistically preferable (73). Some researchers stretched on extending the duration to 6 months postoperatively to prevent the underestimation (74).

### Model development data

Regardless of the complexity of a risk prediction model, its accuracy depends largely on the quality of the data used. Although administrative data are easily accessible, many important clinical variables may not be available in these databases (75, 76). Data from clinical registries are increasingly used in prediction modelling. Although, such databases are especially prone to missing data and missing important predictors, research shows that models derived from the clinical registry database outperform those derived from administrative data (13, 76). Clinical registries are mostly supported by professional societies, state or federal statutory provisions. Most of them cover a wide section of the population and are multi-centred. Models using data from single or fewer centres are unlikely to capture wider population characteristics. Patients from an institution are likely to be clustered within that centre and are likely to be more similar than patients across centres in terms of treatments and

demography. While using data from multiple centres for model development, random effects due to cluster variation should be addressed in the analysis. Only the STS-CABG (57) model has used mixed-effect models to address the centre variation. Use of suboptimal sample may result 'spurious associations' between predictor and outcome, and may affect model precision (77). Further it restricts the number of predictor to be included in the model, because the data should be large enough to have the number of outcome being at least ten times the number of candidate predictors (78). Relatively smaller EPV—number of people with outcome event in the data relative to the number of variables included—may result in overfitting of the model (20). EPV of ten or more is frequently recommended to avoid overfitting (19).

Missing data is a common problem in medical research, irrespective of the rigor of data collection process. Extent of missing data and the methods used to handle this missing data may greatly influence performance of the model (79). However, reporting on the frequency and type of missing data is poor among the model development studies. Only a few of the currently used CABG risk prediction models addressed the impact of missing data adequately. Although only two models used multiple imputation, it is generally considered as the preferred method for handling missing data in prediction research.

**Plausible predictors**

For predicting surgical outcome of CABG, primarily preoperative patient characteristics are used as risk prediction models are mostly needed, prior to surgery for surgical decision making and patient counselling. Although intraoperative predictors might have improved the prediction, this information is only available at surgery. Preoperative predictors may range from patient demographics and clinical characteristics to preoperative cardiac status.

There have been widespread inconsistencies across models in defining the predictors, which makes the task of standardization of model performance difficult. Diverse definitions and measurement methods of predictors are a potential source of heterogeneity and bias (80). Composite variable 'preoperative critical stage' was included in 12 models; however, its definition is quite diverse across models. In several models shock, haemodialysis, acute renal failure, arrhythmia etc. were considered under 'preoperative critical state'. These variables were used as independent predictors in many models. Renal problems appeared as predictors in different models with varying definition such as, serum creatinine, renal failure and dialysis. Diverse definition of predictors may restrict its use in some settings where the definition does not match. NNE score (50) used co-morbidity score, although a

guideline is defined for the variable; it's difficult to rule out ambiguity. Co-morbidity may range from cardiac condition like, angina, shock, pulmonary disease, cerebrovascular disease, extra-cardiac arteriopathy, diabetes etc. with differing impact on mortality. It is more logical to use individual co-morbidities with a standard definition.

Uniform definition based on standard criteria and cut-off help generalizing the model and ensures its utility across wide range of population and setting. Further improper classification of predictor categories may compromise model performance (81). Although categorization of numeric data may apparently improve the usability of the model, it may cost certain degree of 'predictability'. Categorisation assumes a constant risk up to the cut-point and then a different risk beyond the cut-point (82). Unless categories of a predictor is adequately discriminatory and scientifically plausible, continuous data should be used to prevent loss of valuable information (83). However, for purpose of identifying specific clinical conditions like diabetes, hypertension, hypercholesterolemia etc. categorization may be useful. In such situations, categorical state of the variable should be scientifically plausible. One note of caution, keeping continuous predictor should preclude assessment of non-linearity (84), if required nonlinear term should be kept in the model.

To ensure optimal robustness and validity, the number of candidate predictors in the model should be minimal otherwise co-linearity among predictors may hold back reliable estimation. However, with plethora of published literature, limiting the number of candidate predictor is a big challenge. A systematic review conducted by Head and colleagues (15) identified a set of predictors for short-term mortality following CABG, many of those variables captures the similar attribute. Steyeberg (85) recommended combining similar variables based on subject knowledge as an approach to reduce number of predictors

Univariable association with outcome is an approach used to screen candidate predictors in most models. Ideally candidate predictor should be selected without studying association with the outcome in the data under study (12, 85) as the process carries a great risk of introducing bias (85). Predictor selection based on univariable association may omit risk factors that are not associated individually, but become significant in presence of other factors (4). There may be many models with different combinations of risk factors for the same data which fit equally well, hence it's not prudent to exclude variables not appeared significant in univariable screening.

**Model Development technique**

The Montreal Heart Model (7) relied on clinical acumen only for model selection, whereby all plausible risk factors are kept in the model regardless of their statistical significance. However, model development solely on theoretical grounds is discouraged (78).

Among the model development techniques, 'Bayesian algorithms' are particularly impermeable against poor data quality (ie. missing value) and were used by many models. Logistic regression gained popularity as model development tool for its simplicity of use over 'Bayesian algorithms' (4). Availability of better quality data also paved the pathway of the shift. Evidence also shows logistic regression to perform better over Bayesian models (47). The STS-CABG initially used a Bayesian-algorithm, but later in 1995 adopted logistic regression for model building (4). Parsonnet score (8), Modified Parsonnet score (34) and ACEF-score (53) used only univariable association for predictor selection. Most other models used stepwise logistic regression. However, there are concerns among statisticians about using this automated selection process, as this may produce non-reproducible models and can mask multi-collinearity. They may also result in unstable models and may select noise variables with relatively smaller sample (86). Machine-learning algorithms are thought to overcome some of the limitations of logistic regression as they allow "nonlinear information processing" (87). Among the machine learning techniques, "neural networks" and "bootstrap bagging" are increasingly being used in recent models. Bootstrap re-sampling technique identifies predictors that have stability (87, 88).

No consensus has yet been reached over the number of variables to include in a risk model. A model with excess variables have the potential to demonstrate spurious association between some variables and outcome without scientific plausibility and likelihood of co-linearity also increases (53, 89). Inclusion of a variable should be weighed against the amount of precision it adds to the model (90). Conversion of the coefficients to integer risk scores ensures simplicity in clinical use, however this should be done with extreme caution and with an appropriate mathematical technique. Scaling factors for conversion should also be chosen carefully to ensure minimal loss of precision (91).

**Model performance and validation**

Internal validation of risk model determines its reproducibility (85). Models perform better when they are assessed in the same dataset used to develop the model. Hence, the assessment of the model performance should not rely on the development dataset (92). Splitting the original data into

derivation and test subsets, seen frequently in predictive modelling, generally provides little additional benefit beyond that of the assessment in the development data (93). Further in small datasets, the use of split-sample methods increases the risk of bias as derivation and validation data gets even smaller (80). Bootstrapping resampling is the preferred internal validation method, as it captures the optimism in model performance, and provides a shrinkage factor to adjust the estimated regression coefficients (78). Further the technique is relatively safer in this regards, as they allocate a data for validation equal to the original sample, thus ensures adequate validation sample even with relatively smaller original sample (78).

External validation, is the process of determining predictive performance of a model in data that are not used in model development, assesses generalizability or transportability of a model (85,89). Temporal validation, a form of external validation, is particularly relevant where the model is validated in more recently treated patients from the same cohort to ensure that the model is stable in future application (85).

Calibration measures the accuracy of the predicted risk compared with the observed risk (94). The Hosmer-Lemeshow (H-L test) test assesses whether the observed event rates match expected event rates in subgroups of the model population (89). However, the test has the potential to generate data driven conclusion, which may not be clinically relevant (10). Decile-decile plot instead of H-L test was used by some of the models to provide graphical representation of model calibration. Shrinkage of regression coefficients may also be used to assess calibration (95).

Discrimination is the ability of the model to distinguish between those with and without the outcome and is typically assessed using the *c*-statistic, which is the equivalent to the area-under-the-receiver operating characteristic curve (AUC). The *c*-statistic should not be used as the only performance measure (96, 97). Despite widely being used the *c*-statistic/AUC is not very sensitive measure of discrimination as it depends only on the ranks of the prediction and not on the actual value (98). 'Predictive-ness curve'(99) and 'reclassification table' (97) are among several statistical methods that can be used as alternative to AUC.

Models in this review were not explicitly ranked based on quality or performance, due to lack of agreed criteria for rating risk prediction model's quality. Further this paper aims to review all available models in the field to summarizing status and practice in the risk prediction modelling, hence a pooled performance or a meta-analysis is beyond the scope of the study. Discrimination capacity of the models used for quantitative synthesis are mostly derived from internal validation data, which is

subject to overfitting. Discrimination measure in external data would have been better, however most primary modelling study included in the study did no report validation in external sample, as external validation was not the focus of the review.

## CONCLUSION

For a risk model to be widely accepted, modelling methodology should be robust and the model must be user friendly and clinically relevant. The current review identified several methodological areas where there is scope for improvement. Clear and standardised data definition is necessary for both outcome and predictor variables. Modelling processes must adopt all available robust technologies to maximize model's prediction ability. Use of robust variable selection method along with rational treatment of missing data may enhance mortality risk prediction performance.

## AUTHORS' CONTRIBUTIONS

MNK developed the protocol, analysed data and interpreted the results, and wrote the manuscript; MA, MNH screened the papers and extracted data and edited manuscript. BB supervised research activity, edited manuscript critically and provided overall guidance. CR, LR and AC provided guidance for conceptualization of clinical perspective of the paper, provided critical input to the manuscript. All authors approved the final manuscript.

## ACKNOWLEDGEMENTS

Authors thankfully acknowledge contribution of Lorena Romero, Senior Medical Librarian, The Ian Potter Library for her assistance in formulating search strategy.

## AUTHORS' DECLARATION

M.N. Karim has an Australian Postgraduate Awards Scholarship from Monash University. The funding bodies had no role in the design and conduct of the study, the collection, management, analysis and interpretation of the data, or the preparation, review, or approval of the manuscript. The study received 'no external financial support' and has no conflict of interest to declare.

**REFERENCE**


1. Head SJ, Bogers AJ, Serruys PW, et al. A crucial factor in shared decision making: the team approach. The Lancet. 2011;377(9780):1836.

2. Kolh P. Importance of risk stratification models in cardiac surgery. European heart journal. 2006;27(7):768-9.

3. Granton J, Cheng D, editors. Risk stratification models for cardiac surgery. Seminars in cardiothoracic and vascular anesthesia; 2008: SAGE Publications.

4. Shahian DM, Blackstone EH, Edwards FH, et al. Cardiac surgery risk models: a position article. The Annals of Thoracic Surgery. 2004;78(5):1868-77.

5. O'Connor GT, Plume SK, Olmstead EM, et al. A regional intervention to improve the hospital mortality associated with coronary artery bypass graft surgery. Jama. 1996;275(11):841-6

6. Nilsson J, Algotsson L, Höglund P, et al. EuroSCORE predicts intensive care unit stay and costs of open heart surgery. The Annals of thoracic surgery. 2004;78(5):1528-34

7. Paiement B, Pelletier C, Dyrda I, et al. A simple classification of the risk in cardiac surgery. Can Anaesth Soc J. 1983;30(1):61-8

8. Parsonnet V, Dean D, Bernstein A. A method of uniform stratification of risk for evaluating the results of surgery in acquired adult heart disease. Circulation. 1989;79(6 Pt 2):I3-12

9. Nilsson J, Algotsson L, Höglund P, et al. Comparison of 19 pre-operative risk stratification models in open-heart surgery. European Heart Journal. 2006;27(7):867-74.

10. Asimakopoulos G, Al-Ruzzeh S, Ambler G, et al. An evaluation of existing risk stratification models as a tool for comparison of surgical performances for coronary artery bypass grafting between institutions. European Journal of Cardio-thoracic surgery. 2003;23(6):935-42.

11. Geissler HJ, Hölzl P, Marohl S, et al. Risk stratification in heart surgery: comparison of six score systems. European Journal of Cardio-thoracic surgery. 2000;17(4):400-6.

12. Omar RZ, Ambler G, Royston P, et al. Cardiac surgery risk modeling for mortality: a review of current practice and suggestions for improvement. Annals of thoracic surgery. 2004;77(6):2232-7

13. Prins C, Jonker IDV, Smit FE, et al. Cardiac surgery risk-stratification models. Cardiovascular journal of Africa. 2012;23(3):160

14. Thalji NM, Suri RM, Greason KL, et al. Risk assessment methods for cardiac surgery and intervention. Nature Reviews Cardiology. 2014;11(12):704-14.

15. Head SJ, Osnabrugge RL, Howell NJ, et al. A systematic review of risk prediction in adult cardiac surgery: considerations for future model development. European Journal of Cardio-Thoracic Surgery. 2013;43(5):e121-9.

16. Moons KGM, de Groot JAH, Bouwmeester W, et al. Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies: The CHARMS Checklist. PLOS Medicine. 2014;11(10):e1001744.

17. Ingui BJ, Rogers MAM. Searching for Clinical Prediction Rules in Medline. Journal of the American Medical Informatics Association. 2001;8(4):391-7.

18. Moher D, Liberati A, etzlaff J, et al. The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med. 2009;6(7):e1000097.

19. Vittinghoff E, McCulloch CE. Relaxing the Rule of Ten Events per Variable in Logistic and Cox Regression. American Journal of Epidemiology. 2007;165(6):710-8.

20. Concato J, Peduzzi P, Holford TR, et al. Importance of events per independent variable in proportional hazards analysis I. Background, goals, and general strategy. Journal of Clinical Epidemiology. 1995;48(12):1495-501.

21. Antunes PE, Eugénio L, de Oliveira JF, et al. Mortality risk prediction in coronary surgery: a locally developed model outperforms external risk models. Interactive cardiovascular and thoracic surgery. 2007;6(4):437-41

22. Berg K, Stenseth R, Pleym H, et al. Mortality risk prediction in cardiac surgery: comparing a novel model with the EuroSCORE. Acta Anaesthesiologica Scandinavica. 2011;55(3):313-21

23. Bernstein AD, Parsonnet V. Bedside estimation of risk as an aid for decision-making in cardiac surgery. The Annals of thoracic surgery. 2000;69(3):823-8

24. Billah B, Huq MM, Smith JA, et al. AusSCORE II in predicting 30-day mortality after isolated coronary artery bypass grafting in Australia and New Zealand. The Journal of thoracic and cardiovascular surgery. 2014;148(5):1850-5. e2

25. Billah B, Reid CM, Shardey GC, et al. A preoperative risk prediction model for 30-day mortality following cardiac surgery in an Australian cohort. European Journal of Cardio-Thoracic Surgery. 2010;37(5):1086-92

26. Bridgewater B, Neve H, Moat N, et al. Predicting operative risk for coronary artery surgery in the United Kingdom: a comparison of various risk prediction algorithms. Heart. 1998;79(4):350-5

27. Carosella VC, Navia JL, Al-Ruzzeh S, et al. The first Latin-American risk stratification system for cardiac surgery: can be used as a graphic pocket-card score. Interactive cardiovascular and thoracic surgery. 2009;9(2):203-8

28. Cheng MP, Osuna PP, Santos JG, et al. Designing a cardiac surgery mortality risk model with spanish population. Intensive Care Medicine Experimental. 2015;3(S1):1-

29. Chong C-F, Li Y-C, Wang T-L, et al., editors. Stratification of adverse outcomes by preoperative risk factors in coronary artery bypass graft patients: an artificial neural network prediction model. AMIA; 2003.

30. D'Errigo P, Seccareccia F, Rosato S, et al. Comparison between an empirically derived model and the EuroSCORE system in the evaluation of hospital performance: the example of the Italian CABG Outcome Project. European Journal of Cardio-Thoracic Surgery. 2008;33(3):325-33

31. Eagle KA, Guyton RA, Davidoff R, et al. ACC/AHA Guidelines for Coronary Artery Bypass Graft Surgery: Executive Summary and Recommendations A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee to Revise the 1991 Guidelines for Coronary Artery Bypass Graft Surgery). Circulation. 1999;100(13):1464-80

32. Edwards FH, Albus RA, Zajtchuk R, et al. A quality assurance model of operative mortality in coronary artery surgery. The Annals of thoracic surgery. 1989;47(5):646-9

33. Fortescue EB, Kahn K, Bates DW. Development and validation of a clinical prediction rule for major adverse outcomes in coronary bypass grafting. The American journal of cardiology. 2001;88(11):1251-8

34. Gabrielle F, Roques F, Michel P, et al. Is the Parsonnet's score a good predictive score of mortality in adult cardiac surgery: assessment by a French multicentre study. European journal of cardio-thoracic surgery. 1997;11(3):406-14

35. Grover FL, Johnson RR, Marshall G, et al. Factors predictive of operative mortality among coronary artery bypass subsets. The Annals of thoracic surgery. 1993;56(6):1296-307

36. Hammermeister K, Johnson R, Marshall G, et al. Continuous assessment and improvement in quality of care. A model from the Department of Veterans Affairs Cardiac Surgery. Annals of surgery. 1994;219(3):281

37. Hannan EL, Wu C, Bennett EV, et al. Risk stratification of in-hospital mortality for coronary artery bypass graft surgery. Journal of the American College of Cardiology. 2006;47(3):661-8

38. Higgins TL, Estafanous FG, Loop FD, et al. Stratification of morbidity and mortality outcome by preoperative risk factors in coronary artery bypass patients: a clinical severity score. Jama. 1992;267(17):2344-8

39. Huijskes RV, Rosseel PM, Tijssen JG. Outcome prediction in coronary artery bypass grafting and valve surgery in the Netherlands: development of the Amphiascore and its comparison with the Euroscore. European journal of cardio-thoracic surgery. 2003;24(5):741-9

40. Keogh B, Kinsman R. The Society of Cardiothoracic Surgeons of Great Britain and Ireland. Fifth national adult cardiac surgical database report. 2003

41. Keogh BE, Kinsman R. National adult cardiac data base report 1990–2000. The Society of Cardiothoracic Surgeons of Great Britain and Ireland; 2000.

42. Lippmann RP, Shahian DM. Coronary artery bypass risk prediction using neural networks. The Annals of thoracic surgery. 1997;63(6):1635-43

43. Magovern JA, Sakert T, Magovern GJ, et al. A model that predicts morbidity and mortality after coronary artery bypass graft surgery. Journal of the American College of Cardiology. 1996;28(5):1147-53

44. Marshall G, Grover FL, Henderson WG, et al. Assessment of predictive models for binary outcomes: an empirical approach using operative death from cardiac surgery. Stat Med. 1994;13(15):1501-11.

45. Mejía OA, Lisboa LA, Puig LB, et al. InsCor: a simple and accurate method for risk assessment in heart surgery. Arquivos brasileiros de cardiologia. 2013;100(3):246-54

46. Mozes B, Olmer L, Galai N, et al. A national study of postoperative mortality associated with coronary artery bypass grafting in Israel. The Annals of thoracic surgery. 1998;66(4):1254-62

47. Nashef SA, Roques F, Michel P, et al. European system for cardiac operative risk evaluation (EuroSCORE). European journal of cardio-thoracic surgery. 1999;16(1):9-13

48. Nashef SA, Roques F, Sharples LD, et al. Euroscore II. European journal of cardio-thoracic surgery. 2012;41:12

49. Nilsson J, Ohlsson M, Thulin L, et al. Risk factor identification and mortality prediction in cardiac surgery using artificial neural networks. The Journal of thoracic and cardiovascular surgery. 2006;132(1):12-9. e1

50. O'Connor GT, Plume S, Olmstead E, et al. Multivariate prediction of in-hospital mortality associated with coronary artery bypass graft surgery. Northern New England Cardiovascular Disease Study Group. Circulation. 1992;85(6):2110-8

51. Pitkänen O, Niskanen M, Rehnberg S, et al. Intra-institutional prediction of outcome after cardiac surgery: comparison between a locally derived model and the EuroSCORE. European journal of cardio-thoracic surgery. 2000;18(6):703-10

52. Pons J, Granados A, Espinas J, et al. Assessing open heart surgery mortality in Catalonia (Spain) through a predictive risk model. European journal of cardio-thoracic surgery. 1997;11(3):415-23

53. Ranucci M, Castelvecchio S, Menicanti L, et al. Risk of assessing mortality risk in elective cardiac operations age, creatinine, ejection fraction, and the law of parsimony. Circulation. 2009;119(24): 3053-61

54. Reid C, Billah B, Dinh D, et al. An Australian risk prediction model for 30-day mortality after isolated coronary artery bypass: the AusSCORE. The Journal of thoracic and cardiovascular surgery. 2009;138(4):904-10

55. Roques F, Gabrielle F, Michel P, et al. Quality of care in adult heart surgery: proposal for a self-assessment approach based on a French multicenter study. European journal of cardio-thoracic surgery. 1995;9(8):433-40

56. Sanon S, Lee V-V, Elayda MA, et al. Predicting Early Death after Cardiovascular Surgery by Using the Texas Heart Institute Risk Scoring Technique (THIRST). Texas Heart Institute Journal. 2013;40(2):156

57. Shahian DM, O'Brien SM, Filardo G, et al. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 1—coronary artery bypass grafting surgery. The Annals of thoracic surgery. 2009;88(1):S2-S22

58. Singh M, Gersh BJ, Li S, et al. Mayo Clinic risk score for percutaneous coronary intervention predicts in-hospital mortality in patients undergoing coronary artery bypass graft surgery. Circulation. 2008;117(3):356-62

59. Tu JV, Jaglal SB, Naylor CD. Multicenter validation of a risk index for mortality, intensive care unit stay, and overall hospital length of stay after cardiac surgery. Circulation. 1995;91(3):677-84

60. Verduijn M, Rosseel PM, Peek N, et al. Prognostic bayesian networks: II: An application in the domain of cardiac surgery. Journal of biomedical informatics. 2007;40(6):619-30

61. Wong DT, Cheng DC, Kustra R, et al. Risk Factors of Delayed Extubation, Prolonged Length of Stay in the Intensive Care Unit, and Mortality in Patients Undergoing Coronary Artery Bypass Graft with Fast-track Cardiac Anesthesia A New Cardiac Risk Score. The Journal of the American Society of Anesthesiologists. 1999;91(4):936-

62. Wouters SC, Noyez L, Verheugt FW, et al. Preoperative prediction of early mortality and morbidity in coronary bypass surgery. Vascular. 2002;10(5):500-5

63. Aktuerk D, McNulty D, Ray D, et al. National administrative data produces an accurate and stable risk prediction model for short-term and 1-year mortality following cardiac surgery. International Journal of Cardiology. 2016;203:196-203.

64. Kötting J, Beckmann A, Döbler K, et al. German CABG Score: A Specific Risk Model for Patients Undergoing Isolated Coronary Artery Bypass Grafting. The Thoracic and Cardiovascular Surgeon. 2014;62(4):12.

65. Miyata H, Tomotaki A, Motomura N, et al. Operative Mortality and Complication Risk Model for All Major Cardiovascular Operations in Japan. The Annals of Thoracic Surgery. 2015;99(1):130-9.

66. Motomura N, Miyata H, Tsukihara H, et al. First Report on 30-day and Operative Mortality in Risk Model of Isolated Coronary Artery Bypass Grafting in Japan. The Annals of Thoracic Surgery. 2008;86(6):1866-72.

67. Shroyer ALW, Plomondon ME, Grover FL, et al. The 1996 coronary artery bypass risk model: The Society of Thoracic Surgeons Adult Cardiac National Database. The Annals of Thoracic Surgery. 1999;67(4):1205-8.

68. Shroyer ALWP, Grover FLMD, Edwards FHMD. 1995 Coronary Artery Bypass Risk Model: The Society of Thoracic Surgeons Adult Cardiac National Database. The Annals of Thoracic Surgery. 1995;65(3):879-84.

69. Zheng Z, Zhang L, Li X, et al. SinoSCORE: a logistically derived additive prediction model for post-coronary artery bypass grafting in-hospital mortality in a Chinese population. Front Med. 2013;7(4):477-85.

70. Chung PJ, Carter TI, Burack JH, et al. Predicting the risk of death following coronary artery bypass graft made simple: a retrospective study using the American College of Surgeons National Surgical Quality Improvement Program database. Journal of cardiothoracic surgery. 2015;10:62

71. Hannan EL, Farrell LS, Wechsler A, et al. The New York Risk Score for In-Hospital and 30-Day Mortality for Coronary Artery Bypass Graft Surgery. The Annals of Thoracic Surgery. 2013;95(1):46-52.

72. Swinkels B, Plokker H. Evaluating operative mortality of cardiac surgery: first define operative mortality. Netherlands Heart Journal. 2010;18(7):344-5.

73. Daley J. Criteria by which to evaluate risk-adjusted outcomes programs in cardiac surgery. The Annals of thoracic surgery. 1994;58(6):1827-35.

74. Osswald BR, Blackstone EH, Tochtermann U, et al. The meaning of early mortality after CABG. European journal of cardio-thoracic surgery. 1999;15(4):401-7.

75. Iezzoni LI. The risks of risk adjustment. Jama. 1997;278(19):1600-7.

76. Hannan EL, Racz MJ, Jollis JG, et al. Using Medicare claims data to assess provider quality for CABG surgery: does it work well enough? Health services research. 1997;31(6):659.

77. Concato J, Feinstein AR, Holford TR. The risk of determining risk with multivariable models. Annals of internal medicine. 1993;118(3):201-10.

78. Harrell F. Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis. Nashville, TN USA: Springer; 2015.

79. Karim MN, Reid CM, Tran L, et al. Missing Value Imputation Improves Mortality Risk Prediction Following Cardiac Surgery: An Investigation of an Australian Patient Cohort. Heart, Lung and Circulation. 2016.

80. Jacob M, Bruegger D, Conzen P, et al. Development and validation of a mathematical algorithm for quantifying preoperative blood volume by means of the decrease in hematocrit resulting from acute normovolemic hemodilution. Transfusion. 2005;45(4):562-71.

81. Karim MN, Reid CM, Cochrane A, et al. When is 'Urgent'Really Urgent and Does it Matter? Misclassification of Procedural Status and Implications for Risk Assessment in Cardiac Surgery. Heart, Lung and Circulation. 2016;25(2):196-203.

82. Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. Statistics in Medicine. 2006;25(1):127-41.

83. Altman DG. Categorizing continuous variables. British Journal of Cancer. 1991;64:1.

84. Sauerbrei W, Royston P. Building multivariable prognostic and diagnostic models: transformation of the predictors by using fractional polynomials. Journal of the Royal Statistical Society: Series A (Statistics in Society). 1999;162(1):71-94.

85. Steyerberg E. Clinical prediction models: a practical approach to development, validation, and updating. New York: Springer Science & Business Media; 2009.

86. Austin PC, Tu JV. Automated variable selection methods for logistic regression produced unstable models for predicting acute myocardial infarction mortality. Journal of clinical epidemiology. 2004;57(11):1138-46.

87. Breiman L. Bagging predictors. Machine learning. 1996;24(2):123-40

88. Sauerbrei W, Schumacher M. A bootstrap resampling procedure for model building: application to the Cox regression model. Statistics in medicine. 1992;11(16):2093-109.

89. Hosmer DW, Lemeshow S. Applied logistic regression. Hoboken, New Jercy; 2013.

90. Kuppermann N, Willits N. In response to "Statistical Models and Occam's Razor". Acad Emerg Med. 2000;7(1):4

91. Cole TJ. Statistical Algorithm; Scaling and rounding regression-coefficients to integers. Journal of the royal statistical society 1993;42(1):261-8

92. Altman DG, Royston P. What do we mean by validating a prognostic model? Statistics in Medicine. 2000;19(4):453-73.

93. Steyerberg EW, Moons KGM, van der Windt DA, et al. Prognosis Research Strategy (PROGRESS) 3: Prognostic Model Research. PLOS Medicine. 2013;10(2):e1001381.

94. D'Agostino Sr RB, Grundy S, Sullivan LM, et al. Validation of the Framingham coronary heart disease prediction scores: results of a multiple ethnic groups investigation. Jama. 2001;286(2):180-7.

95. Steyerberg E, Eijkemans M, Harrell Jr F, et al. Prognostic modelling with logistic regression analysis: a comparison of selection and estimation methods in small data sets. Statistics in medicine. 2000;19(8):1059.

96. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for some traditional and novel measures. Epidemiology (Cambridge, Mass). 2010;21(1):128-38.

97. Cook NR. Statistical Evaluation of Prognostic versus Diagnostic Models: Beyond the ROC Curve. Clinical Chemistry. 2008;54(1):17-23.

98. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation. 2007;115(7):928-35

99. Pepe MS, Feng Z, Huang Y, et al. Integrating the predictiveness of a marker with its performance as a classifier. American journal of epidemiology. 2008;167(3):362-8

## 3.3 Summary

The review revealed that there are aspects in current risk modelling, where there is room for improvement to reflect current clinical practice and changing patient demography. Subsequent chapters (chapter four to Chapter six) in the thesis have addressed the methodological issues identified in the review. Chapter seven incorporated all the issues identified in the review and knowledge gathered in chapters four, five and six in developing novel risk prediction model.

# Chapter 4: Variable definition and performance of risk prediction model

### 4.1 Introduction

Risk-prediction models rely on the quality of the databases from which they are developed. Incomplete and inaccurate data may result in overestimation or underestimation of surgical risk. Inaccurate or ambiguous definitions make the predictor variables prone to misclassification (81). Predictors with considerable ambiguity or inter-observer variability are prone to misclassification. Due to misclassification, the concerning predictor may lose its ability to predict a patient's risk accurately, which, in turn, can decrease model performance (31). A misclassified variable may even create an opportunity for gaming, since benchmarking is particularly sensitive to the misclassification of key predictors (31).

This chapter aims to determine the extent of misclassification of 'clinical status'—a significant predictor of mortality—alleged to endure misclassification. The chapter also aims to assess the impact that clinical status misclassification can have on estimates of 30-day mortality risk. This chapter includes the peer reviewed article entitled "When is 'Urgent' Really Urgent and Does It Matter? Misclassification of Procedural Status and Implications for Risk Assessment in Cardiac Surgery." which has been published in 'Heart Lung and Circulation'.

## 4.2 Manuscript

**Declaration for Thesis Chapter 4**

**Manuscript: Karim MN,** Reid CM, Cochrane A, Tran L, Billah B. When is 'Urgent' Really Urgent and Does it Matter? Misclassification of Procedural Status and Implications for Risk Assessment in Cardiac Surgery. Heart, Lung and Circulation. 2016;25(2):196-203.

**Declaration by candidate**

In the case of Chapter 4, the nature and extent of my contribution to the work was the following:

| Nature of contribution | Extent of contribution (%) |
|---|---|
| Study concept and design, literature search, data analysis and interpretation, manuscript development and preparation | 75% |

The following co-authors contributed to the work. Where co-authors are students at Monash University, the extent of their contribution in percentage terms is stated:

| Name | Nature of contribution | Extent of contribution (%) for student co-authors only |
|---|---|---|
| Billah, B | Study concept, data analysis and manuscript editing | |
| Reid, C | Study concept, study design and manuscript editing | |
| Cochrane, A | Study concept, Interpretation of result and manuscript editing | |
| Tran, L | Interpretation of result and manuscript editing | |

The undersigned hereby certify that the above declaration correctly reflects the nature and extend o the candidate's and co-author' contributions to this work

| | | |
|---|---|---|
| **Candidate's Signature** | [redacted] | **Date 10/03/2017** |
| **Main Supervisor's Signature** | [redacted] | **Date 10/03/2017** |

Heart, Lung and Circulation (2016) **25**, 196–203
1443-9506/04/$36.00
http://dx.doi.org/10.1016/j.hlc.2015.07.003

ORIGINAL ARTICLE

# When is 'Urgent' Really Urgent and Does it Matter? Misclassification of Procedural Status and Implications for Risk Assessment in Cardiac Surgery

**Md N Karim, MBBS, M Clin Epi [a], Christopher M Reid, PhD [b,c], Andrew Cochrane, MD, FRCS [d], Lavinia Tran, PhD [b], Baki Billah, PhD [a*]**

[a]School of Public Health and Preventive Medicine, Monash University, The Alfred Centre, Melbourne, Vic, Australia
[b]CCRE Therapeutics, School of Public Health & Preventive Medicine, Monash University, Melbourne, Vic, Australia
[c]School of Public Health, Curtin University, Perth, WA, Australia
[d]Department of Cardiothoracic Surgery and Department of Surgery, Monash Medical Centre, Melbourne, Vic, Australia

*Received 28 May 2015; accepted 6 July 2015; online published-ahead-of-print 18 July 2015*

**Background** Many patients classified as ''urgent'' in Australia New Zealand Society of Cardiac and Thoracic Surgeons (ANZSCTS) registry contradict the prescribed definition (surgery within 72 hours of angiogram or unplanned admission). The aim was to examine the impacts of this misclassification on the prediction of 30-day mortality following cardiac surgery.

**Methods** The 'reported clinical status' was compared with a 'corrected clinical status' following reclassification based on the standard definition calculated from raw data. Observed-to-predicted risk ratios (OPRs) of 30-day mortality were calculated for the model using reported status and corrected status and compared. A Bland-Altman plot was generated to examine the level of agreement between the two OPRs.

**Results** Of 18496 cases reported as urgent, 49.9% were operated after 72 hours, leading to misclassification of 14.6% in the registry. Misclassified patients had significantly higher mortality (3.5%) than true urgent patients (2.9%). Underweight (OR:1.6,CI:1.2-2.1), dialysis (OR:1.4,CI:1.1-1.7), endocarditis (OR:2.1,CI:1.7-2.5), shock (OR:1.6,CI:1.3-2.0) and poor ejection fraction (OR:1.2,CI:1.1-1.4) were significant predictors of misclassification. Bland- Altman plot demonstrates significant disagreement between two risk estimates (P<0.001). Misclassification results in overestimation of risk by 9.1%. Observed-to-predicted risk increased with corrected definition (0.8975 vs 0.9875), suggesting poorer calibration with reported status.

**Conclusions** In the ANZSCTS database, misclassification prevalence is 14.6%. Misclassification compromises the discrimination capacity and calibration of the model and results in overestimation of mortality risk.

**Keywords** Clinical status • Misclassification • Global model • Risk prediction • Cardiac surgery • 30-day-mortality

*Corresponding author at: Department of Epidemiology and Preventive Medicine, Faculty of Medicine Nursing and Health Science, Monash University, The Alfred Centre, (level 6), 99 Commercial Road, VIC 3004. Tel.: +61 (0) 3 9903 0160; fax: +61 (0) 3 99030828, Email: [redacted]

# Introduction

The Australian and New Zealand Society of Cardiothoracic Surgery (ANZSCTS) Database Program was established in 2001 with the aim of reporting risk adjusted clinical outcomes for patients undergoing surgery. The registry has developed a number of risk prediction equations that have been shown to provide the best available estimates of pre-operative risk which enables confidence in benchmarking performance at a national and international level [1–4]. Fundamental to the process of establishing the registry was the development and agreement of a standard data set and definitions to be used by all centres participating in the program [5,6]. The variable 'Clinical status' captures data relating to the clinical urgency of a patient (whether Elective, Urgent, Emergency, or Salvage).

It has been observed by the ANZSCTS registry that some cases in the database did not meet the criteria of ''urgent'' because surgery was undertaken more than 72 hours after an angiogram. This misclassification of clinical urgency first surfaced in the report, Victorian Cardiac Surgery Database Program Public Report 2009–2010 [7]. As clinical status classification is one of the major outcome predictor variables in the risk prediction models developed from the database, the misclassification of urgent cases has the potential to affect the prediction of mortality.

We hypothesised that patients classified as ''Urgent'' but where surgery was undertaken more than 72 hours after an angiogram or after unplanned admission, would represent a stable, lower risk group, and that these patients were better classified as elective cases. The aim of the current research is to a) determine the extent of misclassification of ''Urgency''; b) to identify the predictors of urgent status misclassification; and c) to assess its impact on estimates of 30-day mortality risk.

# Material and Methodology

The ANZSCTS database is a large, multicentre registry which has been collecting data for 14 years. Currently, 28 cardiac hospitals across Australia are contributing data on surgical procedures into the registry. The database consists of 287 pre-operative, intra-operative and post-operative variables. Data elements were defined and adapted from internationally standardised data definitions [5,6]. The index outcome of 30-day-mortality is defined as death within 30 days post-procedure. The database contains all information of patients, who had cardiac surgery during 1 July 2001 to 2013, from the participating centres over their period of involvement. The institutional review board of each participating hospital had approved the use of these databases for research; hence, the need for individual patient consent was waived for this study. The study received ethical approval from Monash University, Standing Committee on Ethics in Research Involving Humans (SCERH).

## Definition of 'Clinical Status'

In the ANZSCTS Data definition manual [5] 'clinical status' has been categorised into Elective, Urgent, Emergency and Salvage. *Elective* refers to the status where the procedure could be deferred without increased risk of compromised cardiac outcome. *Urgent* refers to the status where the procedure is not routine, there is a medical reason for operating this admission, a) within 72 hours from angiography if on the same admission that angiography was performed OR b) within 72 hours after an unplanned admission. *Emergency* refers to unscheduled surgery required in next available theatre on the same day due to refractory angina or cardiac compromise. *Salvage* refers to the status where the patient is undergoing CPR en-route to the operating room, that is, prior to surgical incision. Clinical status is recorded as a check-box entry on a web-based entry system or data record form at the time of the procedure. Misclassification of urgent clinical status was calculated by determining the difference between the time of admission and the time of the procedure recorded in the database. Those procedures which were check box recorded as urgent but had a calculated surgery time greater than 72 hours following catheterisation or unplanned admission were identified as misclassified. The data is presented as 'Reported' versus 'Corrected' clinical status.

## Statistical Methods

a) Extent of misclassification. Descriptive statistics were generated to determine the extent of the misclassification.
b) Predictors of misclassification. The association of relevant pre-operative characteristics to misclassification was investigated through cross-tabulation and chi-square analysis. Predictors of misclassification among reported urgent cases were investigated using multiple logistic regression analysis.
c) Impact on estimates of 30-day mortality. The 30-day-mortality risk was re-estimated with all procedure 30-day mortality risk prediction model for cardiac surgery (global model) 4 using both reported and corrected definitions of urgency. Predicted mortality estimates were calculated separately with reported and corrected definitions of 'clinical status'. Observed-to-predicted risk ratios (OPRs) of mortality were calculated for the models with reported and corrected definitions of clinical status. Percentage change of OPR following reclassification of cases was assessed. A Bland-Altman plot [8] was generated to evaluate the agreement between the two OPRs. The 95% limits of agreement for each comparison (average difference $\pm$ 1.96 x standard deviation of the difference) were computed. The difference between the OPR was then regressed on the average of the two risk ratios. Both the risk ratios were then stratified into categories of each variable in the existing all procedures model.
d) Statistical software packages Stata (version 12) [9] and Medcalc 6.1 [10], where appropriate, were used for all analyses.

# Results

a. Rates of misclassification. A total of 41,813 cases (65.8%) were reported to have elective procedures however, using the corrected definition the number was 51,044 (80.3%). Among the reported urgent cases (n=18,496), 9265 (50.1%) had surgery by 72 hours, (Table 1) leaving 49.1% of those by definition who should have been classified as elective. Rates of misclassification (had been admitted more than 72 hours by the time of surgery) were higher in patients with cardiogenic shock (48.2%), preoperative dialysis (26.9%), endocarditis (39.4%), and underweight (BMI<18.5) patients (25.1%). Among patients who died within 30 days of the operation, 25.8% were misclassified as urgent and should by the strict definition have been reclassified as Elective (Figure 1).

b. Factors associated with misclassification of urgency. Multiple logistic regression among reported urgent cases (n=18496) identified the determinants of urgent status misclassification (Table 2). Underweight (OR: 1.6, CI: 1.2-2.1), preoperative dialysis (OR: 1.4, CI: 1.1-1.7), infective endocarditis (OR: 2.1, CI: 1.7- 2.5), cardiogenic shock (OR: 1.6, CI: 1.3-2.0), ejection fraction <30 (OR: 1.2, CI: 1.1-1.4) and surgery other than CABG or valve (OR: 1.3, CI: 1.1-1.4) were significant predictors of misclassification among reported urgent cases. Strict application of the definition of 'Urgent' would re-categorise these patients as 'Elective'. The misclassification was less likely among patients aged over 80 years (OR: 0.8, CI: 0.7-0.9), diabetic on insulin (OR: 0.9, CI: 0.7-0.9), hypertensive (OR: 0.8, CI: 0.7-0.8), NYHA class III (OR: 0.9, CI: 0.8-0.9), previous myocardial infarction (OR: 0.9, CI: 0.8-0.9), valve operation (OR: 0.7,CI: 0.6, 0.8) and combined CABG+valve operation (OR: 0.8, CI: 0.7-0.9).

c. Impact on risk prediction. The Bland-Altman plot demonstrates that the mean risk difference of the two OPR significantly differs from the null value, (P<0.001), indicating the presence of a fixed bias. The 95% limits of agreement for each comparison demonstrate wide variation in the OPRs generated by the model with reported and corrected definitions of clinical status. The average OPR has a significant linear relation with risk difference (β=-0.0013, P<0.001) suggesting disagreement between the two risk ratios. (Figure 2) Stratification of risk by separate Bland-Altman plots for age, gender and procedure type didn't reveal any specific pattern of difference across gender, age group and procedure type, ruling out confounding or effect modification effects of these factors on the discrepancy in two OPRs. Paradoxically, the misclassified patients had significantly higher mortality (3.5%) than the patients who satisfied the time definition of urgency (2.9%) (P<0.001). Mortality risk in the reported urgent category (OR: 2.0, CI: 1.8-2.3) was higher in comparison to the corrected urgent category (OR: 1.5, CI: 1.3-1.7) (Table 3). After correction of the clinical status of urgency, predicted mortality in the new ''Urgent'' category decreased from 3.14% to 2.86%, confirming an overestimation in reported status by 9.1%. The new classification with strict definition based on time actually moved several categories of high-risk patients into the ''Elective'' category, reducing the mortality in the corrected ''Urgent'' category. Observed-to-predicted risk ratios increased when the corrected definition was used (reported: 0.8975, corrected: 0.9875) (Table 4). In general, calibration is perfect when OPR approaches to 1, OPR approaches closer to 1 for corrected definition in comparison to reported status, suggesting poorer calibration with reported status.

**Table 1 Cross tabulation of reported clinical status and reclassified clinical Status.**

| Reported clinical status | Reclassified clinical Status | | | |
|---|---|---|---|---|
| | Elective | Urgent | Emergency/ salvage | Total |
| Elective | 41,813 (100) | 0 (0) | 0 (0) | 41,813 (65.8) |
| Urgent | 9,231 (49.9) | 9,265 (50.1) | 0 (0) | 18,496 (29.1) |
| Emergency/ salvage | 0 (0) | 0 (0) | 3,240(100) | 3240 (5.1) |
| Total | 51,044 (80.3) | 9,265 (14.6) | 3240 (5.1) | 63,549 (100) |

# Discussion

This analysis reveals a large number of patients (14.6%) who are identified at the site as being urgent but whose surgery is not undertaken within 72 hours of angiography or an unplanned admission. This also appears to be a heterogeneous group with a proportion of a) very high risk patients with dialysis, cardiogenic shock, and endocarditis whose surgery is delayed and b) a clinical low risk group who had been check-boxed as being urgent. This sub-group of ''delayed'' surgical patients had poorer outcomes (death within 30 days post operatively) than all patients undergoing surgery within 72 hours. However, when pooled with the low risk misclassified patients, overall risk in the misclassified group was lower than those having surgery with 72 hours.

It is worth stressing that the use of a 72-hour cut-off is fairly arbitrary, and might harbour the grounds for misclassification. Society of Thoracic Surgeons (STS) [11] registry didn't use such a timeframe for defining urgent. They defined the urgent based on conditions that require the patient to remain in the hospital until surgery can take place, but the patient is able to wait for surgery until the next available schedule time. Definition of elective surgery in 'The Australian Institute of

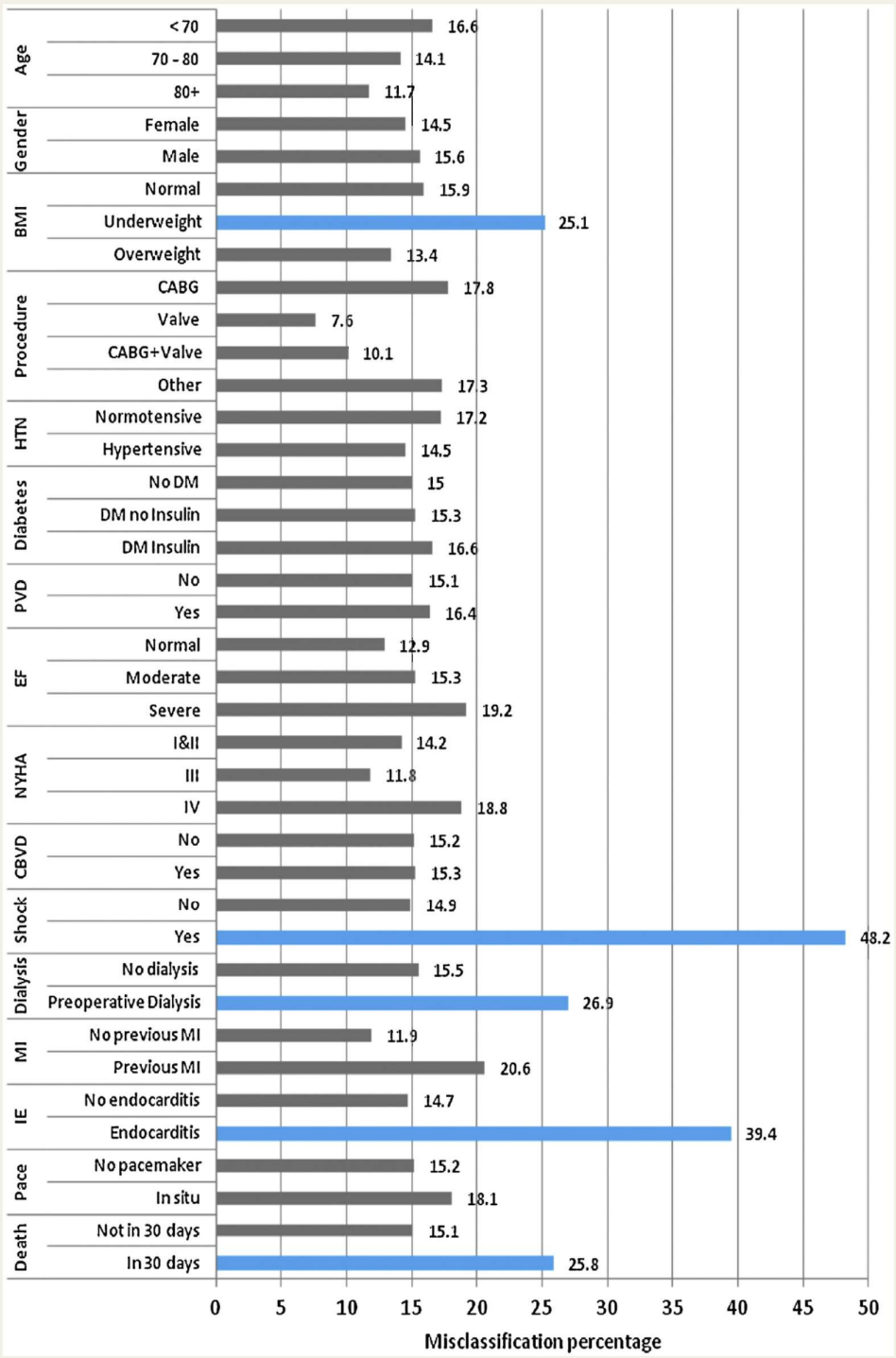



**Figure 1** Proportion of misclassification and preoperative patient characteristics (n=63208).
DM – Diabetes mellitus; MI - Previous myocardia infarction, IE - Infective endocarditis, PVD - Peripheral vascular disease, CVD - cerebro vascular disease, HTN - Hypertension, DIAL - Preoperative Dialysis.

Health and Welfare' national health data dictionary also kept timeframe as a key criteria for demarcating from urgent surgery [12]. However, such a timeframe may not sufficiently describe the exact clinical status of the patient.

Delay in the operation may be necessitated by attempts to improve the patient's condition, a period of medical therapy such as antibiotics for endocarditis, and optimisation of cardiac function. Other reasons for delay might include the

**Table 2** **Determinants of urgent status misclassification among reported urgent cases (n=18496) assessed by binary logistic regression.**

| Factors of misclassification | OR (95% CI) | P value | Referent category |
|---|---|---|---|
| Age 70 - 79 | 0.91 (0.84, 0.98) | 0.010 | *Age< 70 years* |
| Age ≥ 80 | 0.79 (0.70, 0.89) | < 0.001* | *Age< 70 years* |
| Gender (Female) | 1.00 (0.93, 1.08) | 0.929 | *Male* |
| No Medicare | 1.09 (0.93, 1.28) | 0.291 | *Medicare registered* |
| Previous surgery | 1.06 (0.93, 1.21) | 0.352 | *No previous surgery* |
| Peripheral vascular disease (PVD) | 1.06 (0.96, 1.18) | 0.251 | *No PVD* |
| Respiratory disease (LD) | 0.99 (0.91, 1.09) | 0.992 | *No LD* |
| Cerebro-vascular disease (CBVD) | 0.95 (0.86, 1.05) | 0.286 | *No CBVD* |
| Stroke | 0.87 (0.69, 1.08) | 0.199 | *No stroke* |
| Underweight | 1.57 (1.19, 2.10) | 0.002* | *BMI 18.5 - 24* |
| Overweight | 0.99 (0.93, 1.07) | 0.983 | *BMI 18.5 - 24* |
| Diabetes on diet/drug | 1.00 (0.93, 1.08) | 0.959 | *No diabetes* |
| Diabetes on insulin | 0.86 (0.76, 0.96) | 0.010* | *No diabetes* |
| Pre-operative dialysis | 1.39 (1.13, 1.73) | 0.002* | *No pre-operative dialysis* |
| Infective endocarditis (IE) | 2.07 (1.72, 2.49) | < 0.001* | *No pre-operative dialysis* |
| Cardiogenic shock | 1.63 (1.31, 2.01) | < 0.001* | *No IE* |
| Ejection fraction, mild | 0.94 (0.87, 1.01) | 0.080 | *No shock* |
| Ejection fraction (EF), moderate | 1.05 (0.96, 1.16) | 0.281 | *Normal to Mild EF* |
| Ejection fraction (EF), severe | 1.23 (1.08, 1.41) | 0.003* | *Normal to Mild EF* |
| Arrythmia | 1.07 (0.98, 1.17) | 0.149 | *No Arrythmia* |
| Hypertensive | 0.78 (0.72, 0.84) | < 0.001* | *Normotensive* |
| Pacemaker in situ | 1.19 (0.93, 1.52) | 0.166 | *No pacemaker* |
| NYHA class III | 0.87 (0.80, 0.95) | 0.002* | *NYHA class I &II* |
| NYHA class IV | 0.91 (0.81, 1.02) | 0.094 | *NYHA class I &II* |
| Myocardial Infarction (MI) | 0.86 (0.80, 0.92) | < 0.001* | *No MI* |
| Immuno-suppressant use | 1.09 (0.89, 1.35) | 0.406 | *No immunosuppressant* |
| Valve | 0.72 (0.63, 0.83) | < 0.001* | *CABG* |
| CABG+Valve | 0.83 (0.73, 0.93) | 0.002* | *CABG* |
| Other Procedures | 1.27 (1.12, 1.43) | < 0.001* | *CABG* |

*Statistically significant.

availability of results of further investigations, availability of a spouse or parent or guardian for informed consent, and availability of blood products or the availability of results of essential laboratory procedures or tests [13,14].

One of the main reasons for misclassification may be due to the operating surgeon's non-adherence to the ANZSCTS definition of clinical urgency. This non-adherence may be due to the surgeon being unaware of the definition itself or due to the surgeon's preference for any other classification system, or both. Paradoxically, and to our surprise, we found that the 30-day mortality was higher for urgent patients waiting >72 hours than for those <72 hours. The misclassified patients make up 25.8% of the total 30-day mortality. The question arises whether the mortality was higher because they waited a longer period or whether they were a sicker group of patients.

In our data there appear to be at least two groups of patients who have surgery delayed beyond 72 hours. There is a stable group who behave like elective patients and a sicker group with more risk variables, and this may explain why they waited longer if there was a need to optimise the pre-operative medical therapy.

'Clinical urgency status' has been one of the key predictors of 30-day mortality recognised by previous studies [15,16]. We investigated the effect of misclassification on the predicted risk and the performance of global model [4]. From the risk prediction modelling perspective, misclassification of an important predictor is a concern because risk scores are generated based on a beta coefficient of the predictors of the model. The potential of an individual predictor for affecting the model depends on the relative weight of the predictor itself and its level of association with the outcome. In the global model, clinical status was an important predictor both statistically and theoretically. Its association with 30-day mortality was highly significant. Misclassification of this important predictor resulted in an altered level of association

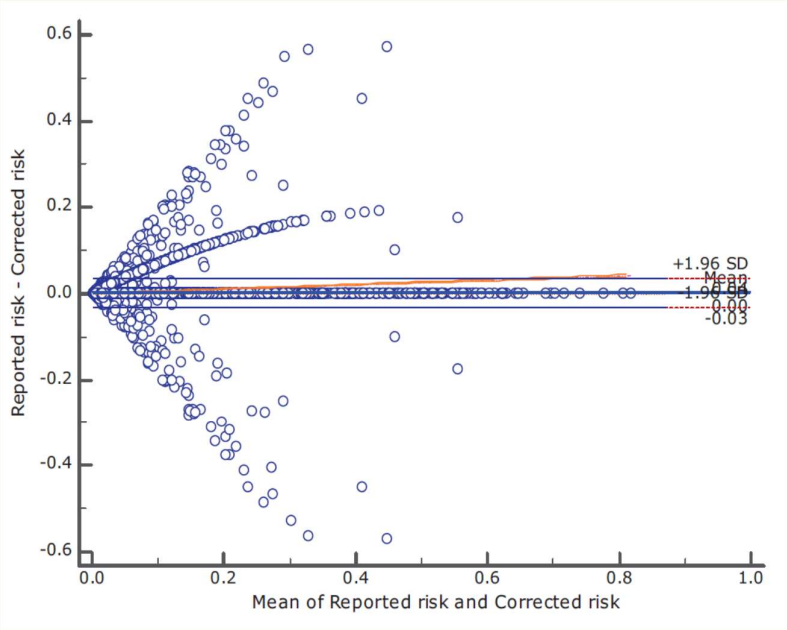



**Figure 2** Bland-Altman plot assessing agreement between predicted and reclassified mortality risk estimates among cardiac surgery patients.

which was reflected in an altered beta coefficient of the predictor as well as other predictors in the model. Altered coefficients of the predictors consequently affected the risk score. In the current analysis, the model generated with reported and corrected clinical status resulted in different beta coefficients and precision estimates.

Running the model for the ANZSCTS database for both before and after reclassification demonstrated that reclassification improved risk prediction by around 10%. This suggests that patients classified into 'urgent <72 hours' were appropriately grouped because the 'clinical status' label added predictive capacity over and above the other known risk variables. By contrast, for the group 'urgent >72 hours' the clinical status label no longer added predictive capacity, suggesting that the existing known risk factors are able to explain almost all of the observed risk. To add in the 'urgent status' for this group would be analogous to 'double counting' because a risk factor is being included twice in different ways.

Put another way, the fact that we need to include 'urgent status' in the ordinary model indicates that the other known and defined risk factors are inadequate to fully explain the observed risk, and the 'urgent' status has some explanatory capacity. Its inclusion is in fact an indicator of the inability of

**Table 3 Comparison of mortality in two definitions of clinical status.**

| Variables | 30-day-mortality | | | |
|---|---|---|---|---|
| | n (%) | n (%) | OR* | P*Value |
| Reported clinical status | | | | |
| Elective | 41,813 (98.4) | 665 (1.6) | Referent | |
| Urgent | 18,496 (96.8) | 590 (3.2) | 2.0 (1.8, 2.3) | <0.001 |
| Emergency/Salvage | 3240 (83.7) | 522 (16.3) | 11.9 (10.6,13.5) | <0.001 |
| Corrected clinical status | | | | |
| Elective | 51,044 (98.0) | 989 (2.0) | Referent | |
| Urgent | 9,265 (97.1) | 266 (2.9) | 1.5 (1.3,1.7) | <0.001 |
| Emergency/Salvage | 3,240 (83.7) | 522 (16.3) | 9.5 (8.7,10.9) | <0.001 |

*OR generated through binary logistic regression.

**Table 4** Effect of misclassification on risk prediction performance and calibration.

| | Risk 1 (Reported) | Risk 2 (Corrected) |
|---|---|---|
| 30-day-mortality risk | | |
| Observed | 0.0282095 | |
| Predicted | 0.0314283 | 0.0285643 |
| Calibration | | |
| Observed to Predicted risk | 0.8975828 | 0.9875789 |
| Change in OPR | 9.1% | |

the model to define all risk factors. If we understood clearly why urgent patients were different, then we could dispense with the need for the 'urgent category' of clinical status variable. From the clinical aspect, it seems important to define a separate 'urgent >72 hours' group, and they should not be included in the elective group because they have a significantly higher mortality than the elective group, they have a different risk factor profile than the elective group, and because they cannot be discharged from hospital or deferred for more than a very short period, hence not meeting the definition of 'elective' patients.

Apart from the prediction itself, misclassification also affects the validity and accuracy of the model [17]. Validation of a risk prediction model requires both calibration [18] and discrimination [19] analyses. The current analysis confirms significant discrepancy in risk prediction performance of the two models generated with the two definitions of clinical status. Further, the model with reported clinical status calibration and discriminatory ability also seems to be compromised.

There should be processes for monitoring categorisation of the patients at an institution level. Whatever the reason for the misclassification, an accurate and uniform system for clinical urgency categorisation is important and it will improve the consistency of decision-making across the surgeons and institutions, improve communication regarding the relative urgency of patients, and improve the allocation of operation theatre resources. Furthermore, a uniform system enables benchmarking across health facilities dealing with cardiac surgery cases [20]. Adequate and regular monitoring and evaluation of the surgical decision should probably be the key strategy for preventing such misclassification.

One limitation of the current analysis is that, the same dataset for calibration and risk estimation was not used from which the original global model was developed. This study uses data that is almost triple the size of the data used by the original global model (23,016 patients). Furthermore, with the advancement of surgical technology and management procedure, a patient population with an altered risk profile is made eligible for surgery. However, altered size and characteristics are unlikely to affect the risk difference as both the models were developed from the same patient population. The standard definition of the clinical status and current data may result in a different set of variables into the model, however investigating the model development process was beyond the remit of the study. The current study limited its aim to assessing the impact on the misclassification on the existing model.

## Conclusion and Recommendation

The current analysis confirms the misclassification of urgent clinical status in 14.6% of the ANZSCTS database. Misclassification of patient to urgent category is more common among patients with cardiogenic shock, preoperative dialysis, endocarditis, and BMI<18.5. Misclassification compromises the discrimination capacity and calibration of the model and results in overestimation of mortality risk in the global model.

This analysis demonstrates that the current simple time-dependent classification of urgent patients is not appropriate, because the misclassified patient group includes several high-risk groups of patients with a high mortality. A purely time-based definition is also inconsistent with the definition used by other large cardiac databases and their models (the STS score and Euroscore). It is clear that many clinicians in practise do not agree with the time-based division between elective and urgent surgery.

We propose a new definition of ''Urgent'' status to include the following categories – (a) Cardiac surgery within 72 hours from angiography, if on the same admission; (b) Cardiac surgery within 72 hours of an unplanned admission; (c) Cardiac surgery for acute valve endocarditis; (d) Cardiac surgery for patients admitted to hospital with cardiogenic shock, or patients with worsening or ongoing chest pain; (e) Cardiac surgery for patients with ejection fraction less than 30% and who have been admitted to hospital before surgery; (f) Surgery for patients on pre-operative dialysis who are admitted to hospital; and (g) Surgery for underweight patients, defined as BMI < 18.5.

## Acknowledgements

All the investigators, data managers and institutions participated in the ANZSCTS Registry.

The study received 'no external financial support' and has no conflict of interest to declare.

## References

[1] Reid C, Billah B, Dinh D, Smith J, Skillington P, Yii M, et al. An Australian risk prediction model for 30-day mortality after isolated coronary artery bypass: The AusSCORE. J Thorac Cardiovasc Surg 2009;138:904–10.

[2] Billah B, Reid CM, Shardey GC, Smith JA. A preoperative risk prediction model for 30-day mortality following cardiac surgery in an Australian cohort. Eur J Cardiothorac Surg 2010;37:1086–92.

[3] Ariyaratne TV, Billah B, Yap C-H, Dinh D, Smith JA, Shardey GC, et al. An Australian risk prediction model for determining early mortality

following aortic valve replacement. Eur J Cardiothorac Surg 2011;39 (6):815–21.
[4] Billah B, Dinh DT, Shardey GC, Reid CM. Preoperative Risk Prediction Model for 30-day Mortality Following Cardiac Surgery in Australia. Heart, Lung and Circ 2008;17 Suppl 3(0):S67.
[5] ANZSCTS National Cardiac Surgery Database Program. Data Definitions Manual. 2008. Version 3.
[6] Reid CM, Rockell M, Skillington PD, Shardey GC, Smith JA, Yii M. Initial Twelve Months Experience and Analysis for 2001–2002 from the Australasian Society of Cardiac and Thoracic Surgeons, Victorian Database Project. Heart, Lung and Circ 2004;13(3):291–7.
[7] Dinh DT, Tran L, Chand V, Ariyaratne T, Newcomb A, Billah B, et al. ASCTS Cardiac Surgery Database. Public report 2009–2010, VIC. Available at http://www.anzscts.org/uploads/39136/ufiles/VIC_Public_report_2009-10.pdf. Accessed on 12 jan 2015
[8] Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;327(8476): 307–10. ion, TX: Stata Corporation; 2004.
[9] StataCorp. Stata Statistical Software, version 9.2 [computer software]. College Stat.
[10] MedCalc. MedCalc Software, version 13.3 [computer software]. MedCalc Software bvba, Acacialaan 22, Ostend, Belgium; 2014.
[11] STS National Database Program. Adult Cardiac Surgery Database Training Manual, v2.73 available at the http://www.sts.org/sites/default/files/documents/Training Manual Update 2012.pdf accessed on 1 Jan 2015.
[12] Australian Institute of Health and Welfare 2008, National health data dictionary, Version 14, available on the web, accessed on 12 Jan 2015.
[13] Hadorn DC, Holmes AC. The New Zealand priority criteria project. Part 2: Coronary artery bypass graft surgery. BMJ 1997;314:135–7.
[14] Kee F, McDonald P, Kirwan JR, Patterson CC, Love AHG. Urgency and priority for cardiac surgery: a clinical judgment analysis. BMJ 1998;316: 925–9.
[15] Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R. European system for cardiac operative risk evaluation (EuroSCORE). Eur J Cardiothorac Surg 1999;16:9–13.
[16] Shahian DM, O'Brien SM, Filardo G, Ferraris VA, Haan CK, Rich JB, et al. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 1–coronary artery bypass grafting surgery. Ann Thorac Surg 2009;88:S2–S22.
[17] Moons KG, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: II. External validation, model updating, and impact assessment. Heart 2012;98(9):691–8.
[18] D'Agostino Sr RB, Grundy S, Sullivan LM, Wilson P. Validation of the Framingham coronary heart disease prediction scores: results of a multiple ethnic groups investigation. JAMA 2001;11(286(2)):180–7.
[19] Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. Circ 2007;20(115(7)):928–35.
[20] Department of Health, Victoria. A framework for emergency surgery in Victorian public health services Available at http://docs.health.vic.gov.au/docs/doc/A-framework-for-emergency-surgery-in-Victorian-public-health-services accessed on 21 Oct 2014

## 4.3 Summary

The study confirms the misclassification of urgent clinical status among patients undergone cardiac surgery, included in the ANZSCTS database. Misclassification was found to compromises the discrimination capacity and calibration of risk prediction models. Major implication of the finding of the study is that the time dependent classification of urgent patients is not discriminative as the misclassified patient group includes several high-risk groups of patients with a high mortality. The paper proposed a new definition of clinical status based on the study finding.

# Chapter 5: Missing value imputation and performance of risk prediction models

## 5.1 Introduction

The systematic review, discussed in Chapter 3, highlighted the necessity for adequate handling of missing data in risk prediction modelling. Risk prediction models are usually developed using data routinely collected in hospitals, general practices or by clinical registries. Such settings for data collection are often prone to missing data. Missing values compromise the quality of the data, and may therefore affect the accuracy of the models that are derived from the data. The potential for missing data to impact on the model development process has mostly been disregarded in risk modelling. Only a few of the risk prediction models that have been developed have handled missing data using an appropriate method.

This chapter aims to assess the impact that missing data values can have on the accuracy of predictions for mortality risk following cardiac surgery, using an existing model as an example. This chapter includes the peer reviewed article entitled "Missing Value imputation improves mortality risk prediction following cardiac surgery: an investigation of an Australian patient cohort" which has been published in the Heart Lung and Circulation.

## 5.2 Manuscript

**Declaration for Thesis Chapter 5**

**Manuscript:** Karim MN, Reid CM, Tran L, Cochrane A, Billah B. Missing Value Imputation Improves Mortality Risk Prediction Following Cardiac Surgery: An Investigation of an Australian Patient Cohort. Heart, Lung and Circulation. 2017;26(3):301-8.

**Declaration by candidate**

In the case of Chapter 5, the nature and extent of my contribution to the work was the following:

| Nature of contribution | Extent of contribution (%) |
|---|---|
| Study concept and design, literature search, data analysis and interpretation, manuscript development and preparation | 75% |

The following co-authors contributed to the work. Where co-authors are students at Monash University, the extent of their contribution in percentage terms is stated:

| Name | Nature of contribution | Extent of contribution (%) for student co-authors only |
|---|---|---|
| Billah, B | Study concept, data analysis and manuscript editing | |
| Reid, C | Study concept, study design and manuscript editing | |
| Cochrane, A | Study concept, Interpretation of result and manuscript editing | |
| Tran, L | Interpretation of result and manuscript editing | |

The undersigned hereby certify that the above declaration correctly reflects the nature and extend o the candidate's and co-author' contributions to this work

| | | |
|---|---|---|
| **Candidate's Signature** | [redacted] | **Date 10/03/2017** |
| **Main Supervisor's Signature** | [redacted] | **Date 10/03/2017** |

Heart, Lung and Circulation (2017) **26,** 301–308
1443-9506/04/$36.00
http://dx.doi.org/10.1016/j.hlc.2016.06.1214

ORIGINAL ARTICLE

# Missing Value Imputation Improves Mortality Risk Prediction Following Cardiac Surgery: An Investigation of an Australian Patient Cohort

**Md Nazmul Karim, MBBS, MClinEpi [a], Christopher M. Reid, PhD [a,b], Lavinia Tran, PhD [a], Andrew Cochrane, MD, FRCS [c], Baki Billah, PhD [a*]**

[a]School of Public Health and Preventive Medicine, Monash University, Melbourne, Vic, Australia
[b]School of Public Health, Curtin University, Perth, WA, Australia
[c]Department of Cardiothoracic Surgery, Monash Medical Centre, Melbourne, Vic, Australia

*Received 15 December 2015; received in revised form 18 June 2016; accepted 27 June 2016; online published-ahead-of-print 8 August 2016*

**Background** The aim of this study was to evaluate the impact of missing values on the prediction performance of the model predicting 30-day mortality following cardiac surgery as an example.

**Methods** Information from 83,309 eligible patients, who underwent cardiac surgery, recorded in the Australia and New Zealand Society of Cardiac and Thoracic Surgeons (ANZSCTS) database registry between 2001 and 2014, was used. An existing 30-day mortality risk prediction model developed from ANZSCTS database was re-estimated using the complete cases (CC) analysis and using multiple imputation (MI) analysis. Agreement between the risks generated by the CC and MI analysis approaches was assessed by the Bland-Altman method. Performances of the two models were compared.

**Results** One or more missing predictor variables were present in 15.8% of the patients in the dataset. The Bland-Altman plot demonstrated significant disagreement between the risk scores ($p<0.0001$) generated by MI and CC analysis approaches and showed a trend of increasing disagreement for patients with higher risk of mortality. Compared to CC analysis, MI analysis resulted in an average of 8.5% decrease in standard error, a measure of uncertainty. The MI model provided better prediction of mortality risk (observed: 2.69%; MI: 2.63% versus CC: 2.37%, $P<0.001$).

**Conclusion** 'Multiple imputation' of missing values improved the 30-day mortality risk prediction following cardiac surgery.

**Keywords** Cardiac surgery • Risk prediction model • Missing data • Multiple imputation

## Background

Risk prediction models for postoperative outcome have become an integral part of cardiac surgical risk assessment and are used for benchmarking quality of care and outcomes. They can also be used to educate and counsel patients as to the preoperative risk associated with the surgery [1]. Risk prediction allows comparison between risks and benefits of the surgery and facilitates evidenced based surgical decision-making [2,3]. Risk prediction models should be precise. To achieve a high level of precision, the model development process should ensure that the predictors are reliably and

*Corresponding author at: Senior Lecturer of Biostatistics School of Public Health and Preventive Medicine, Monash University The Alfred Centre, (level 6), 99 Commercial Road, Melbourne, VIC 3004, Australia. Tel.: +61 3 9903 0160, Email: baki.billah@monash.edu

completely measured. However, irrespective of the design and diligence of those involved in the data collection process, missing data is common in medical research [4].

Missing data compromise the quality of data, and subsequently affect the accuracy of the models derived from the data. Their potential to dent the soundness of research finding has often been disregarded in the medical literature [5]. The issue of missing data is usually addressed by keeping only those individuals who have no missing data in any of the variables required for that analysis (complete cases analysis). In statistical modelling, which typically concerns associations of outcome with several predictors simultaneously, the cumulative effect of missing data in several variables leads to the exclusion of a sizeable proportion of the original sample, which in turn causes a loss of power in risk models [6,7]. Besides, while deleting the incomplete cases, available information of other predictors is lost. Apart from wastage of valuable information, collected with cost and effort, complete case analysis (CC analysis) may lead to biased results [8,9] which subsequently may give rise to an inaccurate prediction of outcome. Hence, missing values should be treated with appropriate method. The need for adequate handling of missing data in medical research is increasingly recognised in recent literature [5].

Filling the gap (imputation) with values (ie, mean, median etc.) generated from the observed data of the same variable [10], is a popular concept for handling missing values. These single imputation approaches may lead to bias, since they do not account for the uncertainty of the missing values. Multiple imputation (MI) is an advanced and robust method of handling missing values. Rubin and colleagues [11] proposed the method several decades ago, however its use remained limited to the field of statistics only because of the lack of computational tools for generating multiple imputations. A notable variety of simulation methods reported in the recent statistical literature has paved the way for its use in medical and other fields.

Multiple imputation accounts for the uncertainty in predicting the missing values. Imputation more than once ensures the randomness of the estimation technique [12]. This technique generates multiple complete datasets, with the missing values replaced by imputed values. These values are the best estimates of missing predictor values generated, based on existing associations between the variables under consideration in the observed data [13]. Each imputed dataset is analysed separately. The parameter estimates of all the imputed datasets are averaged to give an overall estimate [11].

In the area of cardiac surgery, none of the currently used risk prediction models addressed the impact of missing data adequately. The Parsonnet score [14] and modified Parsonnet score [15] dropped predictors which have the potential to generate missing data. The Amphiascore [16] replaced missing values with the most prevalent values. The Pons score [17], Toronto score [18] and UK national score [19] like most prediction models, didn't address the missing value issue. Whilst developing the EuroSCORE I [20] & II [21], cases with incomplete information were excluded from the analysis, assuming it was missing completely at random (MCAR). The Society of Thoracic Surgeons (STS) risk models [22,23] estimated missing values using single imputation. Development of the ANZSCTS risk prediction scores also excluded cases due to missing observations [24–26]. Although the AusSCORE II [27] model incorporated multiple imputation, it didn't assess the impact of imputation on prediction performance of risk models.

The aim of this study was to evaluate the impact of missing values on preoperative risk prediction following cardiac surgery using 30-day mortality as an example. In this study an existing 30-day mortality risk prediction model [25], developed from ANZSCTS database registry using only the complete cases, was re-estimated employing both complete case and multiple imputation approaches to find whether there is any difference in prediction performance.

## Methods

The ANZSCTS database registry collects information on adult patients undergoing cardiac surgery in 28 hospitals across Australia. The database collects preoperative, intraoperative and postoperative variables from each patient undergoing cardiac surgery. Between 2001 and 2014 the database recorded the information of 84,233 patients. Patients with missing information on procedure type were exempted from imputation, as imputation of this particular variable is clinically implausible. Exclusion of the cases with incomplete procedure type information led to a final dataset of 83,309 patient records. A total of 62,737 patients' records between 2001 and 2012 was used for estimation of the model and 20,572 patients' records between 2013 and 2014 were used for external validation of the models. The split resulted in comprising roughly 75% of the records in the estimation set and 25% of the records in the validation set.

Descriptive statistics were generated to assess the pattern and extent of missing-ness. A missing indicator variable was created, where patients with missing information in one or more predictors were categorised as missing. The association of each independent predictor with a missing indicator variable was evaluated using the chi-square test. Multiple imputation of missing values was done using the Imputation by Chained Equations (ICE) method along with multivariable logistic regression. The imputation was repeated five times as suggested by Rubin [11], Schafer and Olsen [28]. The analysis was performed separately on each imputation. The results were then combined into an aggregated MI result.

The regression coefficients of an existing model [25] for predicting the 30-day mortality were estimated on 62,737 patients of the ANZSCTS dataset (a) without (CC: complete cases analysis) and (b) with imputation (MI: multiple imputation analysis). Predicted risk of 30-day mortality was generated for the two risk prediction models (CC model and MI model). The agreement between the predicted risk of MI and CC models was assessed using the Bland-Altman plot which is a graphical method commonly used in medical research to compare two measurement techniques.

Model performance in terms of calibration and discrimination of both the CC and MI models was assessed. Calibration refers to the agreement between observed endpoints and predictions [29]. Calibration was assessed in this study by plotting observed versus expected event proportions within deciles of predicted risk (decile-decile plot) [30]. Deciles of predicted risk were regressed on the deciles of observed mortality proportions. Perfect predictions should be on the diagonal line, described with an intercept of 0 'zero' and slope of 1 'one' [29]. Discrimination refers to the ability of the model to distinguish a patient with the outcome (dead) from a patient without (alive) [31]. The discriminatory power of the two models was assessed by calculating receiver operating characteristics (ROC) from a multifold validation sample. External validation of the two models was performed on a validation dataset comprising patients who underwent operation between 2013 and 2014 (n=20,572).

Statistical software package Stata (version-14), was used for all analyses. The institutional review board of each participating hospital approved the use of these databases for research (Alfred HREC:-262/09). The database approved collection of patient data via opt-out consent approach (MUHREC:-CF08/0322-2008000065). The study received ethical approval from Monash University, Standing Committee on Ethics in Research Involving Humans (SCERH) (MUHREC:-CF14/1117–2014000476).

# Results

## Pattern of Missing Data

The frequency of missing data is shown in Table 1. A total of 9,929 (15.83%) patients had one or more predictors missing. The percentages of missing data for each variable in the existing model are presented in Table 2. Estimated ejection fraction had the highest missing data (11.32%) followed by the NYHA classification (4.53%). The remaining predictors had <1% missing observations. The association of individual predictors with missing indicator variable was reported in Table 3. Age, urgency of procedure, BMI,

**Table 1** Distribution of number of missing observation

| Number of missing | Estimation dataset (n=62,737) | | Validation dataset (n=20,572) | |
|---|---|---|---|---|
| Observations | Frequency | % | Frequency | % |
| 0 | 52,808 | 84.17 | 19,858 | 96.53 |
| 1 | 9,374 | 14.94 | 708 | 3.44 |
| 2 | 377 | 0.60 | 4 | 0.02 |
| 3 | 26 | 0.04 | 1 | <0.01 |
| 4 | 14 | 0.02 | 0 | 0.02 |
| 5 | 25 | 0.04 | 0 | 0.00 |
| 6 | 78 | 0.12 | 0 | 0.00 |
| 7 | 35 | 0.06 | 1 | <0.01 |

**Table 2** Number (%) of missing observations for predictors in the existing model

| Variables | Estimation dataset (n=62,737) | | Validation dataset (n=20,572) | |
|---|---|---|---|---|
| | n | % | n | % |
| Age | 13 | 0.02 | 5 | 0.02 |
| Dialysis | 124 | 0.20 | 2 | 0.01 |
| Hypercholesterolaemia | 172 | 0.27 | 3 | 0.01 |
| BMI | 128 | 0.20 | 51 | 0.25 |
| Peripheral vascular disease | 170 | 0.27 | 2 | 0.01 |
| NYHA classification | 2840 | 4.53 | 9 | 0.04 |
| Ejection fraction | 7100 | 11.32 | 627 | 3.05 |
| Previous surgery | 134 | 0.21 | 0 | 0 |
| Clinical status | 1 | <0.01 | 0 | 0 |
| Ionotropic medication | 157 | 0.25 | 2 | 0.01 |

**Table 3** Association between missing-ness and the variables in the existing model

| Variables & Values | Subject with ≥ 1 missing N (%) | Subject with No missing N (%) | P Value* |
|---|---|---|---|
| Age | | | |
| < 60 years | 3208 (32.35) | 14756 (27.94) | < 0.001 |
| 60 - 69 years | 2905 (29.30) | 15594 (29.53) | |
| 70 - 79 years | 2936 (29.61) | 16705 (31.63) | |
| ≥ 80 years | 867 (8.74) | 5753 (10.89) | |
| Gender | | | |
| Male | 7236 (72.88) | 38035 (72.03) | 0.082 |
| Female | 2693 (27.12) | 14773 (27.97) | |
| Urgency of procedure | | | |
| Elective | 5866 (59.09) | 35659 (67.53) | < 0.001 |
| Urgent | 3389 (34.14) | 14679 (27.80) | |
| Emergency | 673 (6.78) | 2470 (4.68) | |
| BMI | | | |
| < 25 Kg/m2 | 2847 (29.05) | 14507 (27.47) | 0.001 |
| ≥ 25 Kg/m2 | 6954 (70.95) | 38.301 (72.53) | |
| Dialysis | | | |
| No | 9638 (98.30) | 51949 (98.37) | 0.583 |
| Yes | 167 (1.70) | 859 (1.63) | |
| Previous surgery | | | |
| No | 9810 (90.96) | 48312 (91.49) | 0.091 |
| Yes | 885 (9.04) | 4496 (8.51) | |
| Inotropic medication | | | |
| No | 9469 (96.90) | 51440 (97.41) | 0.004 |
| Yes | 303 (3.10) | 1368 (2.59) | |
| Peripheral vascular disease | | | |
| No | 8639 (88.52) | 47416 (89.79) | <0.001 |
| Yes | 1120 (11.48) | 5392 (10.21) | |
| Hypercholesterolaemia | | | |
| No | 3057 (31.33) | 16976 (32.15) | 0.113 |
| Yes | 6700 (68.67) | 35832 (67.85) | |
| Type of procedure | | | |
| CABG | 5898 (59.40) | 30187 (57.16) | < 0.001 |
| Valve | 1248 (12.57) | 9241 (17.50) | |
| Both | 826 (8.32) | 5920 (11.21) | |
| Others | 1957 (19.71) | 7460 (14.13) | |
| NYHA Classification | | | |
| I/II | 5407 (76.27) | 36842 (69.77) | < 0.001 |
| III | 1238 (17.46) | 12256 (23.21) | |
| IV | 444 (6.26) | 3710 (7.03) | |
| Ejection fraction | | | |
| Normal to mild | 2292 (81.02) | 42985 (81.40) | 0.562 |
| Moderate | 400 (14.14) | 7135 (13.51) | |
| Severe | 137 (4.84) | 2688 (5.09) | |
| Mortality | | | |
| No | 9344 (96.65) | 51,454 (97.44) | < 0.001 |
| Yes | 324 (3.35) | 1354 (2.56) | |

*P value was generated by Pearson chi-square test.

inotropic medication use, peripheral vascular disease, type of procedure, NYHA classification, and 30-day mortality were found to be associated with missing indicator variable, suggesting missing observations are related to some observed data. Hence the data are missing at random (MAR).

## Comparison Between CC and MI Models

The regression coefficients and standard error (SE) of the predictors of both CC and MI models are presented in Table 4. Percentage change of SE was calculated to assess the change in uncertainty of model estimates. Multiple imputation results in reduction in uncertainty ranges between 0.3% and 12.5% across categories of all predictors with an average reduction of -8.5%. The Bland and Altman plot demonstrates significant disagreement between the two risk scores generated using CC and MI models ($p<0.0001$) (Figure 1). If the measurements agree most of the data points in the plot would fall within 1.96 standard deviation (SD) of agreement line (zero line). However in our result, a large number of the data-points fell out of 1.96 SD of agreement line suggesting significant disagreement between the risk of MI and CC models. The spread is more apparent with greater average risk (higher x-axis values), giving a fanning-out appearance, demonstrating the trend of increasing disagreement for patients with a higher risk of 30-day mortality. Further, this graph shows that the CC model underestimated the mortality risk compared to the MI model as most data points fell below agreement line (zero line).

## Performance of CC and MI Models

Figure 2 compares the calibration performance of the two models. The graph demonstrates a better agreement between observed mortality and predictions for MI model compared to CC model. Regression of predicted risk on observed mortality also shows better agreement in the MI model (Intercept -0.00071, slope 1.0552) in comparison to the CC model (Intercept -0.00076 slope:1.1115). Intercept and slope of MI model approaches much closer to zero, and 1 respectively relative to those of CC model demonstrating better calibration with the former model. Discrimination of the models was assessed using 10-fold validation. The MI model (ROC=0.8232) showed better discrimination in comparison to the CC model (ROC=0.8156). The observed mortality was 2.69%. The predicted risk of mortality for the CC and MI models was 2.37% and 2.63% respectively ($P<0.001$). Further, the MI model performed better in the validation sample (MI: intercept -0.00014, slope 0.935; CC: intercept -0.00071 slope 1.0551).

# Discussion

Multiple imputation of missing values during model development increased the precision and performance of the model predicting 30-day mortality following cardiac surgery. The results of the current study corroborated evidence of bias associated with the conventional complete case analysis approach of handling missing data during model development.

It is important to consider the reasons, or mechanisms, for which data are missing, since approaches to handling missing data in the statistical analysis rely on assumptions of the mechanism (MCAR or MAR) and these assumptions can be tested. In MCAR there are no systematic differences between the missing values and the observed values [32], hence deletion of incomplete cases doesn't make much difference. Missing at random is a misnomer. This actually means the probability that an observation is missing is related to information for that subject that is present, i.e., observed characteristics [33]. This phenomenon allows simulating a plausible estimate of missing value from observed data.

The results of this study demonstrated a significant association of missing-ness in the individual subject with several predictors and the outcome in the ANZSCTS dataset, indicating that the missing data were MAR. This implies that bias will not arise if the multiple imputation approach is used, because the reason for missing-ness of a variable is associated with other variables in the model [7,34].

The aim of this paper was to highlight the importance of missing value imputation while developing clinical risk prediction models. Multiple imputation is a more advanced and robust method [6,35,36] of handling missing values, which reduces uncertainty, and improves performance of risk prediction models. The technique has become more popular because of its generality and recent software developments [33,37]. Reduction of uncertainty was assessed in this study by percentage reduction of standard error, and the performance was assessed in terms of calibration and discrimination. The MI model excelled over the CC model in both respects. Another way of assessing accuracy of a risk prediction model is to see how much closer the model predicts the observed risk. In the ANZSCTS registry, observed 30-day mortality was 2.69% compared to predicted mortality of 2.37% by the CC model and predicted mortality of 2.63% by the MI model. Imputation of missing value improves the prediction performance of the model and results in much closer prediction to actual mortality.

The agreement between the risk predicted by the CC analysis and MI analysis was assessed using the Bland and Altman plot which demonstrated significant disagreement between the two risk scores. The discrepancy between the two risk estimates showed an increasing trend with greater risk. The striking implication of this finding is, patients with higher risk of outcome (ie. mortality) are expected to incur more bias in prediction with the CC model. With increasing risk, the CC model tends to underestimate the risk compared to the MI model. An imprecise prediction in a high risk patient may have a greater potential impact on the surgical team and the patients' relatives.

Multiple imputation of missing value improves the prediction performance of the model. In the present study a model generated from imputed data showed better calibration, that is the model predicted closest towards the true risk of a patient's death following cardiac surgery. Compared to

**Table 4** Comparison of beta coefficient and its standard errors (SE) in CC and MI models

| Predictors | Complete case analysis Coefficient (SE) | Analysis after Imputation Coefficient (SE) | % change in SE |
|---|---|---|---|
| Age (Ref: < 60 years) | | | |
| Age 60 -70 years | 0.380 (0.094) | 0.412 (0.084) | -10.4 |
| Age 70 - 80 years | 0.907 (0.087) | 0.963 (0.080) | -10.4 |
| Age > 80 years | 1.286 (0.100) | 1.330 (0.091) | -9.3 |
| Gender (Ref: male) | | | |
| Female | 0.415 (0.061) | 0.377 (0.055) | -9.3 |
| Type of procedure (Ref: CABG) | | | |
| Valve (s) only | 0.521 (0.092) | 0.476 (0.086) | -7.0 |
| Valve + CABG | 0.768 (0.086) | 0.762 (0.079) | -8.0 |
| Others | 1.037 (0.080) | 1.058 (0.071) | -11.3 |
| Previous cardiac surgery (Ref: None) | | | |
| Yes | 0.540 (0.079) | 0.528 (0.072) | -9.2 |
| NYHA Class (Ref: GI/GII) | | | |
| G III | 0.400 (0.069) | 0.430 (0.064) | -7.1 |
| G IV | 0.640 (0.085) | 0.640 (0.081) | -4.0 |
| Inotropic medication use (Ref: No) | | | |
| Yes | 0.913 (0.099) | 0.869 (0.087) | -11.4 |
| Ejection fraction (Ref: Normal to mild) | | | |
| Moderate | 0.380 (0.076) | 0.379 (0.077) | 0.3 |
| Severe | 0.772 (0.090) | 0.747 (0.085) | -6.1 |
| Dialysis (Ref: No) | | | |
| Yes | 0.807 (0.144) | 0.807 (0.131) | -8.9 |
| Hypercholesterolaemia (Ref: No) | | | |
| Yes | 0.088 (0.065) | 0.017 (0.058) | -9.8 |
| Peripheral vascular disease (Ref: No) | | | |
| Yes | 0.517 (0.077) | 0.547 (0.070) | -10.1 |
| BMI >25 (Ref: No) | | | |
| Yes | -0.142 (0.061) | -0.151 (0.055) | -9.5 |
| Urgency of operation (Ref: Elective) | | | |
| Urgent | 0.612 (0.069) | 0.657 (0.062) | -9.1 |
| Emergency/salvage | 1.683 (0.090) | 1.854 (0.079) | -12.5 |
| Constant | | | |
| | -5.765 (0.115) | -5.754 (0.104) | -9.4 |
| Area under ROC | 0.8149 | 0.8210 | - |

the CC model, the MI model's overall predicted risk was also closer to the overall observed mortality. In fact, MI models are reported to have a greater ability to produce unbiased estimates of outcome and hence better ability to predict risk for individual patients [38,39]. Recent literature support that multiple imputation also improves the discrimination power [31], which is also evident from the current study.

Risk prediction modelling has become an integral part of cardiac surgery, a high risk field of medicine. The potential for its application in the diverse field of medical science ranging from administrative (i.e. schedule for surgery, estimation of length of hospital or intensive care unit stay, prediction of adverse postoperative outcomes, bench marking of institution and surgeons etc.) to fiscal issues (costing for procedure, resource and manpower allocation etc.) is becoming increasingly popular. Risk prediction scores are commonly used tools for patients' preoperative risk stratification which depends on prediction performance of the model. Both overestimation and underestimation of the risk scores may reduce the credibility of the risk stratification. Hence risk prediction modelling should endeavour to improve the prediction capacity, no matter how subtle the improvements are. For any risk prediction, missing values should be treated with an appropriate technique to maximise the prediction performance of the model.

This is the first study that has highlighted the importance of missing data imputation to improve the performance of risk prediction modelling in cardiac surgery. The prospective nature of the data that was collected from 28 centres across Australia is a very strong aspect of this study. The finding of

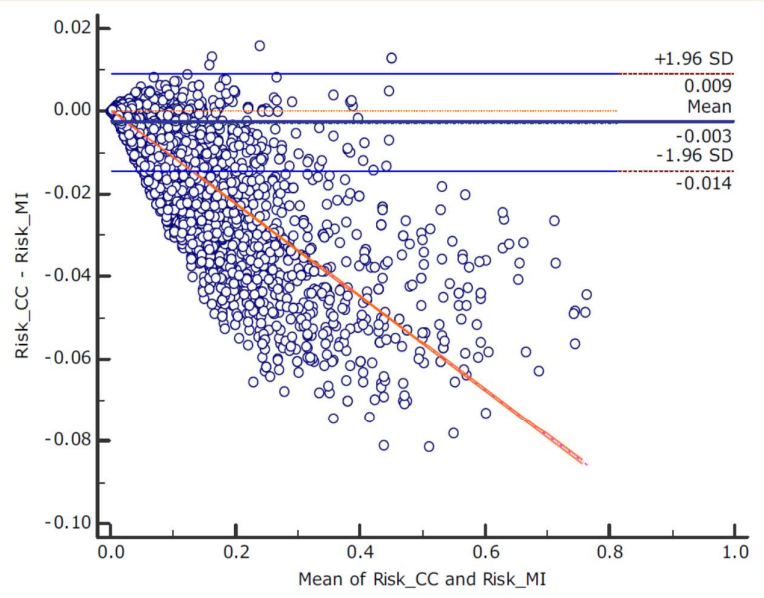



**Figure 1** Bland-Altman plot showing discrepancy between risk estimates of CC and MI model.

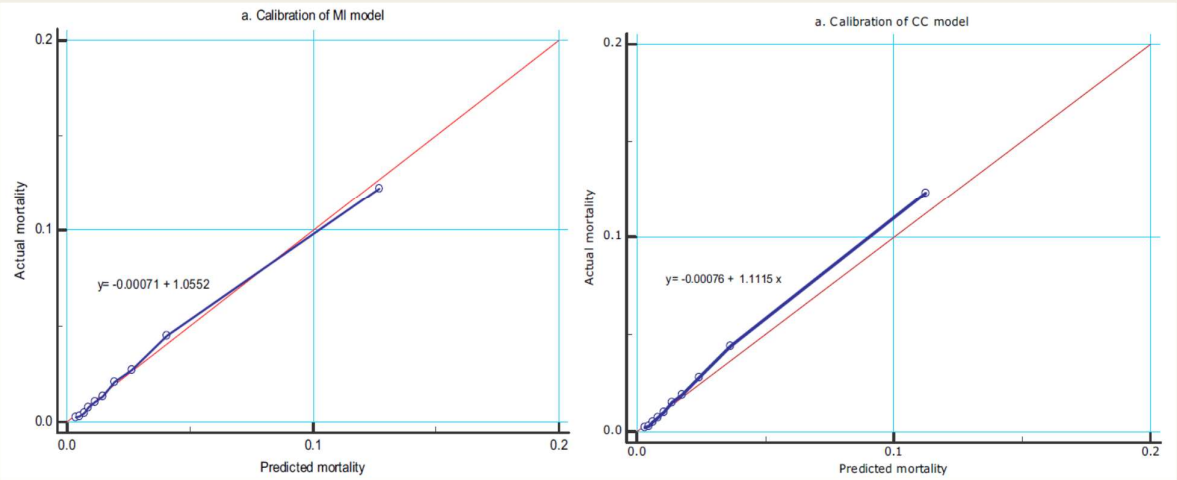



**Figure 2** Calibration plot (decile-decile) of CC and MI models in the validation dataset.

this study may be generalised to other cardiac surgery outcomes such as new renal failure, stroke, deep sternal wound infection, prolonged ventilation, atrial fibrillation and long-term mortality. Further, the methodology discussed in this study can also be applied to other large volumes of datasets and registries.

One limitation of this study is that the missing values were imputed for the selected variables in the risk model [25] used for estimation in this study. This consequently leads to the assumption that only the variables in this model were related to missing-ness. As a general rule of model development, use of all available plausible variables yields multiple imputations that have minimal bias and maximal certainty [34]. However such bias is unlikely in the current analyses as the same set of variables was used to estimate parameters with both CC and MI models. Another limitation was that imputation should have been performed prior to variable selection for the risk model, however variable selection was beyond the remit of this study.

In conclusion, multiple imputation of missing values improves the performance of the model for predicting risk of 30-day mortality following cardiac surgery. Missing data should be imputed before development of risk prediction models in the field of cardiac surgery.

## Authors' Declaration

The study received 'no external financial support' and has no conflict of interest to declare.

## Acknowledgements

The Australian and New Zealand Society of Cardiac and Thoracic Surgeons (ANZSCTS) National Cardiac Surgery Database Program is funded by the Department of Health, Victoria, the Health Administration Corporation (GMCT) and the Clinical Excellence Commission (CEC), NSW, and funding from individual Units. ANZSCTS Research activities are supported through a National Health and Medical Research Council Senior Research Fellowship and Program Grant awarded to C.M. Reid. M.N. Karim has an Australian Postgraduate Awards Scholarship from Monash University. The funding bodies had no role in the design and conduct of the study, the collection, management, analysis and interpretation of the data, or the preparation, review, or approval of the manuscript.

## References

[1] Kolh P. Importance of risk stratification models in cardiac surgery. Eur Heart J 2006;27:768–9.
[2] Cleveland Jr JC, Shroyer AL, Chen AY, Peterson E, Grover FL. Off-pump coronary artery bypass grafting decreases risk-adjusted mortality and morbidity. Ann Thorac Surg 2001;72(4):1282–8.
[3] Ferguson Jr TB, Coombs LP, Peterson ED. Preoperative beta-blocker use and mortality and morbidity following CABG surgery in North America. JAMA 2002;287(17):2221–7.
[4] Little RJ, Rubin DB. Statistical analysis with missing data, 2nd ed., New York: Wiley; 2002.
[5] Wood A, White IR, Thompson SG. Are missing outcome data adequately handled? A review of published randomised controlled trials. Clin Trials 2004;1:368–76.
[6] Greenland S, Finkle WD. A critical look at methods for handling missing covariates in epidemiologic regression analyses. Am J Epidemiol 1995; 142:1255e64.
[7] Soley-Bori M. Dealing with missing data: Key assumptions and methods for applied analysis. Boston University School of Public Health; May 6 2013, Technical Report No. 4.
[8] Moons KG, Donders RA, Stijnen T, Harrell Jr FE. Using the outcome for imputation of missing predictor values was preferred. J Clin Epidemiol 2006;59:101–1092.
[9] Donders AR, van der Heijden GJ, Stijnen T, Moons KG. Review: a gentle introduction to imputation of missing values. J Clin Epidemiol 2006;59: 1087–91.
[10] Vach W, Blettner M. Biased estimation of the odds ratio in case-control studies due to the use of ad hoc methods of correcting for missing values for confounding variables. Am J Epidemiol 1991;134:895–907.
[11] Rubin DB. Multiple imputation for nonresponse in surveys. New York: Wiley; 1987.
[12] Altman DG, Bland JM. Missing data. BMJ 2007;334:424.
[13] Schafer JL, Graham JW. Missing data: our view of the state of the art. Psychol Methods 2002;7:147–77.
[14] Parsonnet V, Dean D, Bernstein AD. A method of uniform stratification of risk for evaluating the results of surgery in acquired adult heart disease. Circulation 1989;79:I3–I12.
[15] Hattler BG, Madia C, Johnson C, Armitage JM, Hardesty RL, Kormos RL, et al. Risk stratification using the Society of Thoracic Surgeons programme. Ann Thorac Surg 1994;52:1348–52.
[16] Huijskes RVHP, Rosseel PMJ, Tijssen JGP. Outcome prediction in coronary artery bypass grafting and valve surgery in the Netherlands: development of the Amphiascore and its comparison with the Euroscore. Eur J Cardiothorac Surg 2003;24:741–9.
[17] Pons JM, Granados A, Espinas JA, Borras JM, Martin I, Moreno V. Assessing open heart surgery mortality in Catalonia (Spain) through a predictive risk model. Eur J Cardiothorac Surg 1997;11:415–23.
[18] Ivanov J, Tu JV, Naylor CD. Ready-made, recalibrated, or remodeled? Issues in the use of risk indexes for assessing mortality after coronary artery bypass graft surgery. Circulation 1999;99:2098–104.
[19] Bridgewater B, Neve H, Moat N, Hooper T, Jones M. Predicting operative risk for coronary artery surgery in the United Kingdom: a comparison of various risk prediction algorithms. Heart 1998;79:350–5.
[20] Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R. European system for cardiac operative risk evaluation (EuroSCORE). Eur J Cardiothorac Surg 1999;16:9–13.
[21] Nashef SAM, Roques F, Sharples LD, Nilsson J, Smith C, Goldstone AR, et al. EuroSCORE II. Eur J Cardiothorac Surg 2012;(41):734–45.
[22] Shahian DM, O'Brien SM, Filardo G, Ferraris VA, Haan CK, Rich JB, et al. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: Part 1 – coronary artery bypass grafting surgery. Ann Thorac Surg 2009;89(1 Suppl 1):S2–S22.
[23] Shahian DM, Edwards FH. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: Introduction. Ann of Thorac Surg 2009;88(1 Suppl 1):S1.
[24] Reid C, Billah B, Dinh D, Smith J, Skillington P, Yii M, et al. An Australian risk prediction model for 30-day mortality after isolated coronary artery bypass: The AusSCORE. J Thorac Cardiovasc Surg 2009;138:904–10.
[25] Billah B, Reid CM, Shardey GC, Smith JA. A preoperative risk prediction model for 30-day mortality following cardiac surgery in an Australian cohort. Eur J Cardiothorac Surg 2010;37:1086–92.
[26] Ariyaratne TV, Billah B, Yap C-H, Dinh D, Smith JA, Shardey GC, et al. An Australian risk prediction model for determining early mortality following aortic valve replacement. Eur J Cardiothorac Surg 2011;39 (6):815–21.
[27] Billah B, Haq MM, Smith JA, Sufu F, Tran L, Shardy GC, et al. AusScore II in Predicting 30-day mortality following isolated coronary artery bypass grafting in Australia and New Zealand. J Thorac Cardiovasc Surg 2014 Nov;148(5):1850–5.
[28] Schafer JL, Olsen MK. Multiple imputation for multivariate missing data problems: A data analyst's perspective. Multivariate Behavioral Research 1998;33:545–71.
[29] Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 2010;21:128–38.
[30] Miller ME, Langefeld CD, Tierney WM, Hui SL, McDonald CJ. Validation of probabilistic predictions. Med Decis Making 1993;13(1):49–58.
[31] Royston P, Altman DG. Visualizing and assessing discrimination in the logistic regression model. Stat Med 2010;29:2508–20.
[32] Rubin DB. Inferences and missing data. Biometrika 1976;63:581–90.
[33] Van Buuren S, Boshuizen SC, Knook DL. Multiple imputation of missing blood pressure covariates in survival analysis. Stat Med 1999;18:681–94.
[34] Unnebrink K, Windeler J. Intention-to-treat: methods for dealing with missing values in clinical trials of progressively deteriorating diseases. Stat Med 2001;20:3931–46.
[35] Rubin DB. Multiple Imputation after 18 years. J Am Stat Assoc 1996; 91:473–89.
[36] Schafer JL. Analysis of Incomplete Multivariate Data. London: Chapman & Hall; 1997.
[37] Royston P. Multiple imputation of missing values: update. Stata Journal 2005;5(4):527–36.
[38] Miettinen OS. Theoretical epidemiology. Principles of occurrence research in medicine. New York: Wiley; 1985.
[39] Vergouwe Y, Royston P, Moons KG, Altman DG. Development and validation of a prediction model with missing predictor data: a practical approach. J Clin Epidemiol 2010;63(2):205–14.

## 5.3 Summary

Despite recommendations for imputing missing values prior to developing a risk prediction model (55), the currently available risk prediction models do not seem to adequately address the issue of missing data. A common approach to dealing with missing data in model development is to drop cases with incomplete data from the analysis. This complete case analysis approach of handling missing data during model development is generally not satisfactory, since it can introduce bias in the estimated model parameters. The study described in this chapter found that imputation of missing values improves the performance of a risk prediction model with regards to predictive accuracy. The study resulted in a recommendation that multiple imputation of missing values be carried out prior to the development of a risk prediction model.

# Chapter 6: Approaches to variable selection and parsimony of risk prediction model

## 6.1 Introduction

Introduction

This chapter addresses the balance between the predictive performance of a model and parsimony, a key issue identified in the systematic review discussed in Chapter 3. A parsimonious model is one which is computationally simpler and has a relatively smaller number of predictor variables, and is therefore easier for a clinician to use in their day-to-day practice (58). However, omitting important prognostic factors has the potential to result in inaccurate predictions (59). A risk prediction model therefore needs to maintain a balance between the number of variables in the model and the predictive accuracy. This trade-off between parsimony and predictive performance is a major challenge in risk prediction modelling (9).

Several variable selection techniques are often used when developing cardiac surgery post-operative mortality risk prediction models. There is no agreement about a method that provides the optimum balance between a models parsimony and performance. The aim of the study in this chapter was to compare the parsimony and predictive performance of risk prediction models generated using different variable selection methods. This chapter includes the article entitles “Variable selection methods for multiple regressions influence the parsimony of risk prediction models for cardiac surgery” which has been published in the Journal of Thoracic and Cardiovascular Surgery.

## 6.2 Manuscript

**Declaration for Thesis Chapter 6**

**Manuscript:** Karim MN, Reid CM, Tran L, Cochrane A, Billah B. Variable selection methods for multiple regressions influence the parsimony of risk prediction models for cardiac surgery. The Journal of Thoracic and Cardiovascular Surgery. 2017. (Article in press)

**Declaration by candidate**

In the case of Chapter 6, the nature and extent of my contribution to the work was the following:

| Nature of contribution | Extent of contribution (%) |
|---|---|
| Study concept and design, literature search, data analysis and interpretation, manuscript development and preparation | 75% |

The following co-authors contributed to the work. Where co-authors are students at Monash University, the extent of their contribution in percentage terms is stated:

| Name | Nature of contribution | Extent of contribution (%) for student co-authors only |
|---|---|---|
| Billah, B | Study concept, data analysis and manuscript editing | |
| Reid, C | Study concept, study design and manuscript editing | |
| Cochrane, A | Study concept, Interpretation of result and manuscript editing | |
| Tran, L | Interpretation of result and manuscript editing | |

The undersigned hereby certify that the above declaration correctly reflects the nature and extend o the candidate's and co-author' contributions to this work

| | | |
|---|---|---|
| **Candidate's Signature** | [redacted] | **Date 10/03/2017** |
| **Main Supervisor's Signature** | [redacted] | **Date 10/03/2017** |

# Variable selection methods for multiple regressions influence the parsimony of risk prediction models for cardiac surgery

Md Nazmul Karim, MBBS, M Clin Epi,[a] Christopher M. Reid, PhD,[a,b] Lavinia Tran, PhD,[a] Andrew Cochrane, MD, FRCS,[c] and Baki Billah, PhD[a]

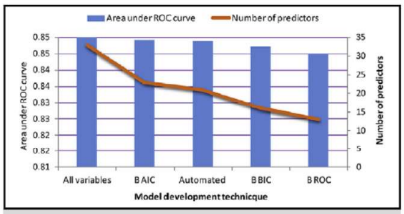
## ABSTRACT

**Objective:** To compare the impact of different variable selection methods in multiple regression to develop a parsimonious model for predicting postoperative outcomes of patients undergoing cardiac surgery.

**Methods:** Data from 84,135 patients in the Australian and New Zealand Society of Cardiac and Thoracic Surgeons registry between 2001 and 2014 were analyzed. Primary outcome was 30-day-mortality. Mixed-effect logistic regressions were used to build the model. Missing values were imputed by the use of multiple imputations. The following 5 variable selection methods were compared: bootstrap receiver-operative characteristic (ROC), bootstrap Akaike information criteria, bootstrap Bayesian information criteria, and stepwise forward and stepwise backward methods. The final model's prediction performance was evaluated by the use of Frank Harrell's calibration curve and using a multifold cross-validation approach.

**Results:** Stepwise forward and backward methods selected same set of 21 variables into the model with the area under the ROC (AUC) of 0.8490. The bootstrap ROC method selected 13 variables with AUC of 0.8450. Bootstrap Bayesian information criteria and Akaike information criteria respectively selected 16 (AUC: 0.8470) and 23 (AUC: 0.8491) variables. Bootstrap ROC model was selected as the final model which showed very good discrimination and calibration power.

**Conclusions:** Clinical suitability in terms of parsimony and prediction performance can be achieved substantially by using the bootstrap ROC method for the development of risk prediction models. (J Thorac Cardiovasc Surg 2017;153:1128-35)




The bootstrap receiver operating characteristic method provides parsimony of risk prediction model.

### Central Message

The balance between parsimony and performance can be improved using the bootstrap method with receiver operating characteristics for developing risk prediction models.

### Perspective

The trade-off between parsimony and performance is a major challenge in risk prediction modeling. Different approaches to variable selection may be an avenue for improving the parsimony of a cardiac surgical risk prediction model. This work compared the clinical suitability of models generated by the use of different variable selection methods with regard to parsimony and performance to predict 30-day mortality.

See Editorial Commentary page 1136.

In the past decades, the field of cardiac surgery has made significant progress in the development of risk prediction models to enable outcome prediction and clinical quality monitoring. National cardiac surgical registries have been established in many countries, and many have developed risk prediction models suitable for local populations.[1-3] The aim of these models was to provide an estimate of postoperative mortality risk based on preoperative risk factors.

In risk prediction modeling of cardiac surgery, outcomes data are adjusted for preoperative risk factors. When so many variables are included in the model, however, the accuracy along with parsimony of the model may be compromised.[4] The principle of parsimony states that

From the [a]School of Public Health and Preventive Medicine, Department of Epidemiology and Preventive Medicine, Monash University, Melbourne; [b]School of Public Health, Department of Health Policy and Management, Curtin University, Western Australia; and [c]Department of Cardiothoracic Surgery, Monash Medical Centre, Clayton, Victoria, Australia.

Received for publication Dec 1, 2015; revisions received Oct 24, 2016; accepted for publication Nov 14, 2016; available ahead of print Dec 19, 2016.

Address for reprints: Baki Billah, PhD, Department of Epidemiology and Preventive Medicine, Faculty of Medicine Nursing and Health Science, Monash University, The Alfred Centre, (level 6), 99 Commercial Rd, Melbourne, VIC 3004, Australia (E-mail: [redacted]).

0022-5223/$36.00



http://dx.doi.org/10.1016/j.jtcvs.2016.11.028

Scanning this QR code will take you to supplemental tables for this article.

**Abbreviations and Acronyms**

AIC = Akaike information criteria
ANZSCTS = Australian and New Zealand Society of Cardiac and Thoracic Surgeons Registry
AUC = Area under ROC curve
BIC = Bayesian information criteria
CI = confidence interval
ROC = receiver operating characteristic

simpler explanations are preferred over more complex explanations.[5] A parsimonious model is computationally simpler for the clinician to implement in day-to-day practice.[5] A smaller number of variables in a risk prediction model makes the model simpler to use. Furthermore, when many variables are included in the model, the burden of clinical staff is increased considerably in regard to data collection, which can compromise data quality as well. The prevalence of cases with missing data also may increase with increasing number of variables in the model. If a simple model can explain a phenomenon with a similar level of accuracy compared with a complex model, the former model should be preferred on the ground of parsimony, unless the complex model outperforms the former.[6,7]

Several methodologic approaches are currently in practice for identifying the predictor variables for risk prediction models. These methods excel in different tasks and have their inherent limitations. Automated variable selection methods are used widely for model development.[8] Automated variable selection, forward selection, and backward elimination are very easy to use and available in most statistical software packages. Although popularly used, a number of concerns have been identified with the application of these methods in the field of risk prediction modeling.[9]

Bootstrap resampling is another approach used for variable selection in risk prediction modeling.[10-12] Usually bootstrap resampling is used in conjunction with criteria like Akaike information criteria (AIC),[13] Bayesian information criteria (BIC),[14] and receiver operating characteristics (ROCs)[15] to develop models; however, the question of whether developing models using bootstrapping method in combination with these criteria improves the parsimony of a risk prediction model remains unclear. Therefore, comparison of different methodologic approaches for variable selection may be an avenue for improving the parsimony of risk prediction models.

The aim of this paper was to compare the clinical suitability of risk prediction models generated by the use of different variable selection methods with regard to their parsimony and performance to predict postoperative outcomes of cardiac surgery patients.

## METHODS

### Australian and New Zealand Society of Cardiac and Thoracic Surgeons Registry (ANZSCTS) Database

The ANZSCTS collects information on adult patients undergoing cardiac surgery in 28 hospitals across Australia. The database collects 287 preoperative, intraoperative, and postoperative variables via internationally standardized data definitions (Table E1).[3] The data collection and its audit methods have been discussed elsewhere.[3,16] Current analysis included information of 84,135 patients who underwent cardiac surgery between 2001 and 2014. Of them, 63,523 patient records between 2001 and 2012 were used for model development and 20,625 patient records between 2013 and 2014 were used for validation. The primary outcome variable for this study was 30-day mortality.[17] The variable "New York Heart Association classification" (5.28%) had the highest missing data, followed by "ejection fraction" (3.07%), "procedure type" (1.24%), and "number of diseased vessels" (1.11%). The remaining predictors had <1% missing observations (Table E2).

### Plausible Predictor Identification and Model Development Methods

Of 287 variables (data field) in the ANZSCTS registry, 101 were preoperative variables. These preoperative variables included administrative data and several stems of variables. On the basis of extensive literature review on cardiac surgery risk prediction models and clinical judgment, a total of 52 variables were identified primarily as preoperative risk factors. With the use of bivariate mixed-effect logistic regression, 33 variables finally were identified as plausible risk factors for 30-day mortality.

Multiple imputation of missing value was performed with the imputation by chained equations method along with multivariable mixed-effect logistic regression. Data were "filled in" with imputed values generated by the use of a specified regression model. The process was repeated 3 times to generate 3 completed datasets. Studies have shown that repeating imputation 3 times is sufficient for data with less missing data (<20%).[18] The mixed-effect logistic regression was fitted separately on each of the imputed dataset. The results were then pooled into an aggregated estimate following Rubin's rule[19] through the MIM estimation command option in Stata 14.[20] Multivariable mixed-effect logistic regression was used for the risk prediction. The multilevel modeling accounts for potential between-hospital variations.[21] The first order interaction effect between clinically relevant risk factors was investigated.

Nonlinearity of continuous predictors (age, body mass index, and estimated glomerular filtration rate) was addressed by fitting fractional polynomial in the mixed-effect logistic regression model.[22,23] Sensitivity analysis was performed for nonlinear term in the model. Little improvement in discrimination and calibration was seen with inclusion of nonlinear terms in the model; hence, linear terms of the continuous variables were kept in the final model to keep the model simple and user friendly. The following 5 variable selection methods were compared: (1) bootstrap along with ROC, (2) bootstrap along with AIC, (3) bootstrap along with BIC, (4) stepwise forward selection method, and (5) stepwise backward elimination method. A model with all 33 variables also was developed.

### Bootstrap Model Selection

A bootstrap sample of the same size of the original sample was drawn from each of the 3 imputed datasets. The 33 plausible risk factors were entered into the mixed-effect logistic regression and were applied to the bootstrap sample to test the significance of the variables. A variable with

*P* value of less than or equal to .05 was considered as significant. The process was repeated 1000 times, and the percentage of times that each variable appeared as significant in 1000 bootstraps (bootstrap coverage) in the imputed datasets was recorded. Bootstrap coverage of each predictor in 3 imputed dataset was averaged to generate an overall coverage of individual predictors. The predictors were then ranked depending on the overall bootstraps coverage.[24]

Fourteen plausible models were developed from variables that were significant in at least 50% of the bootstrap samples.[25] Model 1 comprised 10 predictors that appeared as significant in 100% of bootstrap samples. Then, 13 subsequent models were generated through adding one variable at a time of decreasing rank according to the bootstrap coverage. The area under ROC curve (AUC), AIC, and BIC values were calculated for all of these 14 models. Both AIC and BIC methods select the model with their lowest values. Bootstrap ROC method selects the model with the greatest AUC value.

### Automated Model Selection

Variables in the forward and backward logistic regression were selected for the imputed data as suggested by Wood and colleagues.[24] The selected variables were then entered into the mixed-effect logistic regression and the model was estimated with multiple imputation, MIM estimation command option, in Stata 14.[20] The AUC, AIC, and BIC values were calculated for final automated model.

### Discrimination and Calibration of the Models

Selected models' prediction performance was evaluated by the use of discrimination and calibration powers. Model discrimination was evaluated with AUC. The calibration was evaluated with a decile-decile plot of the observed and predicted 30-day mortality. To calculate the calibration intercept and slope parameters, a linear regression model was fitted with the deciles of observed outcome as the dependent variable and the deciles of predicted outcome as the independent variable. Calibration of final model was also assessed using the method suggested by Frank Harrell with Regression Modeling Strategies (RMS) package in R statistical software.[7]

### External Validation of the Final Model

External validation of the final model was done on the validation dataset comprising patients treated during 2013 to 2014.

### Ethical Approval

The institutional review board of each participating hospital approved the use of these databases for research (Alfred HREC:262/09). The ANZSCTS database registry approved collection of patient data via opt-out consent approach (MUHREC:CF08/0322-2008000065). The study received Ethical approval from Monash University, Standing Committee on Ethics in Research Involving Humans (SCERH) (MUHREC:CF14/1117–2014000476).

## RESULTS

A total of 33 preoperative variables were identified as potential predictors (Tables E3 and E4). There was no first-order interaction effect between these predictors. The percentage of times each of these candidate variables appeared as significant in the multiple mixed-effect logistic regression models in 1000 bootstraps were summarized in Table 1.

Figure 1 shows the AUC of 14 bootstrap models plotted against number of variables based on bootstrap coverage increment. The model with 10 predictors those appeared as significant in 100% of the bootstrap samples had an AUC of 0.8392. With each addition of predictors, the AUC increased steeply until the model with 13 predictors (AUC 0.8450) appeared as significant in at least 99.6% of the samples and apparently reached a plateau (Figure 1). The addition of a further variable into the model didn't result in a steep increase in AUC (0.8454), which also corresponded to the sharp decrease of bootstrap coverage (99.6% with 13 variables vs 90.2% with 14 variables). Hence, the model with 13 predictors was selected as the final bootstrap ROC model. AIC and BIC values for all of the 14 competing bootstrap models are presented in Figure 2. The model with 23 variables had the lowest AIC (12,326.5) and the model with 16 variables had the lowest BIC (12,584.8). Hence, the model with 23 variables and the model with 16 variables were selected as bootstrap

**TABLE 1. Summary of appearance of variables as independent predictors of mortality following cardiac surgery in bootstrap resampling**

| No. | Predictors | %* |
|---|---|---|
| 1 | Age | 100.0 |
| 2 | Sex | 100.0 |
| 3 | Previous cardiac surgery | 100.0 |
| 4 | Peripheral vascular disease | 100.0 |
| 5 | Stroke | 100.0 |
| 6 | Glomerular filtration rate | 100.0 |
| 7 | Infective endocarditis | 100.0 |
| 8 | Urgency of procedure | 100.0 |
| 9 | Type of procedure | 100.0 |
| 10 | Ejection fraction | 100.0 |
| 11 | Inotropic medication | 99.93 |
| 12 | Angina | 99.83 |
| 13 | NYHA class | 99.57 |
| 14 | Dialysis | 90.17 |
| 15 | Congestive heart failure | 89.80 |
| 16 | Shock | 87.13 |
| 17 | Myocardial infarction | 86.27 |
| 18 | Resuscitation | 86.80 |
| 19 | Arrhythmia | 81.00 |
| 20 | Left main disease <50% | 70.77 |
| 21 | Hypertension | 68.00 |
| 22 | Respiratory disease | 56.00 |
| 23 | Hypercholesterolemia | 55.63 |
| 24 | Cerebrovascular disease | 30.00 |
| 25 | Steroid use | 31.17 |
| 26 | Pacemaker in situ | 27.27 |
| 27 | Body mass index | 21.20 |
| 28 | Diabetes | 19.97 |
| 29 | Number of disease vessel | 14.10 |
| 30 | Immunosuppressant use | 11.50 |
| 31 | Family history of heart disease | 12.87 |
| 32 | Anticoagulant use | 7.53 |
| 33 | Intra venous nitrate use | 8.13 |

*NYHA*, New York Heart Association. *Percentage of times of appearance in bootstrap resampling.

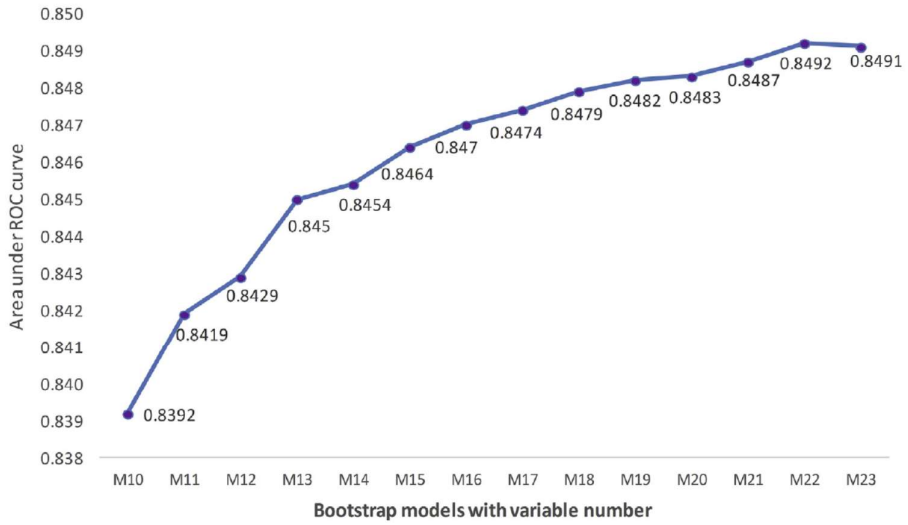



**FIGURE 1.** AUC gradient with increasing number of variables based on bootstrap coverage increment. *AUC*, Area under the ROC curve.

AIC and BIC model, respectively. The AUC values for bootstrap AIC and BIC models were, respectively, 0.8470 and 0.8491.

The stepwise forward selection and backward elimination methods selected same set of 21 variables into the model with AUC of 0.8490. The model with all 33 predictors had an AUC of 0.8504.

The observed 30-day mortality in development sample was 2.693%.The predicted mortality for all variable model was 2.524% (95% confidence interval [CI], 2.481-2.567) and for the automated model was 2.524 (95% CI, 2.481-2.567). The predicted risk for the bootstrap AIC, BIC, and ROC models were 2.529% (95% CI, 2.487-2.572), 2.584% (95% CI, 2.540-2.627), and 2.574% (95% CI, 2.531-2.616), respectively.

Figure 3 shows the calibration of the competing 5 models and the final model in validation sample. Bootstrap BIC, AIC, and ROC models showed calibration slopes of 1.0128, 1.0219, and 1.0191, respectively. The automated and all variable models had calibration slopes of 1.0271 and 1.0133, respectively. All of these models have intercepts close to zero.

Based on the number of variables and discrimination and calibration, the bootstrap ROC model with 13 variables was chosen as the final model (Table 2). The overfitting-corrected loess nonparametric calibration curve demonstrated excellent calibration for the bootstrap ROC

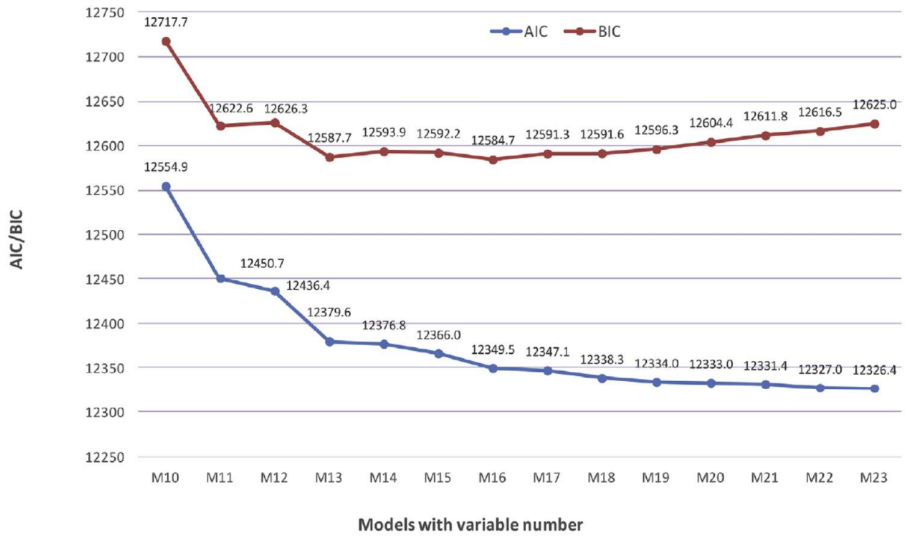



**FIGURE 2.** AIC and BIC gradient with increasing number of variables based on bootstrap coverage increment. *AIC*, Akaike information criteria; *BIC*, Bayesian information criteria.

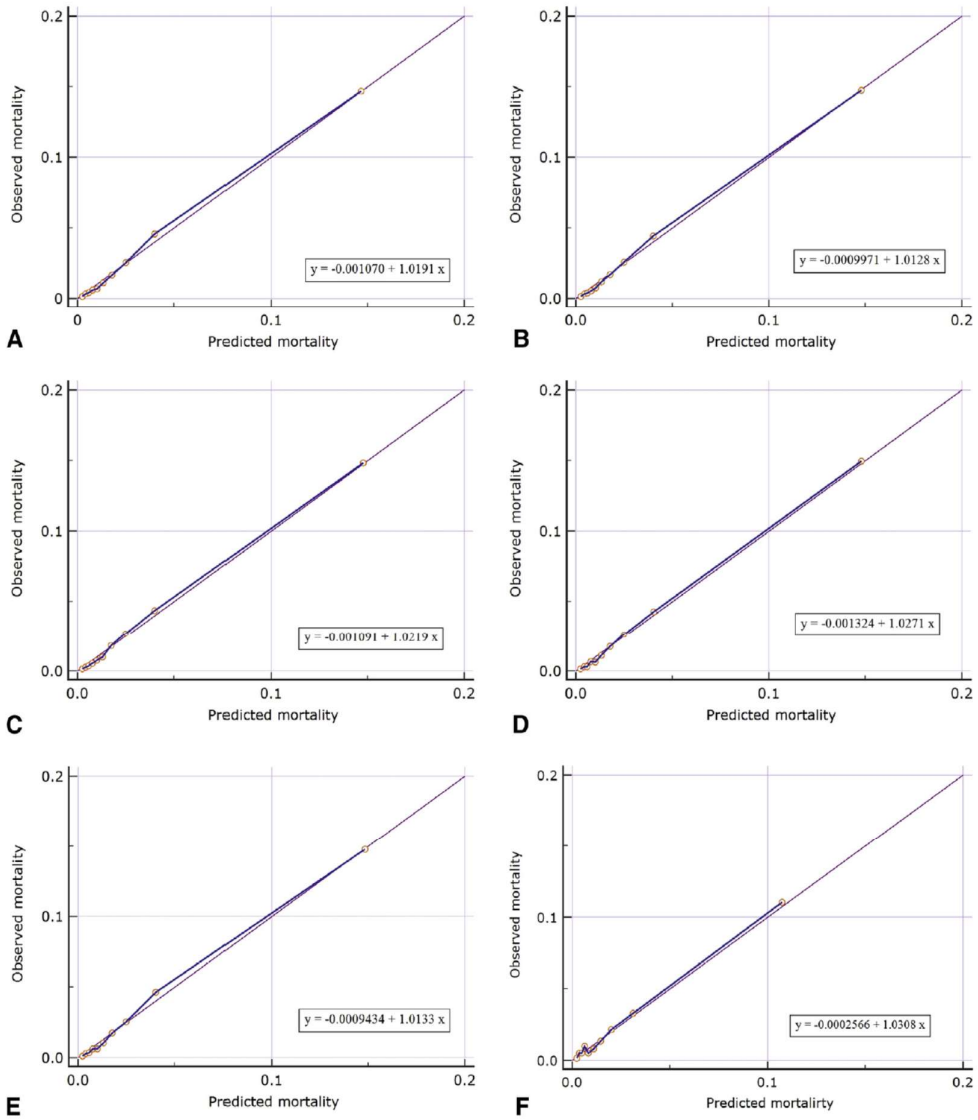



**FIGURE 3.** Decile-decile plots of bootstrap and automated models and validation model. A, Bootstrap ROC model; B, bootstrap BIC model; C, bootstrap AIC model; D, automated model; E, all-variable model; and F, bootstrap ROC model in validation data.

model, especially for low-risk patients (Figure 4). The model also showed very good discrimination in a multifold (10) validation (AUC, 0.8384; 95% CI, 0.8356-0.8413) in creation data as well as in external validation (AUC, 0.8156; 95% CI, 0.7936-0.8375) (Table E5).

## DISCUSSION

This study is the first demonstration in the surgical literature that compares different statistical approaches for developing a parsimonious risk prediction model. A risk prediction model needs to maintain a balance between the number of variables in the model and predictive accuracy.[26] It is important to have a model without variables that may add little or no useful information. Omitting important prognostic factors has the potential to result in a biased estimation of the regression coefficients and inaccurate prediction.[27] Hence, the trade-off between parsimony and prediction performance is a major challenge in risk prediction modeling. The inclusion of a predictor in the model should be judged against the amount of prediction power it adds to the model and one should refrain from including the new predictor if the gain seems negligible.[5] When a complicated model with many predictors in practice is applied, many patients will likely need to be excluded on the basis of missing

**TABLE 2. Comparison of clinical suitability (parsimony and prediction performance) of bootstrap, automated and all predictor models**

| Plausible variables | All variables | Backward elimination | Forward selection | Bootstrap AIC | Bootstrap BIC | Bootstrap ROC |
|---|---|---|---|---|---|---|
| Age | √ | √ | √ | √ | √ | √ |
| Sex | √ | √ | √ | √ | √ | √ |
| Previous cardiac surgery | √ | √ | √ | √ | √ | √ |
| Peripheral vascular disease | √ | √ | √ | √ | √ | √ |
| Stroke | √ | √ | √ | √ | √ | √ |
| Glomerular filtration rate | √ | √ | √ | √ | √ | √ |
| Infective endocarditis | √ | √ | √ | √ | √ | √ |
| Urgency of procedure | √ | √ | √ | √ | √ | √ |
| Ejection fraction | √ | √ | √ | √ | √ | √ |
| Type of procedure | √ | √ | √ | √ | √ | √ |
| Inotropic medication use | √ | √ | √ | √ | √ | √ |
| Angina | √ | √ | √ | √ | √ | √ |
| NYHA class | √ | | | √ | √ | √ |
| Dialysis | √ | | | √ | √ | |
| Congestive heart failure | √ | √ | √ | √ | √ | |
| Shock | √ | √ | √ | √ | √ | |
| Myocardial infarction | √ | √ | √ | √ | | |
| Resuscitation | √ | | | √ | | |
| Arrhythmia | √ | √ | √ | √ | | |
| Left main disease <50% | √ | √ | √ | √ | | |
| Hypertension | √ | | | √ | | |
| Respiratory disease | √ | √ | √ | √ | | |
| Hypercholesterolemia | √ | | | √ | | |
| Steroid use | √ | | | | | |
| Cerebrovascular disease | √ | | | | | |
| Diabetes | √ | | | | | |
| Pacemaker in situ | √ | | | | | |
| Body mass index | √ | √ | √ | | | |
| Number of disease vessel | √ | √ | √ | | | |
| Family history of heart disease | √ | | | | | |
| Immunosuppressant use | √ | | | | | |
| Intravenous nitrate use | √ | | | | | |
| Anticoagulant within 7 d | √ | | | | | |
| Number of variables | 33 | 21 | 21 | 23 | 16 | 13 |
| AUC | 0.8504 | 0.8490 | 0.8490 | 0.8491 | 0.8470 | 0.8450 |

*AIC*, Akaike information criteria; *BIC*, Bayesian information criteria; *ROC*, receiver operating characteristic; *NYHA*, New York Heart Association; *AUC*, area under the ROC curve.

predictors. Furthermore, if many variables are included in a model, the likelihood of co-linearity increases which may result in biased or unstable estimation and model uncertainty.[28]

In this study, models generated by the use of automated variable selection approaches did not show better parsimony in comparison with bootstrap ROC model. The variable selection for risk prediction modeling should consider the clinical plausibility rather than solely relying on statistical variable selection methods.[4] Automated model selection processes act as a black box, which can result in blind selection and hence using this method, a researcher possesses less control over what predictors are included and what is being eliminated.[29] The use of automated variable selection methods sometimes may even produce nonreproducible regression models.[7] Furthermore, automated selection has also the potential to create bias in the estimated regression coefficients.[30] Particularly when the prevalence of events is low, the variable selection may be unstable, the estimated regression coefficients are too extreme, and the performance of the selected model is overestimated.[7] Automated model-building methods also can mask multicollinearity. In such models, the number of noise (unimportant) variables increases with an increase in the number of plausible predicators, which translates into a decrease in the probability of correctly identifying true predictor variables.[27] A study on simulated data demonstrated that automated model selection may result in unstable models and may select noise variables.[9] In the present study, the automated selection method chooses the variables body mass index and number of diseased vessel which had very poor bootstrap coverage (21.2%) and (14.1%), respectively (Table 2). There are potential that variable selection method

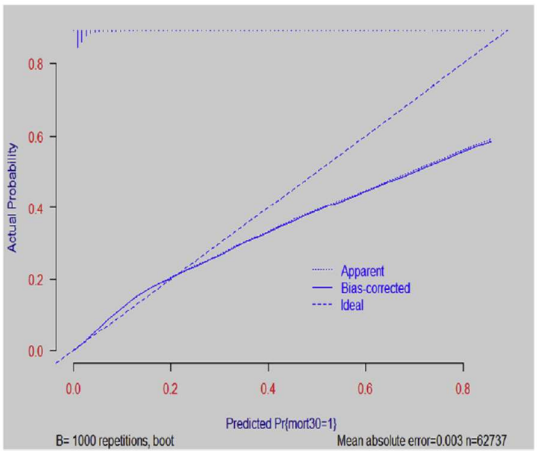



**FIGURE 4.** Overfitting-corrected loess nonparametric calibration curve demonstration calibration of the model.

may eliminate variables those are historically important. For instance, bootstrap ROC model didn't include variables like "cardiogenic shock," "myocardial infarction," and "angina." The model building process actually identifies the best predictors, among the set of similar predictors, which explains most variation in the outcome. In this process, it may eliminate a variable that is historically important; however, the model keeps a surrogate that explains the variation from the excluded variable. It could be that inclusion either or all of "New York Heart Association class," "urgency of operation," and "ejection fraction" is enough to capture the variation of those eliminated variables. For example, the model with 16 and 18 variables included "'cardiogenic shock'" and "myocardial infarction" but didn't perform better than the models without these variables.

To minimize the limitation of automated model selection methods, backward elimination in combination with bootstrap resampling was proposed by Austin and Tu[31]; however, this approach did not improve the ability of variable selection to identify the true predictors of an outcome.[32] Reasons why this combination method did not work could be that the automated model selection was repeated in all 1000 bootstrap samples, and hence the limitation of the automated variable selection persisted. To overcome this limitation, in this study, within each bootstrap sample the multivariable logistic regression was run independently of automated method. A similar approach was used to develop several models for predicting 30-day mortality in the ANZSCTS database.[3,16,33] In this approach, the researcher has the freedom to decide on inclusion and exclusion of predictors for the final model based on theoretical and clinical plausibility and parsimony.

Small degrees of random variation in one dataset can have a substantial influence on the variables that are identified as independent predictors. Thus, it is likely that no one regression model estimated on one dataset can conclusively identify the independent predictors.[11] Bootstrap resampling has been considered as a solution for the sampling variation. The bootstrap method can provide insight in the distribution of a summary measure from a sample. The bootstrap method draws samples from the original sample to introduce a random element.[12,34]

This study showed that bootstrap ROC method generates a parsimonious risk prediction. The model performed very well in the validation dataset. With only 13 predictors, this model showed a competitive discrimination power (AUC: 0.8156) compared with other commonly used risk prediction models such as The Society of Thoracic Surgeons (AUC: 0.8120, 31 predictors)[1] and European System for Cardiac Operative Risk Evaluation II (AUC: 0.8095, 18 predictors).[35]

The bootstrap ROC method also can be used for developing parsimonious risk prediction model for other postoperative outcomes such as new renal failure, stroke, postoperative atrial fibrillation, pneumonia, and long-term mortality in cardiac surgery. This method also could be a useful tool for risk prediction modeling in other disciplines of research.

One limitation of this research was that all the competing models were developed for all procedures and not for a specific procedure type. It is generally hypothesized that procedure specific models performs better than the global model. Furthermore, many of the popularly used models in the field are for general cardiac surgery. The objective of this study, however, was to compare the model development methods. All the models were developed with the same process; hence, the findings in this study were not affected by case-mix. Another limitation of this study was that a number of other methods are available for screening and identification of potential candidate variables including random forest classification, classification and regressing tree (CART) etc. In this study, however, we only compared the methods currently popularly being used in the field of risk prediction modeling.

In conclusion, clinical suitability in terms of parsimony and prediction performance can be achieved by the use of bootstrap resampling in conjunction with ROC for the development of risk prediction models. We recommend this method for future risk prediction model development.

## Conflict of Interest Statement

Authors have nothing to disclose with regard to commercial support.

The Australian and New Zealand Society of Cardiac and Thoracic Surgeons (ANZSCTS) National Cardiac Surgery Database Program is funded by the Department of Health, Victoria, the Health Administration Corporation (GMCT) and the Clinical Excellence Commission (CEC), NSW, and funding

from individual Units. ANZSCTS Research activities are supported through a National Health and Medical Research Council Senior Research Fellowship and Program Grant awarded to C.M. Reid. M.N. Karim has an Australian Postgraduate Awards Scholarship from Monash University. The funding bodies had no role in the design and conduct of the study, the collection, management, analysis and interpretation of the data, or the preparation, review, or approval of the manuscript.

## References


1. Shahian DM, O'Brien SM, Filardo G, Ferraris VA, Haan CK, Rich JB, et al. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 1—coronary artery bypass grafting surgery. *Ann Thorac Surg*. 2009; 89(suppl 1):S2-22.
2. Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R, et al. European system for cardiac operative risk evaluation (EuroSCORE). *Eur J Cardiothorac Surg*. 1999;16:9-13.
3. Reid C, Billah B, Dinh D, Smith J, Skillington P, Yii M, et al. An Australian risk prediction model for 30-day mortality after isolated coronary artery bypass: the AusSCORE. *J Thorac Cardiovasc Surg*. 2009;138:904-10.
4. Steyerberg EW. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. New York: Springer Science; 2009.
5. Kuppermann N, Willits N. In response to "Statistical Models and Occam's Razor." *Acad Emerg Med*. 2000;7:100-3.
6. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J*. 2014;35:1925-31.
7. Harrell FE Jr, Lee KL, Califf RM, Pryor DB, Rosati RA. Regression modelling strategies for improved prognostic prediction. *Stat Med*. 1984;3:143-52.
8. Miller A. *Subset Selection in Regression*. 2nd ed. Boca Raton, FL: Chapman & Hall/CRC; 2002.
9. Austin PC, Tu JV. Automated variable selection methods for logistic regression produced unstable models for predicting acute myocardial infarction mortality. *J Clin Epidemiol*. 2004;57:1138e46.
10. Manor O, Segal E. Predicting disease risk using bootstrap ranking and classification algorithms. *PLOS Comput Biol*. 2013;9:e1003200.
11. Austin PC. Using bootstrap to improve estimation and confidence intervals for regression coefficients selected using backward variable elimination. *Stats Med*. 2008;27:3286-300.
12. Breiman L. Bagging predictors. *Machine Learning*. 1996;24:123-40.
13. Akaike H. Statistical predictor identification. *Ann Inst Math Stat*. 1970;22: 203-17.
14. Schwarz GE. Estimating the dimension of a model. *Ann Stat*. 1978;6:461-4.
15. Billah B, Huq MM, Smith JA, Sufi F, Tran L, Shardey GC, et al. AusSCORE II in predicting 30-day mortality after isolated coronary artery bypass grafting in Australia and New Zealand. *J Thorac Cardiovasc Surg*. 2014;148:1850-5.e2.
16. Billah B, Reid CM, Shardey GC, Smith JA. A preoperative risk prediction model for 30-day mortality following cardiac surgery in an Australian cohort. *Eur J Cardiothorac Surg*. 2010;37:1086-92.
17. Maximus S, Milliken JC, Danielsen B, Khan J, Shemin R, Carey JS. Defining operative mortality: impact on outcome reporting. *J Thorac Cardiovasc Surg*. 2016;151:1101-10.
18. van Buuren S, Boshuizen HC, Knook DL. Multiple imputation of missing blood pressure covariates in survival analysis. *Stat Med*. 1999;18:681-94.
19. Rubin DB. *Multiple Imputation for Nonresponse in Surveys*. Hoboken, NJ: John Wiley & Sons; 1987.
20. StataCorp. *Stata Statistical Software, Release 14*. College Station, TX: Statacorp LP; 2015.
21. Miyata H, Motomura N, Yozu R, Kyo S, Takemoto S. Cardiovascular surgery risk prediction from the patient's perspective. *J Thorac Cardiovasc Surg*. 2011;142: e71-6.
22. Royston P, Ambler G, Sauerbrei W. The use of fractional polynomials to model continuous risk variables in epidemiology. *Int J Epidemiol*. 1999;28: 964-74.
23. Sauerbrei W, Royston P. Building multivariable prognostic and diagnostic models: transformation of the predictors by using fractional polynomials. *J Roy Stat Soc Ser A*. 1999;162:71-94.
24. Wood AM, White IR, Royston P. How should variable selection be performed with multiply imputed data? *Stat Med*. 2008;27:3227-46.
25. Sauerbrei W, Schumacher M. A bootstrap re-sampling procedure for model building: application to the Cox regression model. *Stat Med*. 1992;11: 2093-109.
26. Wears RL, Lewis RJ. Statistical models and Occam's razor. *Acad Emerg Med*. 1999;6:93-4.
27. Murtaugh PA. Methods of variable selection in regression modelling. *Commun Stat Simulation Computation*. 1998;27:711-34.
28. Greenland S. Modelling and variable selection in epidemiologic analysis. *Am J Public Health*. 1989;79:340-9.
29. Derkson S, Keselman HJ. Backward, forward and stepwise automated subset selection algorithms: frequency of obtaining authentic and noise variables. *Br J Math Stat Psychol*. 1992;45:265-82.
30. Hurvich CM, Tsai CL. The impact of model selection on inference in linear regression. *Am Stat*. 1990;44:214-7.
31. Austin PC, Tu JV. Bootstrap methods for developing predictive models. *Am Stat*. 2004;58:131-7.
32. Austin PC. Bootstrap model selection had similar performance for selecting authentic and noise variables compared to backward variable elimination: a simulation study. *J Clin Epidemiol*. 2008;61:1009-17.e1.
33. Ariyaratne TV, Billah B, Yap C-H, Dinh D, Smith JA, Shardey GC, et al. An Australian risk prediction model for determining early mortality following aortic valve replacement. *Eur J Cardiothorac Surg*. 2011;39:815-21.
34. Efron B, Tibshirani R. *An Introduction to the Bootstrap*. New York: Chapman & Hall; 1993.
35. Nashef SAM, Roques F, Sharples LD, Nilsson J, Smith C, Goldstone AR, et al. EuroSCORE II. *Eur J Cardiothorac Surg*. 2012;41:734-45.


**Key Words:** bootstrap resampling, automated model, risk prediction model, cardiac surgery

**TABLE E1. ANZSCTS registry variable information**

| | Variables | Data field number |
|---|---|---|
| 1 | Preoperative (101) | |
| | Patient demography | 24 |
| | Patient risk factor | 19 |
| | Preoperative cardiac status | 33 |
| | Previous intervention | 16 |
| | Hemodynamic state | 09 |
| 2 | Intra- operative (114) | |
| | Operative status | 41 |
| | Minimally invasive | 04 |
| | CPB support | 17 |
| | Coronary bypass | 12 |
| | Valve surgery | 40 |
| 3 | Postoperative (72) | |
| | Postoperative support | 15 |
| | Complication | 36 |
| | Mortality/readmission | 21 |
| Total | | 287 |

*CPB*, Cardiopulmonary bypass.

**TABLE E2. Number (%) of missing observations for predictors in the development dataset**

| | Missing values | |
|---|---|---|
| **Plausible variables** | **n** | % |
| Age | 13 | 0 |
| Sex | 0 | 0 |
| Peripheral vascular disease | 215 | 0.34 |
| Stroke | 0 | 0 |
| Glomerular filtration rate | 576 | 0.91 |
| Infective endocarditis | 207 | 0.33 |
| Urgency of procedure | 37 | 0.06 |
| Ejection fraction | 1951 | 3.07 |
| Inotropic medication use | 207 | 0.33 |
| Type of procedure | 786 | 1.24 |
| Previous cardiac surgery | 182 | 0.29 |
| NYHA class | 3355 | 5.28 |
| Angina | 581 | 0.91 |
| Shock | 199 | 0.31 |
| Congestive heart failure | 195 | 0.31 |
| Myocardial infarction | 172 | 0.27 |
| Resuscitation | 201 | 0.32 |
| Arrhythmia | 220 | 0.35 |
| Hypertension | 207 | 0.33 |
| Left main disease <50% | 581 | 0.91 |
| Hypercholesterolemia | 217 | 0.34 |
| Respiratory disease | 195 | 0.31 |
| Steroid use | 211 | 0.33 |
| Diabetes | 327 | 0.52 |
| Pacemaker in situ | 429 | 0.68 |
| Body mass index | 339 | 0.53 |
| Number of disease vessel | 703 | 1.11 |
| Immunosuppressant use | 206 | 0.32 |
| Intravenous nitrate use | 211 | 0.33 |
| Anticoagulant use | 221 | 0.35 |
| Mortality in 30 d | 280 | 0.44 |

*NYHA*, New York Heart Association.

**TABLE E3. Descriptive statistic s of plausible categorical variables**

| Variables | Frequency | Percent |
|---|---|---|
| Sex | | |
| Male | 45,854 | 72.18 |
| Female | 17,669 | 27.82 |
| Family history of heart disease | | |
| No | 38,480 | 60.58 |
| Yes | 19,318 | 30.41 |
| Previous cardiac surgery | | |
| No | 57,918 | 91.18 |
| Yes | 5423 | 8.54 |
| Peripheral vascular disease | | |
| No | 56,723 | 89.3 |
| Yes | 6585 | 10.37 |
| Respiratory disease | | |
| No | 54,558 | 85.89 |
| Yes | 8770 | 13.81 |
| Cerebrovascular disease | | |
| No | 56,116 | 88.34 |
| Yes | 7198 | 11.33 |
| Stroke | | |
| No | 62,198 | 97.91 |
| Yes | 1325 | 2.09 |
| Hypertension | | |
| No | 17,273 | 27.19 |
| Yes | 46,043 | 72.48 |
| Diabetes | | |
| No DM | 45,282 | 71.28 |
| DM + no insulin | 13,442 | 21.16 |
| DM + insulin | 4534 | 7.14 |
| Hypercholesterolemia | | |
| No | 20,351 | 32.04 |
| Yes | 42,955 | 67.62 |
| IE | | |
| No IE | 61,754 | 97.22 |
| IE treated | 553 | 0.87 |
| IE active | 1009 | 1.59 |
| Left main disease | | |
| No | 51,664 | 81.33 |
| Yes | 11,278 | 17.75 |
| Number of diseased vessels | | |
| None | 15,875 | 24.99 |
| One | 5264 | 8.29 |
| Two | 11,347 | 17.86 |
| Three | 30,334 | 47.75 |
| NYHA class | | |
| I, II | 42,367 | 66.7 |
| III | 13,599 | 21.41 |
| IV | 4202 | 6.61 |
| Cardiogenic shock | | |
| No | 61,820 | 97.32 |
| Yes | 1504 | 2.37 |
| Resuscitation | | |
| No | 62,559 | 98.48 |
| Yes | 763 | 1.2 |
| Congestive heart failure | | |
| Yes | 47,953 | 75.49 |
| No | 15,375 | 24.2 |
| Ejection fraction | | |
| Normal, >60% | 31,391 | 49.42 |
| Mild, 46%-60% | 18,321 | 28.84 |
| Moderate, 31%-45% | 8689 | 13.68 |
| Severe, ≤30 | 3171 | 4.99 |
| Arrhythmia | | |
| No | 53,078 | 83.56 |
| Yes | 10,225 | 16.1 |
| Pacemaker in situ | | |
| No | 61,890 | 97.43 |
| Yes | 1204 | 1.9 |
| Angina | | |
| No angina | 21,225 | 33.41 |
| Stable angina | 30,664 | 48.27 |
| Unstable angina | 11,053 | 17.4 |
| Previous MI | | |
| No | 38,817 | 61.11 |
| Yes | 24,534 | 38.62 |
| Inotrope use | | |
| No | 61,630 | 97.02 |
| Yes | 1686 | 2.65 |
| Intravenous nitrates | | |
| No | 60,078 | 94.58 |
| Yes | 3234 | 5.09 |
| Anticoagulant use | | |
| No | 50,973 | 80.24 |
| Yes | 12,329 | 19.41 |
| Steroids use | | |
| No | 61,806 | 97.3 |
| Yes | 1506 | 2.37 |
| Immunosuppressive use | | |
| No | 61,784 | 97.26 |
| Yes | 1533 | 2.41 |
| Procedure type | | |
| Isolated CABG | 36,085 | 56.81 |
| Valve(s) only | 10,489 | 16.51 |
| Valve(s) + CABG | 6746 | 10.62 |
| Other | 9417 | 14.82 |
| Urgency of operation | | |
| Elective | 41,829 | 65.85 |
| Urgent | 18,478 | 29.09 |
| Emergency | 2924 | 4.6 |
| Salvage | 255 | 0.4 |
| 30-d mortality | | |
| No | 61,540 | 97.31 |
| Yes | 1703 | 2.69 |

*(Continued)*

**TABLE E3. Continued**

*DM*, Diabetes mellitus; *IE*, infective endocarditis; *NYHA*, New York Heart Association; *MI*, myocardial infarction; *CABG*, coronary artery bypass grafting.

**TABLE E4. Descriptive statistic s of plausible numerical variables**

| Predictors | n | Mean | SD |
|---|---|---|---|
| Age, y | 63,510 | 65.67 | 12.83 |
| Body mass index | 63,351 | 28.48 | 9.74 |
| Estimated glomerular filtration rate | 62,947 | 82.15 | 37.10 |

*SD*, Standard deviation.

**TABLE E5. Bootstrap ROC model**

| Predictors | OR (95% CI) | *P* value |
|---|---|---|
| Age | 1.024 (1.018-1.029) | <.001 |
| Sex (female) | 1.414 (1.265-1.581) | <.001 |
| eGFR | 0.990 (0.987-0.992) | <.001 |
| Infective endocarditis (ref: none) | | |
| Treated endocarditis | 1.205 (0.745-1.948) | .446 |
| Active endocarditis | 2.437 (1.874-3.169) | <.001 |
| Urgency (ref: elective) | | |
| Urgent | 1.567 (1.371-1.791) | <.001 |
| Emergency/salvage | 4.611 (3.884-5.475) | <.001 |
| Ejection fraction (ref: >60%) | | |
| Mild (46%-60%) | 1.242 (1.081-1.427) | .002 |
| Moderate (31%-45%) | 1.634 (1.397-1.910) | <.001 |
| Severe (≤30%) | 2.346 (1.895-2.904) | <.001 |
| NYHA class (ref: I and II) | | |
| III | 1.467 (1.279-1.681) | <.001 |
| IV | 1.725 (1.463-2.032) | <.001 |
| Procedure type (ref: CABG) | | |
| Valve | 1.462 (1.209-1.767) | .001 |
| CABG + valve | 2.043 (1.738-2.402) | <.001 |
| Others | 2.968 (2.546-3.460) | <.001 |
| Peripheral vascular disease | 1.617 (1.406-1.859) | <.001 |
| Stroke | 5.865 (4.985-6.901) | <.001 |
| Inotrope administration | 2.293 (1.916-2.743) | <.001 |
| Angina | | |
| Stable | 0.930 (0.810-1.066) | .297 |
| Unstable | 1.260 (1.066-1.489) | .007 |
| Previous cardiac surgery | 1.687 (1.459-1.951) | <.001 |

*OR*, Odds ratio; *CI*, confidence interval; *eGFR*, estimated glomerular filtration rate; *NYHA*, New York Heart Association; *CABG*, coronary artery bypass grafting.

## 6.3 Summary

This chapter investigated different approaches to variable selection as an avenue for balancing the parsimony and predictive performance of a cardiac surgical risk prediction model. The study compared the variable selection methods that have been previously used in cardiac surgical risk prediction modeling. The findings in the paper showed that the balance between parsimony and predictive performance can be best achieved using the bootstrap bagging method, in conjunction with receiver operating characteristics, when developing a risk prediction model. The study resulted in a recommendation for the use of the bootstrap bagging method when developing future risk prediction models. This approach to variable selection has therefore been used in the study described in the following chapter, where we develop a new prediction model.

# Chapter 7: Predicting long-term survival following Coronary Artery Bypass Surgery

## 7.1 Introduction

Short-term mortality predictions are commonly used to evaluate pre-operative risk in patients undergoing coronary artery bypass graft (CABG) surgery (37). However, short-term mortality does not provide adequate information to guide long-term post-CABG patient management (38, 82). Long-term mortality risk is becoming increasingly important in informing patient management strategies following CABG surgery (39, 40), however, there are currently very few prediction models for long-term mortality risk following CABG surgery. In addition, those that do exist have been developed in the US and may not generalise well to other populations. The aim of the study in this chapter was to develop and validate a risk prediction model for long-term mortality risk following CABG surgery using an Australian CABG patient cohort.

The new risk prediction model was developed whilst keeping in mind all of the issues identified in the systematic review (chapter 3) as well as the knowledge accumulated in subsequent chapters (4, 5 and 6). This chapter includes the article entitled 'Predicting long-term survival after coronary artery bypass graft surgery', which has been submitted for publication to the Interactive Cardiovascular and Thoracic Surgery.

## 7.2. Manuscript

**Declaration for Thesis Chapter 7**

**Manuscript**: Predicting long-term survival after coronary artery bypass graft surgery (Under review)

**Declaration by candidate**

In the case of Chapter 7, the nature and extent of my contribution to the work was the following:

| Nature of contribution | Extent of contribution (%) |
|---|---|
| Study concept and design, literature search, data analysis and interpretation, manuscript development and preparation | 70% |

The following co-authors contributed to the work. Where co-authors are students at Monash University, the extent of their contribution in percentage terms is stated:

| Name | Nature of contribution | Extent of contribution (%) for student co-authors only |
|---|---|---|
| Billah, B | Study concept, data analysis and manuscript editing | |
| Reid, C | Study concept, study design and manuscript editing | |
| Cochrane, A | Study concept, Interpretation of result and manuscript editing | |
| Tran, L | Interpretation of result and manuscript editing | |
| Brilleman, S | Data analysis and interpretation | 5% |

The undersigned hereby certify that the above declaration correctly reflects the nature and extend o the candidate's and co-author' contributions to this work

| | | |
|---|---|---|
| **Candidate's Signature** | [redacted] | **Date 10/03/2017** |
| **Main Supervisor's Signature** | [redacted] | **Date 10/03/2017** |

## Predicting long-term survival after coronary artery bypass graft surgery

Md N Karim[1], Christopher M Reid[1,2], Mollah Huq[1], Samuel L Brilleman[1], Andrew Cochrane[3], Lavinia Tran[1], Baki Billah[1]

[1]School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia
[2]School of Public Health, Curtin University, WA, Australia
[3]Dept. of Cardiothoracic Surgery, Monash Medical Centre, Clayton, Australia

Corresponding Author
**Baki Billah**
Senior Lecturer of Biostatistics
School of Public Health and Preventive Medicine, Monash University
The Alfred Centre, (level 6), 99 Commercial Road, Victoria - 3004, Australia
Email [redacted]
Ph [redacted]
Fax: +61(0)399030828

Manuscript text word count: 5197 (Including title page and abstract)

## Predicting long-term survival after coronary artery bypass graft surgery


### Abstract

**Objective**

To develop a model for predicting long-term survival following coronary artery bypass graft (CABG) surgery.

**Methods**

This study included 46,573 patients from the Australian and New Zealand Society of Cardiac and Thoracic Surgeons (ANZCTS) registry, who underwent isolated CABG surgery between 2001 and 2014. Data were randomly split into development (23282) and validation (23291) samples. Cox regression models were fitted separately, using the important pre-operative variables, for four 'time intervals' (31-90 days, 91-365 days, 1-3 years and >3 years), with optimal predictors selected using the bootstrap bagging technique. Model performance was assessed in both validation data and in combined data (development and validation samples). Coefficients of all four final models were estimated on the combined data adjusting for hospital-level clustering.

**Result**

Kaplan-Meier mortality rates estimated in the sample were 1.7% at 90 days, 2.8% at 1 year, 4.4% at 2 years and 6.1% at 3 years. Age, peripheral vascular disease, respiratory disease, reduced ejection fraction, renal dysfunction, arrhythmia, diabetes, hypercholesterolemia, cerebrovascular disease, hypertension, congestive heart failure, steroid use and smoking were included in all 4 models. However, their magnitude of effect varied across the time intervals. Models showed excellent discrimination in both development and validation dataset. Harrell's C-statistic was 0.83, 0.78, 0.75, and 0.74 for the 31-90 days, 91-365 days, 1-3 years and >3 years models, respectively. Overfitting-corrected calibration curves demonstrated excellent model calibration.

**Conclusion**

Models were developed for predicting long-term survival at four time-intervals after isolated CABG surgery. These models can be used in conjunction with the existing 30-day mortality prediction model.


**Word count:** 250


**Key words**

CABG, long-term survival, risk prediction model, risk stratification, cardiac surgery, coronary revascularization.

## Predicting long-term survival after coronary artery bypass graft surgery

### Introduction

The prediction of 30-day or in-hospital mortality is popularly used to evaluate operative risk in cardiac surgery (1-4). However, this short-term mortality does not provide adequate information to guide long-term post-surgery patient management (5). Due to advancements in surgical technologies and perioperative care, operative and 30-day mortality rates have declined over the last few decades and consequently more attention is now required towards improving long-term survival following cardiac surgery, which is becoming increasingly important in informing patient management strategies following CABG surgery (6, 7). Prediction of long-term survival can be used to determine the most appropriate post-discharge care strategies. This would essentially help patients and their doctors to implement behavioural and therapeutic modifications to optimize benefit from surgery (6). Besides, these models can be used for various scientific purposes and to facilitate research.

EuroSCORE, a short-term mortality risk prediction model, has been shown to predict intermediate to long term survival following cardiac surgery (8). It is expected that the short-term models may to some extent predict long-term mortality risk as most predictors are similar. However, this does not justify use short-term risk model for prediction long-term survival. AusSCORE, EuroSCORE etc were not intended for predicting long-term survival and, their development process was not based on survival analysis which allows the time-varying nature of the risk (i.e. hazard) of the event. Shahian et al. (7) showed that the impact of predictor variables on mortality fluctuates as time following surgery increases. Hence separate models for predicting long-term survival may be needed. Two such models have been developed in the United States (US) in the Cardiac Surgery Reporting System (CSRS) (9) and the Society of Thoracic Surgeons (STS) (10) databases.

No model is currently available for predicting long-term survival following coronary artery bypass graft (CABG) surgery in Australian patients. It is widely recognised that a risk prediction model will generally predict outcomes more accurately in the population setting where it was originally developed (11, 12). Therefore, the aim of the current study was to develop a risk prediction model for predicting long-term survival following CABG surgery using an Australian patient cohort.

## Material and Methods

### Dataset

The study used data from 46,573 patients, included in the Australian and New Zealand Society of Cardiac and Thoracic Surgeons (ANZCTS) registry, who underwent isolated CABG surgery between 2001 and 2014. The ANZCTS registry collects preoperative, intraoperative and postoperative variables using internationally standardized data definitions, on adult patients undergoing cardiac surgery in 28 hospitals across Australia. The data collection and its audit methods have been discussed elsewhere (13). In-hospital and 30-day mortality data were collected by the registry. Outcome of the model was long term survival following cardiac surgery. Mortality data outside 30 days post-surgery were collected through linkage with the National Death Index (NDI) database.

The data were divided, at a ratio of 1:1, into development (23282) and validation (23291) set. The analysis in this study involved 30 plausible preoperative variables identified through a variety of methods, including literature review, clinical acumen, or their use in other models developed using the same database.

### Statistical Analysis

Missing data

The variable 'family history of heart disease' (10.8%) had the highest percentage of missing data, followed by 'NYHA classification' (3.8%), 'reduced ejection fraction' (2.2%) and 'renal dysfunction' (1.3%). The remaining predictors each had < 1% missing observations (supplementary table in appendix 4.1). Missing data were imputed using the Multiple Imputation by Chained Equations (MICE) method. Ten imputations were generated. The analysis was performed separately on each imputed dataset and then final parameter estimates were obtained by aggregating across the imputed datasets (14).

Model development

Univariable associations between preoperative patient characteristics and mortality were assessed using Univariable Cox regression. Previous studies have shown that the effects of some variables on

mortality depends on the time since CABG surgery (7). To accommodate the fact that the effect of each of the risk factors on mortality differs across time (non–proportional hazards in single Cox model). Four separate Cox regression models were fitted, to generate piecewise hazard, forcing same set of variables into these models. Selection of the four-time interval was done, using the technique adopted by Shahian et al while developing STS long-term mortality model (7). The first-time interval started at 31 days to maintain continuum with the existing AusSCORE II model that predicts 30-day mortality following CABG surgery (15). The first interval included a range up to 90 days since recent advancements in modern critical care mean there is now an increased capacity for postoperative care and, therefore, the potential for an extension of early postoperative period; some already consider 90-day mortality as a new convention or benchmark (1, 7). The $2^{nd}$, $3^{rd}$ and $4^{th}$ time intervals were decided based on a preliminary analysis that involved fitting Cox regression models with several relatively narrow intervals (each spanning 90 days), then collapsing adjacent intervals in to larger intervals (1years, 3 years and > 3 years), while retaining sufficient events in each merged interval to ensure precise estimation of interval-specific hazard ratio (7).

Bootstrap bagging techniques were used to select predictors for the multivariable models (16). A bootstrap sample of the same size as the development sample was drawn from each of the imputed datasets. For each bootstrap sample, all plausible risk factors were entered into a multivariable Cox regression model and the p-value for each variable in the model was calculated. A predictor with a p-value less than or equal to 0.05 was considered as significant. For each imputed dataset, 1000 bootstrap samples were taken, and the percentage of times that each predictor appeared as significant across the 1000 bootstraps was recorded (bootstrap coverage). Bootstrap coverage of each predictor was averaged across 10 imputed datasets to generate an overall coverage for each predictor (17). The predictors were then ranked per their overall bootstrap coverage (Supplementary table in appendix 4.2).

Fourteen multivariable Cox regression models were then fitted with the predictors that achieved at least 50% overall bootstrap coverage (18). The first model comprised 6 predictors which each achieved 100% overall bootstrap coverage. Thirteen subsequent models were generated through adding one variable at a time to the model, based on decreasing rank per the overall bootstrap coverage (Supplementary table in appendix 4.2). The area under the receiver operating characteristic curve (AUC) was calculated for each of these 14 models to provide an estimate of model discrimination. The model with highest AUC value was selected as the final model.

For the final model, non-linearity of continuous predictors (age) was considered by fitting fractional polynomials in the Cox regression model (19) and using a sensitivity analysis to assess whether the inclusion of a non-linear term changes the model fit. However, there was little improvement in discrimination or calibration with the inclusion of non-linear terms and, hence, the final model retained linear terms for each of the continuous variables. The first order interaction effects between clinically relevant risk factors were also investigated. Interaction effects between some pairs of predictor variables appeared significant ($p < 0.05$), however their inclusion did not improve model performance and therefore, only the main effects were retained in the final model.

Model performance and validation

Model performance was assessed first in the validation dataset. Subsequently multi-fold (k=100) cross validation was done in combined datasets (development and validation set) to avoid optimistic prediction. Finally, Harrell's C-statistics, a global measure for the assessment of a fitted survival model for the continuous event time, (7, 20) was generated in the combined dataset.

Calibration of the final model was assessed using the Regression Modelling Strategies (RMS) package version 4.4-2 in the R statistical software (21). Bootstrap resampling was used to get overfitting-corrected estimates of predicted survival probabilities. Locally weighted scatter-plot smoother (LOESS) calibration curves were generated for each of the four time intervals plotting these probabilities against corresponding Kaplan-Meier survival estimates, stratifying on intervals of predicted survival.

Final model estimation

Coefficients of all four final models were estimated on the combined data (development and validation samples) including a hospital-level random effect in the model to account for hospital-level clustering (22). Coefficients (and standard errors) for the smoothed baseline hazard, was generated using the approach proposed by Royston et al (23).

Statistical software

Statistical software packages Stata version 14 (StataCorp. Release 14; 2015) and R version 3.3.2 were (R core team version 3.3.2, 2013) used for the analyses.

**Ethical approval**

The Institutional Review Board of each participating hospital had approved the use of their data for research purposes (Alfred HREC:262/09). The ANZCTS registry has approved collection of patient data using an 'opt-out consent approach' (MUHREC: CF08/0322-2008000065). The current study received ethical approval from the Monash University Standing Committee on Ethics in Research Involving Humans (SCERH) (MUHREC:CF14/1117–2014000476).

## Results

Supplementary table in the appendix 4.3 presents the preoperative characteristics of the 46,573 patients. Mean ± standard deviation (sd) age of the patients at surgery was 65.9±10.4 and 79.4% of them were male. Median follow-up time was 4.2 (IQR 1.8–7.0) years.

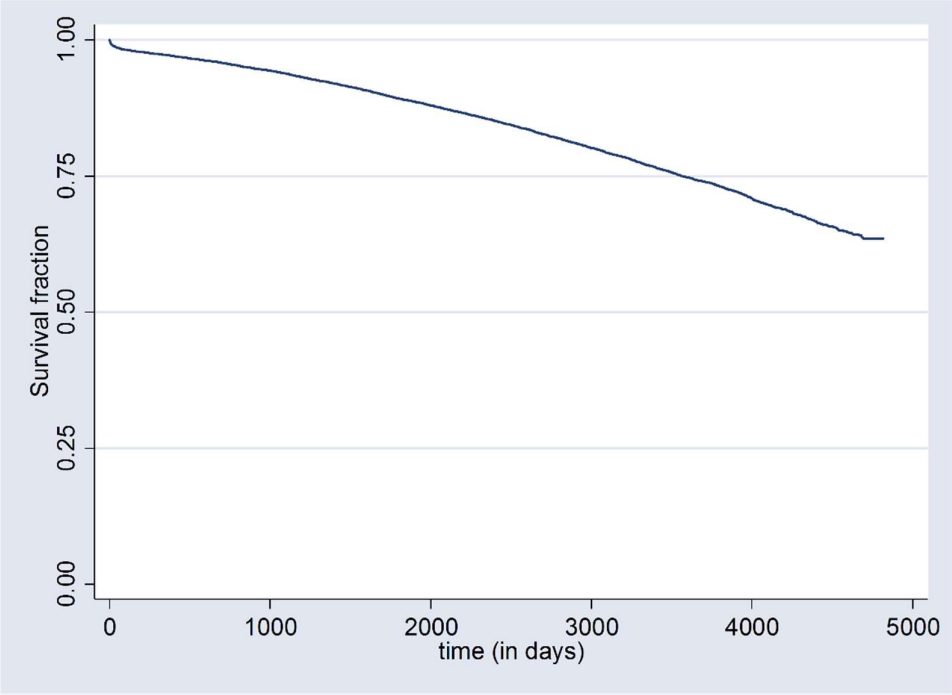



Figure 1: Kaplan-Meier estimates of mortality in the study sample

Figure 1 presents the Kaplan-Meier estimates of mortality in the study sample. Kaplan-Meier mortality rate estimates for the study sample were 1.72 % at 90 days, was 2.81% at 1 year, 4.36 % at 2 years and 6.14% at 3 years (Table 1).

Table 1: Kaplan-Meier mortality rate estimates for the study sample

| | Development | | Validation | | Overall sample | | |
|---|---|---|---|---|---|---|---|
| Time interval | Total patients | Death | Total patients | Death | Total patients | Death | KM estimates |
| 30 days | 21822 | 272 | 21851 | 250 | 43673 | 522 | 1.17 (1.08, 1.28) |
| 90 days | 21258 | 114 | 21294 | 123 | 42552 | 237 | 1.72 (1.60, 1.84) |
| 1 year | 19088 | 221 | 19081 | 231 | 38169 | 452 | 2.81 (2.66, 2.97) |
| 2 years | 16190 | 300 | 16338 | 268 | 32527 | 568 | 4.36 (4.17, 457) |
| 3 years | 13705 | 297 | 13813 | 265 | 27518 | 562 | 6.14 (5.69, 6.39) |
| 4 years | 11167 | 291 | 11309 | 304 | 22476 | 595 | 8.35 (8.06, 8.65) |

*Figure s in parentheses denotes 95% CI

Supplementary table in appendix 4.3 presents the univariable associations between preoperative characteristics and mortality using univariable Cox regression for each of the time intervals (31-90 days, 91-365 days, 1-3 years, and >3 years). EF < 30 was strongly associated with mortality at the '31-90 days' interval (hazard ratio (HR) = 7.82, 95% confidence intervals (CI): 5.24 to 11.67), however the magnitude of its association with mortality diminished steadily over time (91-365 days HR = 5.18, 95% CI: 3.70 to 7.24; 1-3 years HR = 3.55, 95% CI: 2.82 to 4.47; >3 years HR = 2.52, 95% CI: 2.18 to 2.89). Severe renal dysfunction was strongly associated with mortality at the '31-90 days' interval (HR = 21.4, 95% CI: 11.45 to 39.84), whilst its association with mortality diminished over time, (91-365 days HR = 7.31, 95% CI: 4.59 to 11.67; 1-3 years HR = 6.31, 95% CI: 4.71 to 8.45; >3 years HR = 6.08, 95% CI: 5.02 to 7.36). Similar associations were evident for respiratory disease, congestive heart failure, steroid use and New York Heart Association (NYHA) class. Each of Body Mass index (BMI) >25 kg/$m^2$, Inotrope use, previous cardiac surgery, cardiogenic shock, IV nitrite use, resuscitation and urgency of operation showed strong associations with mortality at the earlier time intervals, namely 31-90 days and 91-365 days, however, their associations with mortality were less evident during later time intervals.

Table 2 presents the HR and 95% CI from the multivariable Cox regression models estimated at each of the four time intervals. Thirteen predictors including age, peripheral vascular disease, respiratory disease, reduced EF, renal dysfunction, smoking history, arrhythmia, diabetes, hypercholesterolemia, cerebrovascular disease, hypertension, congestive heart failure and steroid use appeared in all four models. However, the magnitude of their association with mortality varies over time. Peripheral vascular disease (HR = 1.24, 95% CI: 0.91-1.07) and congestive heart failure at current admission (HR = 1.43, 95% CI: 0.99-2.05) were not significantly associated with mortality at 31-90 days, but they were associated with mortality at later periods.

Table 2: Cox proportional hazards models for long-term survival following CABG surgery

| **Predictors** | **31 day - 90 days** HR (95% CI) | **91 days - 1 years** HR (95% CI) | **1+ year - 3 years** HR (95% CI) | **> 3 years** HR (95% CI) |
|---|---|---|---|---|
| Age | 1.05 (1.03, 1.07) | 1.06 (1.05, 1.08) | 1.05 (1.04, 1.06) | 1.07 (1.06, 1.08) |
| Peripheral vascular disease | 1.24 (0.91, 1.67) | 1.76 (1.42, 2.18) | 1.64 (1.42, 1.89) | 1.48 (1.36, 1.61) |
| Respiratory disease | 2.03 (1.53, 2.70) | 1.25 (0.99, 1.58) | 1.40 (1.21, 1.63) | 1.39 (1.27, 1.52) |
| Ejection fraction: 46-60% | 1.34 (0.95, 1.90) | 1.55 (1.21, 1.98) | 1.49 (1.29, 1.73) | 1.23 (1.13, 1.34) |
| Ejection fraction: 30-45% | 2.10 (1.47, 3.01) | 2.42 (1.87, 3.13) | 1.96 (1.67, 2.31) | 1.48 (1.34, 1.63) |
| Ejection fraction: <30% | 4.11 (2.65, 6.36) | 3.12 (2.18, 4.46) | 2.47 (1.94, 3.15) | 1.78 (1.54, 2.07) |
| Renal dysfunction: mild | 1.44 (0.81, 2.55) | 0.88 (0.62, 1.25) | 1.04 (0.85, 1.28) | 1.01 (0.89, 1.15) |
| Renal dysfunction: moderate | 2.21 (1.23, 3.99) | 1.43 (0.98, 2.07) | 1.23 (0.98, 1.54) | 1.28 (1.11, 1.47) |
| Renal dysfunction: severe | 5.99 (3.05,11.78) | 2.14 (1.29, 3.52) | 2.21 (1.61, 3.04) | 1.90 (1.55, 2.34) |
| On dialysis | 9.23 (4.67,18.26) | 4.80 (2.96, 7.77) | 4.17 (3.05, 5.71) | 3.55 (2.76, 4.52) |
| Smoking | 1.36 (1.02, 1.81) | 1.59 (1.28, 1.98) | 1.28 (1.12, 1.47) | 1.37 (1.27, 1.48) |
| Arrhythmia | 2.39 (1.80, 3.18) | 1.68 (1.33, 2.11) | 1.34 (1.14, 1.57) | 1.30 (1.17, 1.44) |
| Diabetes: no treatment | 0.86 (0.46, 1.48) | 1.18 (0.80, 1.73) | 0.88 (0.66, 1.16) | 1.08 (0.94, 1.25) |
| Diabetes: on drug | 1.31 (0.96, 1.79) | 1.29 (1.03, 1.62) | 1.20 (1.03, 1.39) | 1.23 (1.12, 1.35) |
| Diabetes: on insulin | 1.49 (1.03, 2.14) | 1.21 (0.90, 1.62) | 1.38 (1.15, 1.65) | 1.60 (1.43 1.79) |
| Hypercholesterolemia | 1.03 (0.73, 1.46) | 0.86 (0.67, 1.08) | 0.81 (0.70, 0.94) | 0.77 (0.71, 0.84) |
| Cerebrovascular disease | 1.77 (1.32, 2.37) | 1.03 (0.80, 1.33) | 1.45 (1.25, 1.69) | 1.24 (1.13, 1.37) |
| Hypertension | 1.35 (0.89, 2.03) | 1.16 (0.88, 1.54) | 1.10 (0.92, 1.30) | 1.14 (1.04, 1.25) |
| CHF: Past | 1.15 (0.78, 1.69) | 1.69 (1.30, 2.20) | 1.15 (0.95, 1.39) | 1.26 (1.14, 1.39) |
| CHF: current | 1.43 (0.99, 2.05) | 1.45 (1.09, 1.94) | 1.30 (1.07, 1.59) | 1.27 (1.12, 1.44) |
| Steroid use at surgery | 1.89 (1.06, 3.37) | 1.49 (0.87, 2.57) | 2.48 (1.87, 3.28) | 1.53 (1.21, 1.93) |
| Harrell's C statistics | 0.8308 | 0.7813 | 0.7448 | 0.7403 |

Hypertension was associated with mortality only after 3 years post-surgery (HR = 1.16, 95%CI: 1.05 to 1.27). Diabetes on insulin, steroid use, and cerebrovascular disease appeared as significant predictors in the models for 1-3 years and >3 years. Older age and smoking were strongly associated with mortality, with similar magnitudes of hazard ratios across time periods. Respiratory disease, reduced EF, severe renal dysfunction and arrhythmia were significantly associated with mortality, with decreasing magnitude of hazard ratios over time. Hypercholesterolemia was not significant in the first of the two time-intervals, but appeared as protective factor after 1 year onward (HR 0.81 at 1-3 years and HR 0.77 at >3 years). Supplementary table in appendix 4.5 presents baseline hazard coefficients and standard error of four models (22).

Model discrimination (AUC) in the validation set was 0.835 (95% CI: 0.802 to 0.868) for predicting 31–90 days survival, 0.791 (95% CI: 0.763 to 0.818) for predicting 91-365 days survival, 0.747 (95% CI: 0.727 to 0.768) for predicting 1-3-year survival and 0.737 (95% CI: 0.725 to 0.749) for predicting >3-

year survival. ROC curves for model discrimination in validation dataset (Figure 2) and in combined dataset (Appendix 4.6) shows excellent discrimination. Model discrimination AUC) in the in multi-fold cross-validation was 0.833 (95% CI: 0.827 to 0.839) within the 31-90 days interval, 0.791 (95% CI: 0.786 to 0.796) within the 91-365 days interval, 0.753 (95% CI: 0.751 to 0.755) within the 1-3-year interval, and 0.739 (95% CI: 0.737 to 0.742) after 3 years.

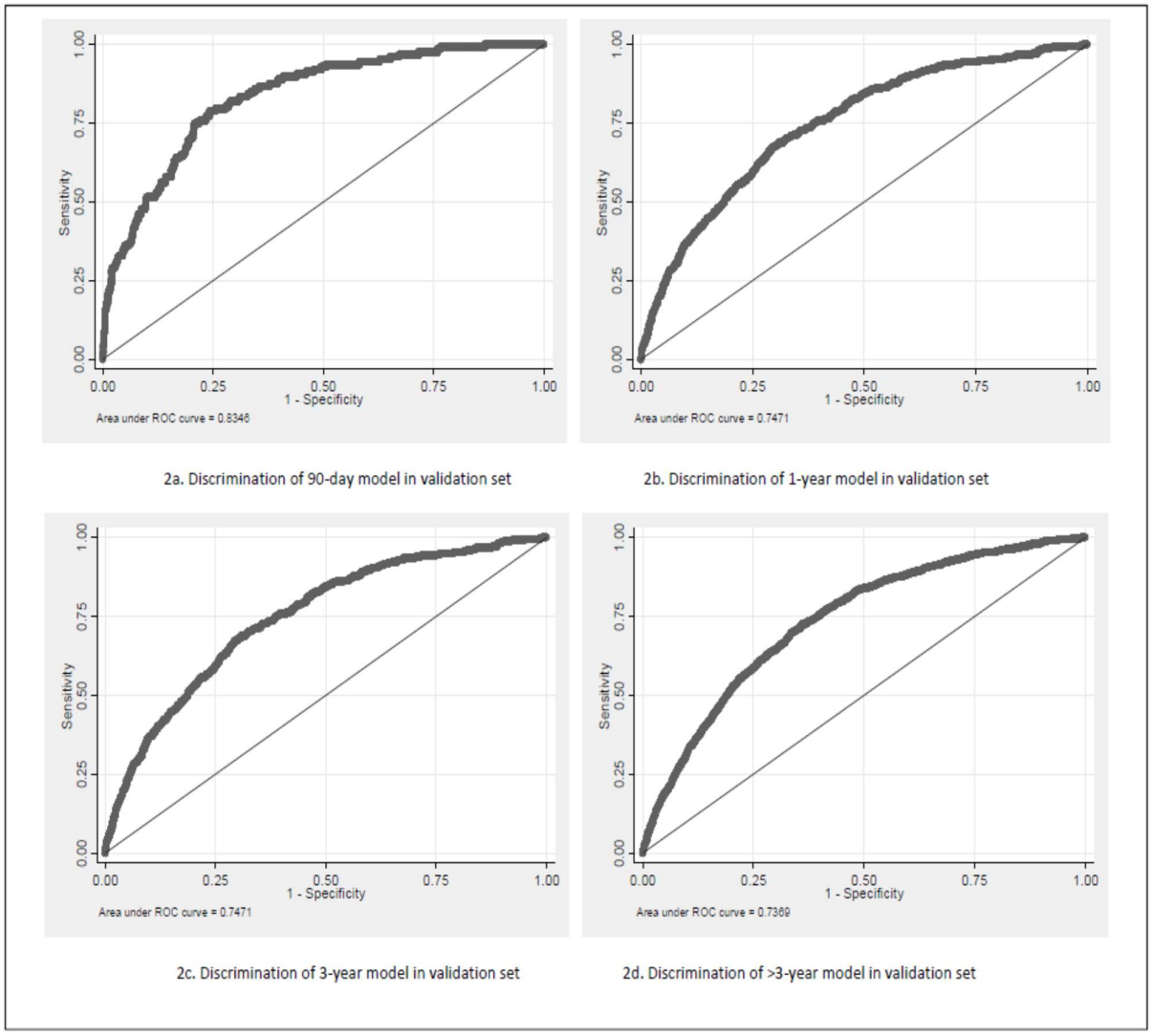



Figure 2: Model discrimination curves for the four time intervals in the validation dataset.

The Harrell's C statistics for the four period-specific Cox regression models were 0.83, 0.78, 0.75, 0.74 at 31–90 days, 91-365 days, 1-3 years, and >3 years respectively. All four LOESS calibration curves show minimal error, where error is defined as the difference between the predicted values and the corresponding bias-corrected calibrated values, demonstrating excellent calibration of the models for all four time intervals (Figure 3).

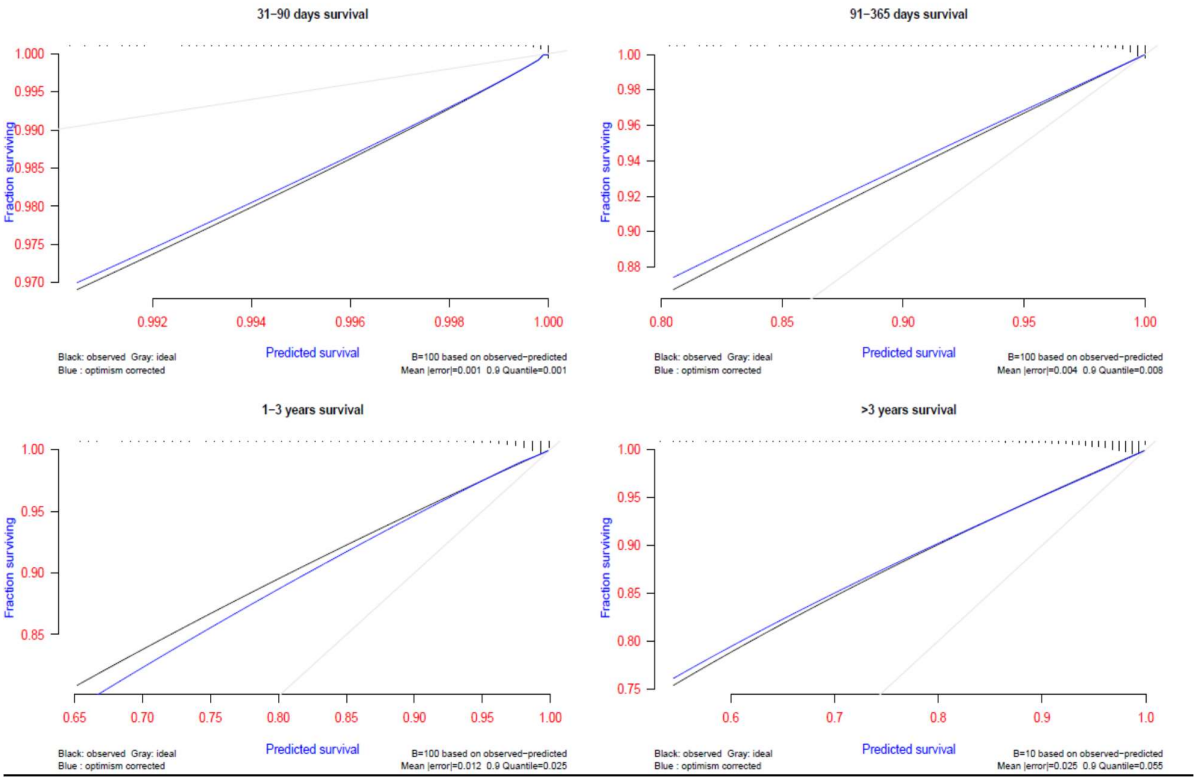



Figure 3: Overfitting-corrected LOESS non-parametric calibration curves, demonstrating calibration of the survival models

## Discussion

In the current study, a set of models have been developed for predicting long-term mortality risk at four distinct time intervals (31-90 days, 91-365 days, 1-3 years, >3 years), recognising the fact that the effect of various risk factors on mortality may differ depending on the time since CABG surgery.

The models developed in this study are expected to supplement the previously published AusSCORE II (15) model that predicts 30-day mortality following CABG surgery. The first of the four models developed in this study was for 31-90 days mortality risk which ensures continuity with the existing AusSCORE II model. The rationale for keeping 90 days as the upper bound of the interval for the first-time period was the potential expansion of the early postoperative period due to improvements in surgical techniques, postoperative care and most importantly the critical care system (7). The increased capacity of the medical system for resuscitating critical postoperative patients, as well as the use of advanced mechanical and pharmacological support, has increasingly delayed the death of many seriously ailing postoperative patients. Given that these patients are now more likely to die

outside of 30 days post-surgery, 30-day mortality alone is likely to underestimate the true rate of operative deaths (1, 7). Accordingly, a 31-90 day mortality risk model should be used to supplement 30-day mortality risk information obtained from a short-term mortality risk prediction model such as the AusSCORE II. The remaining three time intervals provide an opportunity to estimate survival probabilities beyond the period of operative death.

Many of the significant predictors of short-term mortality reported in AusSCORE II (15) did not appear in the long-term models developed in this study and vice-versa. This finding supports the post-surgery mortality risk pattern that long-term outcomes of surgery are less affected by conventional predictors of early mortality, such as emergency status and cardiogenic shock (24). Whereas late mortality is more strongly related to comorbidities and chronic conditions such as diabetes and renal impairment, as well as behavioural characteristics such as smoking (7). Gardner et al. also reported similar pattern, most of their short-term mortality predictors were cardiac-related variables, whereas, most of their longer-term mortality predictors were noncardiac-related variables (25).

The findings in this study underpin the importance of behavioral characteristics, functional status and comorbidities in predicting longer-term survival following CABG surgery. Smoking history—history of any tobacco consumption—which did not appear in the AusSCORE II model, did appear as a significant predictor in all the models developed in this study. Herlitz et al. also showed an association between smoking and 5-year mortality following CABG (26). A study by Saxena et al. using an Australian CABG cohort reported an increased risk of pulmonary complications and reduced long-term survival among smoking patients (27). This may be because of a permanent pre-operative injury due to smoking, or may be because previous smokers are much more likely to restart smoking at some point after surgery than pre-existing non-smokers. Respiratory problems also showed a similar association with mortality in the current study, confirming that respiratory complications may be seen as an intermediate pathway to mortality.

Among the preoperative cardiac conditions, only reduced EF appeared as an independent predictor in all long-term models developed in the current study as well as in the 30-day mortality model reported in AusSCORE II. Among the comorbid conditions, renal impairment, peripheral vascular disease and cerebrovascular disease appeared as independent predictors in all survival models as well as in AusSCORE II. Since preoperative reduced EF and the aforementioned preoperative comorbidities were associated with both short-term and long-term mortality following CABG, these probably form the core set of predictors that contribute to mortality risk after CABG at all times following surgery. Hence

caution should be taken with patients who present with these comorbid conditions prior to CABG surgery.

In the current study, several of the risk factors showed a temporal pattern similar to that reported by Shahian et al. (7). For example, the magnitude of the effects of a reduced EF, severe renal impairment, preoperative dialysis, respiratory disease, and arrhythmia on mortality risk decreased over time. A possible explanation of such trends might be that these predictors are linked to a patient's recovery from surgery during the early postoperative period, and if a patient survives that early postoperative period then the effect of these risk factors on survival diminishes. The opposite trend was seen in some of the other risk factors. The magnitude of the effects of smoking, diabetes, hypertension and congestive heart failure on mortality risk all increased over time, suggesting an accumulation of risk from these debilitating chronic behaviors and diseases (7). The risk with high cholesterol was seen to be progressively falling. Possible explanation for such paradox, may be the use of statin. Published evidence also demonstrated similar evidence that perioperative statin therapy improves outcomes in patients undergoing coronary artery bypass grafting (28). Further research is needed to explore the precise dynamics of the time-varying effects of the risk factors across time.

In general, the prediction models for all four time intervals performed well. However, models for later time intervals showed lower discrimination compared to those for earlier time intervals. This is likely because, as more time passes since the surgery, the relative influence of factors unrelated to surgery increases and thus compromises the discriminatory power of the model.

This is the first study to predict long-term survival after isolated CABG surgery in an Australian patient cohort. One of the major strengths of this study is the use of data from a nationwide cardiac surgery registry and the NDI. Moreover, the use of a bootstrap model selection technique (29), multiple imputation of missing values, and model adjustment for hospital-level variation are major strengths of the model development process used in this study.

Limitations

The long-term survival model presented in this paper was developed based on preoperative patient characteristics. Intraoperative predictors like use of cardiopulmonary bypass use might improve the prediction. However, as intraoperative data are not available prior to surgery, surgeons can only rely on preoperative patient characteristics to foresee long-term prognosis for patient counselling and surgical decision making.

The present study used data from patients who underwent isolated CABG surgery during 2001-2014, and this includes many patients who were operated on a decade ago. Advancements in technology, surgical procedures and postsurgical care may have decreased mortality risks over time, and prolonged survival times among newer patients who have undergone CABG surgery more recently. However, the current study used the latest available ANZSCTS registry and NDI data.

The authors also acknowledge that there is inherent scope for bias due to voluntary data collection and the fact that some risk factors (eg, BMI) may change over time (but only baseline values were used in developing our prediction model). As the mortality data are collected through linkage with the NDI the cause of death was not available and, therefore, there is the potential for an overestimation of cardiac-specific mortality risks due to contamination by all-cause mortality in long-term outcome analyses.

Conclusion

Prediction models were developed in an Australian cohort for predicting mortality risk at 31-90 days, 91-365 days, 1-3 years and >3 years after isolated CABG surgery. These risk prediction models can be used by clinicians in continuum with AusSCORE II 30-day mortality risk model to get complete prognosis and thus facilitate evidenced-based surgical decision making.

**Funding statement and conflict of interest.**

The Australian and New Zealand Society of Cardiac and Thoracic Surgeons (ANZSCTS) National Cardiac Surgery Database Program is funded by the Department of Health, Victoria, the Health Administration Corporation (GMCT) and the Clinical Excellence Commission (CEC), NSW, and funding from individual Units. ANZSCTS Research activities are supported through a National Health and Medical Research Council (NHMRC) Senior Research Fellowship and Program Grant awarded to C.M. Reid. M.N. Karim has an Australian Postgraduate Awards Scholarship from Monash University. S.L. Brilleman is funded by an Australian NHMRC Postgraduate Scholarship (APP1093145). The funding bodies had no role in the design and conduct of the study, the collection, management, analysis and interpretation of the data, or the preparation, review, or approval of the manuscript.

Conflict of interest: None declared

## References


1. Osswald BR, Blackstone EH, Tochtermann U, Thomas G, Vahl CF, Hagl S. The meaning of early mortality after CABG. Eur J Cardiothorac Surg. 1999;15(4):401-7.

2. Billah B, Reid CM, Shardey GC, Smith JA. A preoperative risk prediction model for 30-day mortality following cardiac surgery in an Australian cohort. Eur J Cardiothorac Surg. 2010;37(5):1086-92.

3. Nashef SA, Roques F, Sharples LD, Nilsson J, Smith C, Goldstone AR, et al. Euroscore II. Eur J Cardiothorac Surg. 2012;41:12.

4. D'Errigo P, Seccareccia F, Rosato S, Manno V, Badoni G, Fusco D, et al. Comparison between an empirically derived model and the EuroSCORE system in the evaluation of hospital performance: the example of the Italian CABG Outcome Project. Eur J Cardiothorac Surg. 2008;33(3):325-33.

5. Carr BM, Romeiser J, Ruan J, Gupta S, Seifert FC, Zhu W, et al. Long-Term Post-CABG Survival: Performance of Clinical Risk Models Versus Actuarial Predictions. J Card Surg. 2016;31(1):23-30.

6. Baillot RG, Joanisse DR, Stevens LM, Doyle DP, Dionne B, Lellouche F. Recent evolution in demographic and clinical characteristics and in-hospital morbidity in patients undergoing coronary surgery. Can J Surg. 2009;52(5):394-400.

7. Shahian DM, O'Brien SM, Sheng S, Grover FL, Mayer JE, Jacobs JP, et al. Predictors of Long-Term Survival Following Coronary Artery Bypass Grafting Surgery: Results from The Society of Thoracic Surgeons Adult Cardiac Surgery Database (The ASCERT Study). Circulation. 2012;125(12):1491-500.

8. O'Boyle F, Mediratta N, Fabri B, Pullan M, Chalmers J, McShane J, et al. Long-term survival after coronary artery bypass surgery stratified by EuroSCORE†. Eur J Cardiothorac Surg. 2012;42(1):101-7.

9. Wu C, Camacho FT, Wechsler AS, Lahey S, Culliford AT, Jordan D, et al. Risk Score for Predicting Long-Term Mortality After Coronary Artery Bypass Graft Surgery. Circulation. 2012;125(20):2423-30.

10. Shahian DM, O'Brien SM, Filardo G, Ferraris VA, Haan CK, Rich JB, et al. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 1—coronary artery bypass grafting surgery. Ann Thorac Surg . 2009;88(1):S2-S22.

11. Kolh P. Importance of risk stratification models in cardiac surgery. Eur Heart J. 2006;27(7):768-9.

12. Yap CH, Mohajeri M, Ihle BU, Wilson AC, Goyal S, Yii M. Validation of EuroSCORE model in an Australian patient population. Aust N Z J Surg . 2005;75(7):5.

13. Reid C, Billah B, Dinh D, Smith J, Skillington P, Yii M, et al. An Australian risk prediction model for 30-day mortality after isolated coronary artery bypass: the AusSCORE. J Thorac Cardiovasc Surg. 2009;138(4):904-10.

14. Rubin DB. Multiple Imputation for Nonresponse in Surveys. Hoboken (NJ): John Wiley & Sons; 2004.

15. Billah B, Huq MM, Smith JA, Sufi F, Tran L, Shardey GC, et al. AusSCORE II in predicting 30-day mortality after isolated coronary artery bypass grafting in Australia and New Zealand. J Thorac Cardiovasc Surg. 2014;148(5):1850-5. e2.

16. Austin PC, Tu JV. Bootstrap Methods for Developing Predictive Models. Am Stat. 2004;58(2):131-7.

17. Wood A, White I, Royston P. How should variable selection be performed with multiply imputed data? . Stat Med. 2008;27:20.

18. Sauerbrei W, Schumacher M. A bootstrap resampling procedure for model building: application to the Cox regression model. Stat Med. 1992;11(16):2093-109.

19. Royston P, Ambler G, Sauerbrei W. The use of fractional polynomials to model continuous risk variables in epidemiology. Int J Epidemiol . 1999;28(5):964-74.

20. Harrell FE, Jr, Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. JAMA. 1982;247(18):2543-6.

21. Harrell FEJ, Lee KL, Califf RM, Pryor DB, Rosati RA. Regression modelling strategies for improved prognostic prediction. Stat Med. 1984;3(2):10.

22. Gutierrez RG. Parametric frailty and shared frailty survival models. Stata J. 2002;2(1):23.

23. Royston P, Altman DG. External validation of a Cox prognostic model: principles and methods. BMC Med Res Methodol. 2013;13(1):33.

24. Gao D, Grunwald GK, Rumsfeld JS, Schooley L, MacKenzie T, Shroyer ALW. Time-Varying Risk Factors for Long-Term Mortality After Coronary Artery Bypass Graft Surgery. Ann Thorac Surg.81(3):793-9.

25. Gardner SC, Grunwald GK, Rumsfeld JS, Mackenzie T, Gao D, Perlin JB, et al. Risk factors for intermediate-term survival after coronary artery bypass grafting. Ann Thorac Surg.72(6):2033-7.

26. Herlitz J, Haglid M, Albertsson P, Westberg S, Karlson B, Hartford M, et al. Short- and Long-Term Prognosis after Coronary Artery Bypass Grafting in Relation to Smoking Habits. Cardiology. 1997;88(6):492-7.

27. Saxena A, Shan L, Reid C, Dinh DT, Smith JA, Shardey GC, et al. Impact of smoking status on early and late outcomes after isolated coronary artery bypass graft surgery. J Cardiol. 2013;61(5):336-41.

28. Barakat AF, Saad M, Abuzaid A, Mentias A, Mahmoud A, Elgendy IY. Perioperative Statin Therapy for Patients Undergoing Coronary Artery Bypass Grafting. Ann Thorac Surg..101(2):818-25.

29. Karim MN, Reid CM, Tran L, Cochrane A, Billah B. Missing Value Imputation Improves Mortality Risk Prediction Following Cardiac Surgery: An Investigation of an Australian Patient Cohort. Heart Lung Circ. 2017;26(3):301-8.

## 7.3 Summary

This is the first study to predict long-term survival after isolated CABG surgery in an Australian patient cohort. A set of models have been developed for predicting long-term survival at four distinct time intervals (31-90 days, 91-365 days, 1-3 years, >3 years), recognising the fact that the effect of various risk factors on mortality may differ depending on the time since CABG surgery. One of the major strengths of this study is the use of data from a nationwide cardiac surgery registry and the national death index. Moreover, the use of a bootstrap model selection technique, multiple imputation of missing values, and model adjustment for hospital-level variation are major strengths of the model development process used in this study. These risk prediction models can be used by clinicians in continuum with existing short-term mortality risk prediction model (AusSCORE II) to get complete prognosis of patients undergoing CABG surgery and thus facilitate evidenced-based surgical decision making.

# Chapter 8: General discussion and conclusion

## 8.1 Introduction

The overarching aim of this thesis was to study the aspects of risk prediction modelling. First to understand the current practice and knowledge gaps, a review of existing models predicting coronary artery bypass surgery was undertaken. The review focused on critical appraisal of the methods that were used for developing these models.

The information of current practice and the knowledge gaps, revealed by the review, were used to guide the aim and structure of this thesis. Chapter four, five and six investigated the issues identified in the review. Chapter seven aimed to apply all the issues revealed in the review and the knowledge gained in the subsequent chapters (chapters four, five and six) in risk prediction model development.

This chapter includes a summary of strengths and limitations of the thesis and the key findings. Based on key findings, future direction is proposed to improve the risk prediction in the field of cardiac surgery.

## 8.2 Strength and limitation of the thesis

The thesis has four major strengths, predominantly in the novelty of several projects, backed by appropriate methods and reliable data. This thesis identified knowledge gaps in the field of cardiac surgery risk prediction modelling and proceeded to advance that knowledge. First of the major strength of the thesis was, the systematic approach to assessing the current practices prevailing in the cardiac surgery risk prediction modelling. Secondly, use of nationwide large multi-centred cardiac surgery registry data for the research ensures the generalizability and adequacy of models. Thirdly, this was the first research to investigate impacts of different methodological characteristics (standardized variable definition, missing value imputation and variable selection method) on model performances. Fourthly, the model developed for prediction of long-term survival following CABG was the first in Australian population and is among the few in the world. The model development process employed most updated dataset and robust methods to suit contemporary clinical practice and to optimize performance.

One limitation of the thesis was, it focused on mortality as outcome. There is argument that mortality by itself is not a sufficient indicator of surgical performance and emphasis on mortality as the only endpoint, may reduce access to surgery for people with high risk of mortality. Morbidity could have been the suitable alternative. However, unlike mortality, morbidity data are difficult to collect, and there is problem with standardization of morbidity definitions. Another limitation of the thesis was it did not consider the intraoperative or postoperative variable. Inclusion of intraoperative predictors might improve the prediction. However, as intraoperative data are not available prior to surgery. The present study used data from patients who underwent isolated CABG surgery during 2001-2014, and this includes many patients who were operated on a decade ago. Advancements in technology, surgical procedures and postsurgical care may have decreased mortality risks over time, and prolonged survival times among newer patients who have undergone the CABG operation.

## 8.3 Key findings and their implications

The approach taken to pursue the aim of the research was to identify knowledge gap and research need in cardiac surgery risk prediction modelling, to address those gaps and to incorporate the knowledge gained in model development. The key finding in relation to the aims are as follows:

### 8.3.1 Current scenario of risk prediction models for coronary artery bypass graft surgery

A total of 53 risk-prediction models for short-term mortality following CABG were identified. Many of these models didn't vividly detail their development methodology and validation process. Wide variation exists in the development methodology of the risk prediction models. Only few of them employed most appropriate statistical methods required to optimize prediction. Ambiguous predictor and outcome definition, sub-optimum sample size, inappropriate handling of missing data and inefficient predictor selection technique were major issues identified in the review.

Findings of the review were used to guide the structure of this thesis. In subsequent chapters, issues identified in the review were investigated. Subsequently the knowledge gained from the research are applied to develop a novel long-term survival model for CABG patients.

### 8.3.2 Misclassification of procedural status and implications for risk assessment

In the ANZSCTS database, prevalence of procedural status misclassification was quite high (14.4%). Misclassification of patient to urgent category was prevalent more among patients with certain preoperative conditions (cardiogenic shock, preoperative dialysis, endocarditis, and BMI<18.5). Higher prevalence of misclassification in several high-risk groups of patients with a high mortality contradicts the existing time-dependent classification of urgent status in ANZSCTS registry database. Misclassification compromises the discrimination capacity and calibration of the model and results in overestimation of mortality risk.

This study proposes a new definition of 'urgent' status to include the following categories – (a) Cardiac surgery within 72 hours from angiography, if on the same admission; (b) Cardiac surgery within 72 hours of an unplanned admission; (c) Cardiac surgery for acute valve endocarditis; (d) Cardiac surgery for patients admitted to hospital with cardiogenic shock, or patients with worsening or ongoing chest pain; (e) Cardiac surgery for patients with ejection fraction less than 30% and who have been admitted to hospital before surgery; (f) Surgery for patients on pre-operative dialysis who are admitted to hospital; and (g) Surgery for underweight patients, defined as BMI < 18.5.

### 8.3.3 Impact of missing values on the prediction performance of the model

In the ANZSCTS database, one or more missing predictor variables were present in 15.8% of the patients. Conventional complete case analysis approach of handling missing data during model development results in bias in prediction estimate. Patients with higher risk of mortality are expected to incur more bias in prediction. Multiple imputation of missing values during model development increases the precision and performance of the risk prediction models.

Risk prediction modelling should endeavor to treat missing values with an appropriate technique to maximize the prediction performance of the model.

### 8.3.4 Variable selection methods and the parsimony of risk prediction models

This study compared the parsimony and performance of models generated using five commonly used variable selection methods. As a variable selection technique, bootstrap bagging in conjunction with ROC outperformed popularly used stepwise logistic regression and Bayesian algorithm. Clinical

suitability in terms of parsimony and prediction performance can best be achieved using this technique for the development of risk prediction models.

The study recommends the use of bootstrap bagging technique in conjunction with ROC for risk prediction model development in cardiac surgery patients.

### 8.3.5 Predicting long-term survival after coronary artery bypass graft surgery

The long-term outcomes of surgery are less affected by conventional cardiac-related predictors of early mortality, such as NYHA class, previous myocardial infraction, Inotrope use, cardiogenic shock, urgency of operation. Rather, it is mostly related with comorbidities and chronic conditions such as respiratory disease, smoking history, diabetes status, hypertension, hypercholesterolemia, cerebrovascular disease. This finding underpins the importance of risk factors, functional status and non-cardiac comorbidities for predicting longer-term survival following CABG. Several of the predictors showed a temporal pattern of mortality risk. Magnitude of effect of these factors including smoking, diabetes, hypertension, hypercholesterolemia and congestive heart failure increased over time following surgery. Further research in this issue is needed to explore the precise dynamics of the changing risk across time after surgery.

Four separate models were developed for predicting survival after isolated CABG surgery at 31-90 days, 91-365 days, 1-3 years and >3 years. These risk models can be used by clinician in continuum with 30-day mortality risk prediction model (AusScore II) to get complete prognosis thus to facilitate evidenced based surgical decision making.

#### 8.3.6 Implication of the finding

The potential for the application of Risk prediction models in medical science is vast and is not any more restricted to surgical outcome assessment only. Nowadays they are used for wide range of purposes ranging from administrative to fiscal issues. Risk modelling requires endeavoring to improve the prediction capacity, no matter how subtle the improvements are. The modeling methodology should be correct and robust and the proposed model must be straightforward to implement and clinically relevant. The finding of the current research investigated aspects of model development methods where there are room for improvement. Recommendation based on the finding of current research are likely to improve performance of these models.

Another major implication of the current study finding is that it necessitates many of the currently used models to undergo methodological refurbishment and upgradation to cope with the current clinical practice and to incorporate divers aim.

Further, a major implication lies in the development of long term survival models, which helps overcoming of shortcomings of short-term mortality as outcome indicator. Prediction of long-term survival can aid determination of the most appropriate post-discharge care strategies. This would essentially help patients and their doctors to implement behavioural and therapeutic modifications to optimize benefit from surgery. By using both short and long-term models in tandem clinician can get complete prognosis thus to facilitate evidenced based surgical decision making.

## 8.4 Future direction

At a practical level, it would be most useful to construct a calculator, so that relevant health professional could electronically calculate individual patients' mortality risk for each of the four time periods. Future effort should include conversion of these models into simple calculator, which can be either used at bedside or on the web or can even be both. The benefit of preoperative risk prediction can be maximized by incorporating risk prediction process into guidelines to stratify patients per risk level and identify patients who may benefit most from a specific treatment strategy.

Current research focused primarily on developing model for predicting long-term survival following CABG surgery. Further research is needed to develop models for other procedures (eg. valve surgery). As procedure-specific models are preferred over all procedure models in cardiac surgery (83), Further research in this issue is also needed to resolve, how specific a model must be to attain the best possible prediction.

Current thesis used mortality, conventionally and popularly used in the field, as outcome indicator of surgical performance. With the growing emphasis for cost-effectiveness of care and quality of life, models are also expected to predict cost of care, hospital resource need and post-operative adverse event (39). Postoperative morbidity and length of hospital stay are important determinant of cost of care and quality of life after surgery. Future research should be directed towards prediction of morbidity following surgery and resource need.

## 8.5 Concluding remarks

The aim of the thesis was to study the aspects of development of risk prediction models with a view to improve cardiac surgical outcome assessment. This research has provided new knowledge about the existing practice in the risk prediction modelling for cardiac surgery patients. Currently there was no consensus on model development method for generating parsimonious model. Current research provided a range of evidence based suggestion regarding model development practices for improving outcome assessment following cardiac surgery.

The research also provided a set of novel risk prediction models for predicting long-term survival at four distinct time intervals following CABG surgery. These prediction models are generated in continuum with the existing 30-day mortality risk prediction model developed on the same datasets. These models along with the existing short-term mortality model will provide surgeons and patients greater confidence in surgical decision making.

# Reference


1. Bridgewater B, Neve H, Moat N, Hooper T, Jones M. Predicting operative risk for coronary artery surgery in the United Kingdom: a comparison of various risk prediction algorithms. Heart. 1998;79(4):350-5.

2. Head SJ, Bogers AJ, Serruys PW, Takkenberg JJ, Kappetein AP. A crucial factor in shared decision making: the team approach. The Lancet. 2011;377(9780):1836.

3. Shahian DM, Blackstone EH, Edwards FH, Grover FL, Grunkemeier GL, Naftel DC, et al. Cardiac surgery risk models: a position article. The Annals of Thoracic Surgery. 2004;78(5):1868-77.

4. Asimakopoulos G, Al-Ruzzeh S, Ambler G, Omar R, Punjabi P, Amrani M, et al. An evaluation of existing risk stratification models as a tool for comparison of surgical performances for coronary artery bypass grafting between institutions. European Journal of Cardio-thoracic surgery. 2003;23(6):935-42.

5. Shahian DM, O'Brien SM, Filardo G, Ferraris VA, Haan CK, Rich JB, et al. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 1—coronary artery bypass grafting surgery. The Annals of thoracic surgery. 2009;88(1):S2-S22.

6. Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R, et al. European system for cardiac operative risk evaluation (EuroSCORE). European journal of cardio-thoracic surgery. 1999;16(1):9-13.

7. Kappetein AP, Head SJ. Predicting prognosis in cardiac surgery: a prophecy? European Journal of Cardio-Thoracic Surgery. 2012;41:2.

8. Kolh P. Importance of risk stratification models in cardiac surgery. European heart journal. 2006;27(7):768-9.

9. Granton J, Cheng D, editors. Risk stratification models for cardiac surgery. Seminars in cardiothoracic and vascular anesthesia; 2008: SAGE Publications.

10. Grover FL, Hammermeister KE, Shroyer ALW. Quality initiatives and the power of the database: What they are and how they run. The Annals of Thoracic Surgery. 1995;60(5):1514-21.

11. Heijmans J, Maessen J, Roekaerts P. Risk stratification for adverse outcome in cardiac surgery. European journal of anaesthesiology. 2003;20(07):515-27.

12. Nilsson J, Algotsson L, Höglund P, Lührs C, Brandt J. EuroSCORE predicts intensive care unit stay and costs of open heart surgery. The Annals of thoracic surgery. 2004;78(5):1528-34.

13. Antunes PE, Eugénio L, de Oliveira JF, Antunes MJ. Mortality risk prediction in coronary surgery: a locally developed model outperforms external risk models. Interactive cardiovascular and thoracic surgery. 2007;6(4):437-41.

14. Kennedy J, Kaiser G, Fisher L, Maynard C, Fritz J, Myers W, et al. Multivariate discriminant analysis of the clinical and angiographic predictors of operative mortality from the Collaborative Study in Coronary Artery Surgery (CASS). The Journal of thoracic and cardiovascular surgery. 1980;80(6):876-87.

15. Kouchoukos N, Oberman A, Kirklin J, Russell Jr R, Karp R, Pacifico A, et al. Coronary bypass surgery: analysis of factors affecting hospital mortality. Circulation. 1980;62(2 Pt 2):I84-9.

16. Chaitman BR, Rogers WJ, Davis K, Tyras DH, Berger R, Bourassa MG, et al. Operative risk factors in patients with left main coronary-artery disease. New England Journal of Medicine. 1980;303(17):953-7.

17. Paiement B, Maillé J-G, Boulanger M, Taillefer J, Sahab P, Pelletier C, et al. La visite pre-operatoire en chirurgie cardio-vasculaire. Canadian Anaesthetists' Society Journal. 1980;27(6):584-93.

18. Paiement B, Pelletier C, Dyrda I, Maille JG, Boulanger M, Taillefer J, et al. A simple classification of the risk in cardiac surgery. Can Anaesth Soc J. 1983;30(1):61-8.

19. Parsonnet V, Dean D, Bernstein A. A method of uniform stratification of risk for evaluating the results of surgery in acquired adult heart disease. Circulation. 1989;79(6 Pt 2):I3-12.

20. Hattler BG, Madia C, Johnson C, Armitage JM, Hardesty RL, Kormos RL, et al. Risk stratification using the Society of Thoracic Surgeons program. The Annals of thoracic surgery. 1994;58(5):1348-52.

21.Keogh BE, Kinsman R. National adult cardiac surgical database report. Playhatch, Berkshire, UK: : The Society of Cardiothoracic Surgeons of Great Britain and Ireland. Dendrite Clinical Systems Ltd 1998.

22. Society of Cardiothoracic Surgeons of Great Britain and Ireland National Database : Preliminary report 1996.: Society of Cardiothoracic Surgeons of Great Britain, Ireland National Database and MRC Biostatistics Unit Cambridge and Dendrite Clinical systems (UK) Ltd.; 1996.

23. Nashef SA, Roques F, Sharples LD, Nilsson J, Smith C, Goldstone AR, et al. Euroscore II. European journal of cardio-thoracic surgery. 2012;41:12.

24. Yap CH, Mohajeri M, Ihle BU, Wilson AC, Goyal S, Yii M. Validation of Euroscore model in an Australian patient population. ANZ Journal of Surgery. 2005;75(7):508-12.

25. Reid C, Billah B, Dinh D, Smith J, Skillington P, Yii M, et al. An Australian risk prediction model for 30-day mortality after isolated coronary artery bypass: the AusSCORE. The Journal of thoracic and cardiovascular surgery. 2009;138(4):904-10.

26. Billah B, Huq MM, Smith JA, Sufi F, Tran L, Shardey GC, et al. AusSCORE II in predicting 30-day mortality after isolated coronary artery bypass grafting in Australia and New Zealand. The Journal of thoracic and cardiovascular surgery. 2014;148(5):1850-5. e2.

27. Billah B, Reid CM, Shardey GC, Smith JA. A preoperative risk prediction model for 30-day mortality following cardiac surgery in an Australian cohort. European Journal of Cardio-Thoracic Surgery. 2010;37(5):1086-92.

28. Ariyaratne TV, Billah B, Yap C-H, Dinh D, Smith JA, Shardey GC, et al. An Australian risk prediction model for determining early mortality following aortic valve replacement. European Journal of Cardio-Thoracic Surgery. 2011;39(6):815-21.

29. Prins C, Jonker IDV, Smit FE, Botes L. Cardiac surgery risk-stratification models. Cardiovascular journal of Africa. 2012;23(3):160.

30. Christenson JT, Simonet F, Schmuziger M. Extensive endarterectomy of the left anterior descending coronary artery combined with coronary artery bypass grafting. Coronary Artery Disease. 1995;6(9):731-7.

31. Siregar S, Groenwold RHH, Versteegh MIM, Noyez L, ter Burg WJPP, Bots ML, et al. Gaming in risk-adjusted mortality rates: effect of misclassification of risk factors in the benchmarking of cardiac surgery risk-adjusted mortality rates. The Journal of thoracic and cardiovascular surgery. 2013;145(3):781-9.

32. Andrianopoulos N, Dinh D, Shardey G, Smith J, Cameron P, Reid C. Application of variable life adjusted displays (VLAD) on isolated coronary artery bypass from Australian society of cardiac and thoracic surgeons (ASCTS) database. Circulation: Cardiovascular Quality and Outcomes Conference: Quality of Care and Outcomes Research in Cardiovascular Disease and Stroke. 2010;4(6 MeetingAbstracts2010).

33. Gameren MV, Kliverik LMA, Struijs A, Venema AC, Kappetein AP, Bogers AJJC, et al. Impact of the definition of renal dysfunction on EuroSCORE performance. Journal of Cardiovascular Surgery. 2009;50(5):7.

34. Chassin MR, Hannan EL, DeBuono BA. Benefits and Hazards of Reporting Medical Outcomes Publicly. New England Journal of Medicine. 1996;334(6):394-8.

35. Omar RZ, Ambler G, Royston P, Eliahoo J, Taylor KM. Cardiac surgery risk modeling for mortality: a review of current practice and suggestions for improvement. The Annals of thoracic surgery. 2004;77(6):2232-7.

36. Swinkels B, Plokker H. Evaluating operative mortality of cardiac surgery: first define operative mortality. Netherlands Heart Journal. 2010;18(7):344-5.

37. Osswald BR, Blackstone EH, Tochtermann U, Thomas G, Vahl CF, Hagl S. The meaning of early mortality after CABG. European Journal of Cardio-Thoracic Surgery. 1999;15(4):401-7.

38. Shahian DM, Edwards FH, Ferraris VA, Haan CK, Rich JB, Normand S-LT, et al. Quality Measurement in Adult Cardiac Surgery: Part 1—Conceptual Framework and Measure Selection. The Annals of Thoracic Surgery. 2007;83(4, Supplement):S3-S12.

39. Shahian DM, O'Brien SM, Sheng S, Grover FL, Mayer JE, Jacobs JP, et al. Predictors of Long-Term Survival Following Coronary Artery Bypass Grafting Surgery: Results from The Society of Thoracic Surgeons Adult Cardiac Surgery Database (The ASCERT Study). Circulation. 2012;125(12):1491-500.

40. Baillot RG, Joanisse DR, Stevens LM, Doyle DP, Dionne B, Lellouche F. Recent evolution in demographic and clinical characteristics and in-hospital morbidity in patients undergoing coronary surgery. Canadian Journal of Surgery. 2009;52(5):394-400.

41. Parsonnet V. Risk stratification in cardiac surgery: is it worthwhile? Journal of cardiac surgery. 1995;10(6):690-8.

42. Jones R. In search of the optimal surgical mortality. Circulation. 1989;79(6 Pt 2):I132-6.

43. Shahian DM, Normand S-L, Torchiana DF, Lewis SM, Pastore JO, Kuntz RE, et al. Cardiac surgery report cards: comprehensive review and statistical critique. The Annals of thoracic surgery. 2001;72(6):2155-68.

44. Thakar CV, Arrigain S, Worley S, Yared J-P, Paganini EP. A Clinical Score to Predict Acute Renal Failure after Cardiac Surgery. Journal of the American Society of Nephrology. 2005;16(1):162-8.

45. Brown JR, Cochran RP, Leavitt BJ, Dacey LJ, Ross CS, MacKenzie TA, et al. Multivariable Prediction of Renal Insufficiency Developing After Cardiac Surgery. Circulation. 2007;116(11 suppl):I-139-I-43.

46. Badheka AO, Chothani A, Mehta K, Patel NJ, Deshmukh A, Hoosien M, et al. Utilization and adverse outcomes of percutaneous left atrial appendage closure for stroke prevention in atrial fibrillation in the United States: influence of hospital volume. Circulation: Arrhythmia and Electrophysiology. 2015;8(1):42-8.

47. Roques F, Gabrielle F, Michel P, De Vincentiis C, David M, Baudet E. Quality of care in adult heart surgery: proposal for a self-assessment approach based on a French multicenter study. European journal of cardio-thoracic surgery. 1995;9(8):433-40.

48. Chou S, Deily ME, Li S. Travel distance and health outcomes for scheduled surgery. Medical Care. 2014;52(3):250-7.

49. Singh M, Gersh BJ, Li S, Rumsfeld JS, Spertus JA, O'Brien SM, et al. Mayo Clinic risk score for percutaneous coronary intervention predicts in-hospital mortality in patients undergoing coronary artery bypass graft surgery. Circulation. 2008;117(3):356-62.

50. van Gameren M, Kappetein AP, Steyerberg EW, Venema AC, Berenschot EA, Hannan EL, et al. Do we need separate risk stratification models for hospital mortality after heart valve surgery? The Annals of thoracic surgery. 2008;85(3):921-30.

51. Little RJA, Rubin DB. Statistical Analysis with Missing Data. 2nd Edition ed. New York: Willey; 2002.
52. Greenland S, Finkle WD. A Critical Look at Methods for Handling Missing Covariates in Epidemiologic Regression Analyses. American Journal of Epidemiology. 1995;142(12):1255-64.
53. Moons KGM, Donders RART, Stijnen T, Harrell Jr FE. Using the outcome for imputation of missing predictor values was preferred. Journal of Clinical Epidemiology. 2006;59(10):1092-101.
54. Yu LM, Burton A, Rivero-Arias O. Evaluation of software for multiple imputation of semi-continuous data. Statistical Methods in Medical Research. 2007;16(3):243-58.
55. Janssen KJM, Vergouwe Y, Donders ART, Harrell FE, Chen Q, Grobbee DE, et al. Dealing with Missing Predictor Values When Applying Clinical Prediction Models. Clinical Chemistry. 2009;55(5):994-1001.
56. Greenland S. Principles of multilevel modelling. International Journal of Epidemiology. 2000;29(1):158-67.
57. Amorim LDAF, Fiaccone RL, Santos CAST, Santos TNd, Moraes LTLPd, Oliveira NF, et al. Structural equation modeling in epidemiology. Cadernos de Saúde Pública. 2010;26:2251-62.
58. Kuppermann N, Willits N. In response to "Statistical Models and Occam's Razor". Acad Emerg Med. 2000 7(1):4.
59. Wears RL, Lewis RJ. Statistical Models and Occam's Razor. Academic Emergency Medicine. 1999;6(2):93-4.
60. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. European Heart Journal. 2014;35(29):1925-31.
61. Harrell FEJ, Lee KL, Califf RM, Pryor DB, Rosati RA. Regression modelling strategies for improved prognostic prediction. Statistic in Medicine. 1984;3(2):10.
62. Congdon P. Bayesian Statistical Modelling. 2nd ed: John Wiley and Sons; 2006.
63. Breiman L. Statistical Modeling: The Two Cultures (with comments and a rejoinder by the author). 2001:199-231.
64. Steen PM. Approaches to predictive modeling. The Annals of Thoracic Surgery. 1994;58(6):1836-40.
65. Hammermeister K, Johnson R, Marshall G, Grover FL. Continuous assessment and improvement in quality of care. A model from the Department of Veterans Affairs Cardiac Surgery. Annals of surgery. 1994;219(3):281.
66. Ivanov J, Tu JV, Naylor CD. Ready-made, recalibrated, or Remodeled? Issues in the use of risk indexes for assessing mortality after coronary artery bypass graft surgery. Circulation. 1999;99(16):2098-104.
67. Azarasa M, Azarfarin R, Changizi A, Alizadehasl A. Substance use among Iranian cardiac surgery patients and its effects on short-term outcome. Anesthesia & Analgesia. 2009;109(5):1553-9.

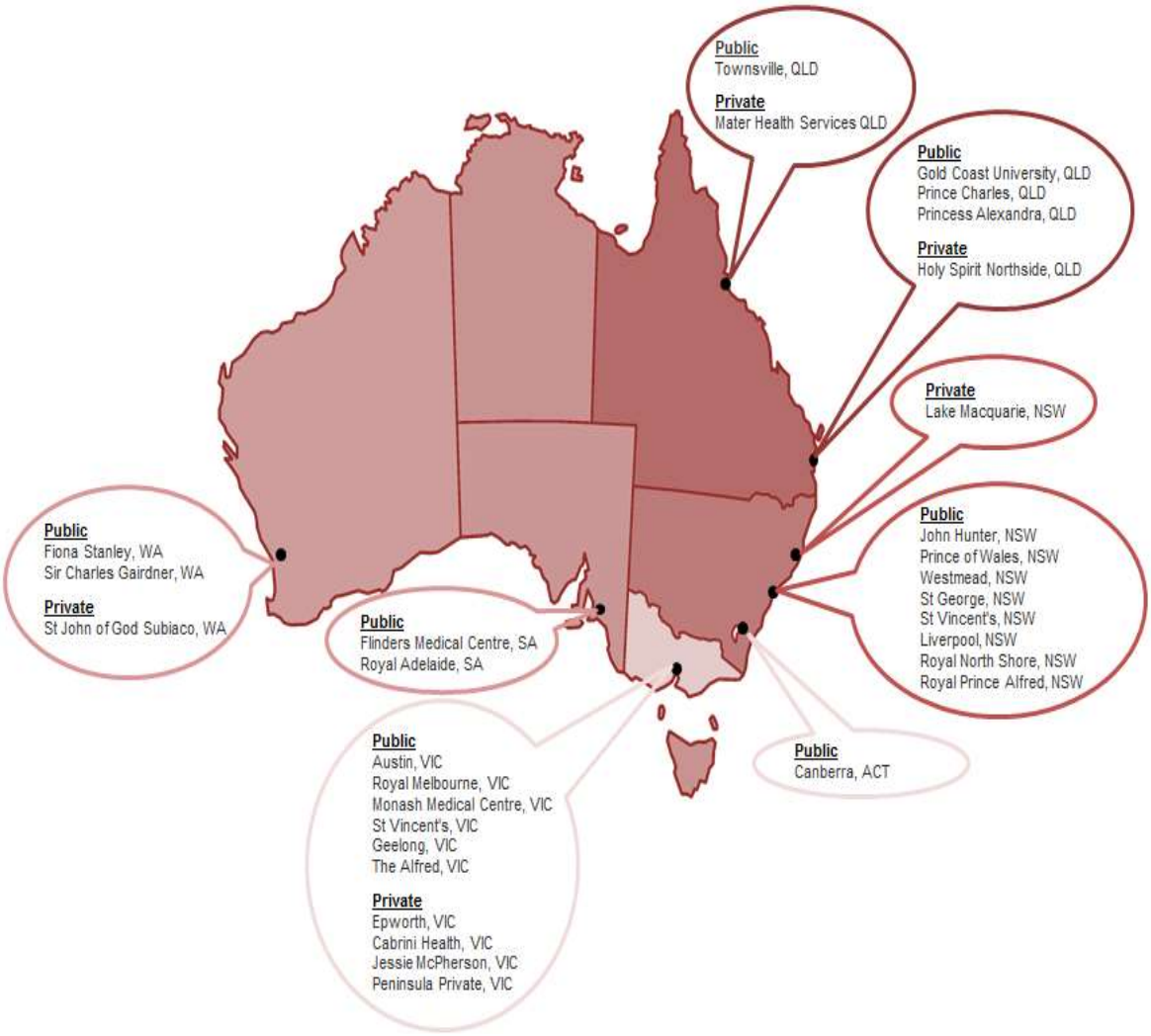
68. Nilsson J, Algotsson L, Höglund P, Lührs C, Brandt J. Comparison of 19 pre-operative risk stratification models in open-heart surgery. European Heart Journal. 2006;27(7):867-74.
69. Reid CM, Rockell M, Skillington PD, Shardey GC, Smith JA, M.S MY, et al. Initial Twelve Months Experience and Analysis for 2001–2002 from the Australasian Society of Cardiac and Thoracic Surgeons—Victorian Database Project. Heart, Lung and Circulation. 2004;13(3):291-7.
70. Skillington P. A national cardiac surgery database: Current achievements and future challenges. Heart, Lung and Circulation. 2001;10(1):S2-S4.
71. ANZSCTS. National Cardiac Surgery Database Program. Data Definitions Manual. 2008.
72. Moons KGM, de Groot JAH, Bouwmeester W, et al. Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies: The CHARMS Checklist. PLOS Medicine. 2014;11(10):e1001744.
73. Steyerberg EW. Clinical prediction models: a practical approach to development, validation, and updating. New York: Springer; 2009.
74. Rubin DB. Multiple Imputation for Nonresponse in Surveys. Hoboken (NJ): John Wiley & Sons; 2004.
75. Wood AM, White IR, Royston P. How should variable selection be performed with multiply imputed data? Statistics in Medicine. 2008;27(17):3227-46.
76. Sauerbrei W, Schumacher M. A bootstrap resampling procedure for model building: application to the Cox regression model. Statistics in medicine. 1992;11(16):2093-109.
77. Austin PC, Tu JV. Bootstrap Methods for Developing Predictive Models. The American Statistician. 2004;58(2):131-7.
78. Royston P, Ambler G, Sauerbrei W. The use of fractional polynomials to model continuous risk variables in epidemiology. International journal of epidemiology. 1999;28(5):964-74.
79. Gutierrez RG. Parametric frailty and shared frailty survival models. The Stata Journal. 2002;2(1):23.
80. Sullivan LM, Massaro JM, D'Agostino RB. Presentation of multivariate data for clinical use: The Framingham Study risk score functions. Statistics in Medicine. 2004;23(10):1631-60.
81. Simon R, Altman DG. Statistical aspects of prognostic factor studies in oncology. British Journal of Cancer. 1994;69(6):979-85.
82. Carr BM, Romeiser J, Ruan J, Gupta S, Seifert FC, Zhu W, et al. Long-Term Post-CABG Survival: Performance of Clinical Risk Models Versus Actuarial Predictions. Journal of Cardiac Surgery. 2016;31(1):23-30.
83. van Gameren M, Kappetein AP, Steyerberg EW, Venema AC, Berenschot EAJ, Hannan EL, et al. Do We Need Separate Risk Stratification Models for Hospital Mortality After Heart Valve Surgery? The Annals of Thoracic Surgery. 2008;85(3):921-30

# Appendices

**Appendix 1: Preoperative risk prediction model developed from ANZSCTS database**

| Year | Model | Procedure type | Outcome | Sample size | Data collection period | Predictor Number | Predictors |
|---|---|---|---|---|---|---|---|
| 2009 | An Australian risk prediction model for 30-day mortality after isolated coronary artery bypass: The AusSCORE | Coronary Artery Bypass Grafting | 30-day mortality | 11823 | 2001-2005 | 8 | Age, NYHA class, ejection fraction estimate, urgency of procedure, previous cardiac surgery, hypercholesterolemia, peripheral vascular disease, and cardiogenic shock |
| 2010 | A preoperative risk prediction model for 30-day mortality following cardiac surgery in an Australian cohort (Global Model). | Case mix of cardiac surgery | 30-day mortality | 23016 | 2001-2008 | 12 | Age, sex, NYHA class, urgency of procedure, ejection fraction estimate, lipid-lowering treatment, preoperative dialysis, previous cardiac surgery, procedure type, inotropic medication, peripheral vascular disease BMI. |
| 2011 | An Australian risk prediction model for determining early mortality following aortic valve replacement. (AVR score) | Aortic valve replacement | 30-day mortality | 3544 | 2001-2008 | 9 | Age, NYHA class, left main disease, infective endocarditis, cerebrovascular disease, renal dysfunction, previous cardiac surgery and estimated ejection fraction |
| 2014 | AusSCORE II in predicting 30-day mortality after isolated coronary artery bypass grafting in Australia and New Zealand | Coronary Artery Bypass Grafting | 30-day mortality | 31250 | 2001-2011 | 13 | Age, gender, ejection fraction estimate, previous cardiac surgery, urgency of procedures, eGFR, NYHA class, inotrope administration, MI, peripheral vascular disease, anticoagulant medication, cardiogenic shock, and IV nitrate administration. |

## Appendix 2.1 Participating Sites of ANZCTS Database program




*Source : ANZSCTS database program. https://anzscts.org/database*

**Appendix 2.2: ANZSCTS database data collection form**

**ANZSCTS CARDIAC SURGERY DATABASE**
**DATA COLLECTION FORM**

# ANZSCTS CARDIAC SURGERY DATABASE DATA COLLECTION FORM

1.0 Medical Record No.

First character here

## SECTION 1: PATIENT DEMOGRAPHICS

2.0 Last Name

3.0 First Name

4.0 Middle Name

5.0 Date of Birth (d d / m m / y y y y)

6.0 Sex ◯ Male ◯ Female

7.0 Address

8.0 Suburb

9.0 State

10.0 Post Code

11.0 Ph Number 1

12.0 Ph Number 2

13.0 E-mail Address

14.0 Insurance ◯ Private ◯ DVA ◯ Medicare ◯ Self-Insured ◯ Overseas ◯ Other

15.0 Medicare No. OR → 15.1 Patient does not have a Medicare No. ◯ Not Registered

16.0 Department of Veteran Affairs No.

17.0 Is patient Aboriginal or Torres Strait Islander ◯ Yes ◯ No If YES → Indicate indigenous group -*select all that apply*
◯ Aboriginal ◯ Torres Strait Islander

18.0 Elective Day of Surgery Admit? ◯ Yes ◯ No

19.0 Admission Date (d d / m m / y y y y)

20.0 Surgery Date (d d / m m / y y y y)

21.0 Discharge Date (d d / m m / y y y y)

22.0 Cardiac operation number on day for this patient (1-5) -1 = unobtainable

## SECTION 2: PATIENT RISK FACTORS

24.0 Smoking History ◯ Yes ◯ No ◯ Unknown If YES → 24.1 Current Smoker ◯ Yes ◯ No ◯ Unknown

25.0 Diabetes ◯ Yes ◯ No If YES → 25.1 Control Method ◯ None ◯ Diet ◯ Oral ◯ Insulin

26.0 Hypercholestrolaemia ◯ Yes ◯ No

**RENAL**

27.0 Last Pre-Operative Creatinine: µmol/l
(For conversion from mmol/l see overleaf)

28.0 Dialysis ◯ Yes ◯ No

29.0 Transplant ◯ Yes ◯ No

30.0 Pre-Operative Haemoglobin: g/L

31.0 Hypertension ◯ Yes ◯ No

32.0 Cerebrovascular Disease ◯ Yes ◯ No If YES → 32.1 Type ◯ Coma ◯ CVA ◯ RIND or TIA ◯ Carotid Test
If Type = CVA → 32.2 When ◯ Recent ◯ Remote

33.0 Carotid Test Result ◯ Yes ◯ No

34.0 Peripheral Vascular Disease ◯ Yes ◯ No

35.0 Respiratory Disease ◯ Yes ◯ No If YES → 35.1 Type ◯ Mild ◯ Moderate ◯ Severe

36.0 Infective Endocarditis ◯ Yes ◯ No If YES → 36.1 Type ◯ Active ◯ Treated

37.0 Immunosuppressive Therapy ◯ Yes ◯ No

# ANZSCTS CARDIAC SURGERY DATABASE

Medical Record No.

First character here

| SECTION 1: PATIENT DEMOGRAPHICS -*DEFINITIONS* | |
|---|---|
| 15.0 Medicare Number | The full Medicare number of the patient if the patient is registered with Medicare.<br>Note: The full Medicare number is comprised of the family number AND the individual reference number). |
| 16.0 DVA Number | The full DVA number of the patient if they are admitted as a DVA patient. If the patient is registered with Medicare and is admitted as a private patient enter MEDC and not DVA. |
| 18.0 DOSA Patient | Patient admitted for scheduled elective procedure on same day as procedure.<br>Note: Patients admitted to a Medihotel on the night prior to surgery still qualify as a DOSA. |
| 19.0 Admission Date | Date patient admitted/transferred to hospital where surgery performed. |
| 20.0 Surgery Date | Date on which the first surgical incision was made for the current cardiac surgical procedure. |
| 21.0 Discharge Date | Date patient discharged from being an inpatient at the hospital where the procedure was performed.<br>Discharge to hospital in the home, rehabilitation hospital or unit or to a local referring hospital is considered as discharge from hospital. |
| 22.0 Operation Number | Sequential number of cardiac operation(s) performed on the day of the index operation. Note: where a patient has two (or more) cardiac operations on the same day that warrants a new form, code '1' on the first CRF and '2' on the second CRF. |

| SECTION 2: PATIENT RISK FACTORS -*DEFINITIONS* | |
|---|---|
| 24.0 Smoking History | A history confirming any form of tobacco use in the past. |
| 24.1 Current Smoker | Smoked within one month of surgery. |
| 25.0 Diabetes | A history of diabetes, regardless of duration of disease or need for anti-diabetic agents. |
| 25.1 Diabetes Control | The most aggressive diabetes control therapy at the time of surgery (insulin>oral>diet). |
| 26.0 Hypercholesterolaemia | A history of hypercholesterolaemia diagnosed and/or treated by a physician and/or cholestrol >5.0mmol/L, HDL<1.0mmol/L or triglycerides>2.0mmol/L. |
| 27.0 Pre-Operative Creatinine | Last serum creatinine in µmol/L recorded prior to surgery. To convert from mmol/L multiply by 1000 (i.e. move decimal point 3 spaces to the right). |
| 30.0 Pre-Operative Haemoglobin | Last haemoglobin recorded prior to surgery. |
| 31.0 Hypertension | Patient has a diagnosis of hypertension, documented by **one or more** of the following:<br>a.) History of hypertenion diagnosed and treated with medication, diet, and/or exercise;<br>b.) Blood pressure exceeding 140 systolic or 90 diastolic on at least two occasions;<br>c.) Current use of antihypertensive medication |
| 32.0 Cerebrovascular Disease | Documentation by any of the following; unresponsive coma >24hrs at any time prior to the index admission OR CVA with symptoms remaining >72 hours after onset OR RIND (recovery within 72hrs) OR TIA with recovery within 24 hours OR non-invasive carotid test with 50% diameter stenosis (equivalent to 75% cross-sectional area stenosis). |
| 33.0 Cerebrovascular Disease - Carotid Test Result | Non-invasive/invasive carotid test result indicating 50% or greater diameter stenosis (equivalent to 75% cross-sectional area stenosis). |
| 34.0 Peripheral Vascular Disease | Examples include: a.) Claudication either with exertion or rest or b.) Amputation for arterial insufficiency or c.) vascular reconstruction, bypass surgery or percutaneous intervention to the extremities or d.) documented aortic aneurysm or e.) documented renal artery stenosis or f.) positive non-invasive testing documented. |
| 35.1 Respiratory Disease Type | Specify the severity of the chronic lung/respiratory disease.<br>Mild = on chronic inhaled or oral bronchodilator therapy.<br>Moderate = chronic oral steroid therapy aimed at lung disease<br>Severe = room air $pO_2$ <60 or Room air $pCO_2$ >50 or mechanical ventilation for chronic lung disease |
| 36.0 Infective Endocarditis<br>36.1 Infective Endocarditis Type | A patient presenting with valvular disease of infectious aetiology with past or present positive blood culture or postop histology or microbiological confirmation.<br>Active = currently on antibiotic therapy for endocarditis<br>Treated = no antibiotic medication (other than prophylactic medication) is being given at time of surgery |
| 37.0 Immunosuppressive Therapy | Use of any form of immunosuppressive therapy within 30 days of the operative procedure or chronic long term steroid use (eqv. to Prednisolone dosage more than or equal to 5mg within 30 days, anti-rejection medication or chemotherapy). |

# ANZSCTS CARDIAC SURGERY DATABASE

Medical Record No. ☐☐☐☐☐☐☐☐☐☐

First character here

## SECTION 3: PRE-OPERATIVE CARDIAC STATUS

38.0 Previous Myocardial Infarction ◯ Yes ◯ No **If YES →** 38.1 Type ◯ NSTEMI ◯ STEMI ◯ Unknown

**If YES →** 38.2 When ◯ <=6hrs ◯ >6 but <24hrs ◯ 1-7 days ◯ >7 to 21 days ◯ >21 days

39.0 Angina CCS Classification ☐
*Value must be between 0 - 4*

**If CCS > 0 → Treatment of Angina**

39.1 IV GTN *(day of surgery)* ◯ Yes ◯ No

39.2 IV Heparin *(<=12 hours prior to surgery)* ◯ Yes ◯ No

39.3 Full Dose Low MW Heparinoids *(<=24 hours prior to surgery)* ◯ Yes ◯ No

40.0 History of Congestive Heart Failure (CHF) ◯ Yes ◯ No **If YES →** 40.1 CHF at Current Admission ◯ Yes ◯ No

41.0 NYHA Class ☐
*Value must be between 1 - 4*

42.0 Cardiogenic Shock (at time of op) ◯ Yes ◯ No

43.0 Resuscitation (within one hour prior to op) ◯ Yes ◯ No

44.0 Arrhythmia ◯ Yes ◯ No **If YES →** Type ◯ Atrial ◯ Heart Block ◯ Ventricular ◯ Other

**If Atrial →** Type ◯ Paroxysmal ◯ Permanent ◯ Unknown

45.0 Permanent Pacemaker In Situ ◯ Yes ◯ No

**Medications at Time of Surgery**

46.0 Inotropes ◯ Yes ◯ No

47.0 IV Nitrates (GTN) ◯ Yes ◯ No

48.0 Anticoagulation Therapy *(see list below)* ◯ Yes ◯ No

49.0 Steroids ◯ Yes ◯ No

***Antiplatelet Therapy (within last 7 days)***

50.0 Aspirin Only ◯ Yes ◯ No **If YES →** 50.1 When (cessation) ☐ days

51.0 Thienopyridine *(see list below)* ◯ Yes ◯ No **If YES →** 51.1 When (cessation) ☐ days

*must be between 0 and 7 - If less than 24 hours write '0'*

52.0 Ticagrelor ◯ Yes ◯ No **If YES →** 52.1 When (cessation) ☐ days

53.0 Tirofiban or Eptifibatide ◯ Yes ◯ No **If YES →** 53.1 When (cessation) ☐ days

54.0 Abciximab ◯ Yes ◯ No **If YES →** 54.1 When (cessation) ☐ days

55.0 Other Antiplatelet ◯ Yes ◯ No **If YES →** 55.1 When (cessation) ☐ days

*Examples of Anticoagulants include (but are not limited to)*

| Brand Name | Generic Name |
|---|---|
| **HEPARIN -UNFRACTIONATED** | |
| | Heparin |
| **HEPARIN -LMW (INJECTABLE)** | |
| Fragmin | Dalteparin |
| Lovenox | Enoxaparin |
| Tinzaparin | Innohep |
| **PARENTERAL THROMBIN INHIBITORS** | |
| Angiomax | Bivalirudin |
| Argatroban | Argatroban |
| Anxtra | Fondaparinux |
| Iprivask | Desirudin |
| Refludan | Lepirudin |
| **ORAL THROMBIN INHIBITORS** | |
| Pradaxa | Dabigatran |
| Coumadin, Marevan | Warfarin |
| **FACTOR Xa INHIBITORS** | |
| Xarelto | Rivaroxaban |
| Axrixtra | Fondaparinux |
| Eliquis | Apixaban |

*Thienopyridine Agents*

| Brand Name | Generic Name |
|---|---|
| Plavix | Clopidogrel |
| Ticlid | Ticlopidine |
| Effient | Prasugrel |

## SECTION 3: PRE-OPERATIVE CARDIAC STATUS - *DEFINITIONS*

**38.0 Previous Myocardial Infarction-** Patient hospitalised at any time for a Myocardial Infarction (MI) documented in the medical record or during the current admission..

**Non ST- Elevation MI (NSTEMI)**

***AT LEAST one of the following biomarkers for detecting myocardial necrosis MUST be present (refer to note regarding Reference Control Limits)***

1.Troponin T or I: Maximal concentration of troponin T or I > the MI diagnostic limit on at least one occasion within the first 24 hours from the index clinical event;

2.CK-MB:
- Maximal value of CK-MB > 2x the upper limit of normal (ULN) on one occasion during the first hours after the index clinical event; OR
- Maximal value of CK-MB (preferable CK-MB mass) > ULN on two successive samples.

3.Total CK: Only where troponin or CK-MB assays are unavailable, total CK> 2x the ULN (or the B fraction of CK) may be employed.

***NOTE-*** The preferred assays to use as biomarkers for the myocardial necrosis are troponin, CK-MB or total CK (in that order).

***AND ONE*** of the following:

1.Either ST segment depression or T wave abnormalities in the ECG; or

2. In the presence or absence of chest discomfort. Ischaemic symptoms may include;
   - Unexplained nausea and vomiting; or
   - Persistent shortness of breath secondary to left ventricular failure; or
   - Unexplained weakness, dizziness, light headedness, or syncope

**ST- Elevation MI (STEMI)**

***AT LEAST ONE of the following biochemical indicators for detecting myocardial necrosis MUST be present (see below for a definition of Reference Control Limits)***

1.Troponin T or I: Maximal concentration of troponin T or I > the MI diagnostic limit on at least one occasion within the first 24 hours from the index clinical event;

2.CK-MB:
- Maximal value of CK-MB > 2x the upper limit of normal (ULN) on one occasion during the first hours after the index clinical event; OR
- Maximal value of CK-MB (preferable CK-MB mass) > ULN on two successive samples.

3.Total CK: Only where troponin or CK-MB assays are unavailable, total CK > 2x the ULN (or the B fraction of CK) may be employed.

***NOTE-*** The preferred assays to use as biomarkers for the myocardial necrosis are troponin, CK-MB or total CK (in that order).

***AND ONE*** of the following ECG changes:

1.ST segment elevation: New or presumed new ST segment elevation at the J point in two or more contiguous leads with the cut-off points >=0.2mV in leads V1, V2, or V3 or >= 0.1 mV in other leads;

2. Development of any Q wave in leads V1 through V3, or the development of a Q-wave >=30ms (0.03s) in leads I, II, aVL, aVF, V4, V5, or V6. (Q wave changes must be present in any two contiguous leads, and be >= 1mm in depth).

***Defining Reference Control Values (MI Diagnostic Limit and Upper Limit of Normal):***
*Reference values must be determined in each laboratory by studies using specific assays with appropriate quality control, as reported in peer-reviewed journals. Acceptable imprecision (coefficient of variation) at the 99th percentile for each assay should be defined as <=10%. Each individual laboratory should confirm the range of reference values in their specific setting.*

# ANZSCTS CARDIAC SURGERY DATABASE

Medical Record No. ☐☐☐☐☐☐☐☐☐☐
First character here

## SECTION 4: PREVIOUS INTERVENTIONS

56.0 Previous Cardiothoracic Intervention ***(open or percutaenous)*** ◯ Yes ◯ No → **If YES continue below, if NO skip to Section 5**

**Previous Cardiac Surgery**

56.1 Previous ***open cardiac*** surgery ◯ Yes ◯ No → **If YES continue answering below, If NO skip to Q 56.2**

56.1.1 No. of prior cardiac operations with cardiopulmonary bypass ☐ (0-9)

56.1.2 No. of prior cardiac operations without cardiopulmonary bypass ☐ (0-9)

56.1.3 - 5 Type of previous surgery -***select all that apply*** ◯ CABG ◯ Off Pump CABG ◯ Valve ◯ Other Cardiac

**Previous Percutaneous Intervention**

56.2 Previous Percutaenous Intervention ◯ Yes ◯ No → **If YES continue answering below, If NO skip to Section 5**

56.2.1 Previous TAVR ◯ Yes ◯ No

56.2.2 Previous PTCA/Stent ◯ Yes ◯ No **If YES** → In which admission? ◯ This Admission ◯ Remote

**If YES to This Admission** Interval ☐☐☐ hrs

56.2.3 Non Surgical Balloon Valvuloplasty ◯ Yes ◯ No

56.2.4 ASD/PFO Device Closure ◯ Yes ◯ No

56.2.5 VSD Device ◯ Yes ◯ No

56.2.6 Left Atrial Appendage Occlusion ◯ Yes ◯ No

56.2.7 Electrophysiology Ablation ◯ Yes ◯ No

56.2.8 Percutaneous Mitral Valve Repair ◯ Yes ◯ No

56.2.9 Previous TMVR ◯ Yes ◯ No

## SECTION 5: HAEMODYNAMICS

57.0 Patient Height ☐☐☐ cm

58.0 Patient Weight ☐☐☐ kg

} *Perfusionist to complete*

59.0 Cardiac Catheterisation (Angiography) ◯ Yes ◯ No **If YES** → 59.1 Date ☐☐ / ☐☐ / ☐☐☐☐ (d d / m m / y y y y)

60.0 LVEF Method ◯ No ◯ Angiogram ◯ Radionuclide ◯ Echocardiogram ◯ MRI ◯ Unknown

60.1 LVEF ☐☐ %

60.2 LVEF Estimate ◯ Normal (>60%) ◯ Mild Impairment (46-60%) ◯ Moderate (30-45%) ◯ Severe (<30%)

61.0 Left Main Coronary Artery Stenosis >50% ◯ Yes ◯ No

62.0 No. Diseased Coronary Systems: ◯ None ◯ One ◯ Two ◯ Three

| SECTION 4: PREVIOUS INTERVENTIONS - *DEFINITIONS* | |
|---|---|
| 56.0 Previous Cardiothoracic Intervention (open or percutaneous) | Has the patient undergone any previous cardiovascular intervention, either open or percutaneous, prior to the index operation? This includes all forms of percutaneous angioplasty and transcathetar procedures for cardiac interventions or prior interventions in the same admission episode. |
| 56.2.4 ASD/PFO Device Closure | Closure by percutaneous technique of atrial septal defect or patient foramen ovale prior to the index admission (including those done in current admission). |
| 56.2.5 VSD Device Closure | Closure by percutaneous technique of Ventricular Septal Defect (including those done in current admission). |

| SECTION 5: HAEMODYNAMICS - *DEFINITIONS* | |
|---|---|
| 60.0 LVEF Method | Was the Left Ventricular Ejection Fraction measured, and how was this information obtained?<br>1 = Not measured<br>2 = Angiogram(angiographic LV gram, obtained during cardiac catheterisation)<br>3 = Radionuclide (nuclear)<br>4 = Echocardiogram (TTE or TOE)<br>5 = Magnetic Resonance Imaging |
| 61.0 Left Main Coronary Artery Stenosis > 50% | Any stenosis that involves any parts of the left main. Left main coronary stenosis is present when there is >50% compromise of vessel diameter in any angiographic view. |
| 62.0 Number of Diseased Coronary Systems | The number of (the three) major coronary systems (LAD system, circumflex system, and/or right coronary system) with >50% narrowing in any angiographic view. The number of diseased systems should be the number of systems requiring surgical approach at that operation.<br><br>NOTE: Left main disease (>50%) is counted as TWO systems (LAD and circumflex).<br>For example, left main and RCA would count as THREE in total.<br>Dominant circumflex counts as TWO systems.<br>LMCAD associated with dominant circumflex counts as THREE systems.<br><br>If a system has not been grafted previously and the graft has no haemodynamically significant stenosis, then that system is NOT counted as diseased. If a previous graft requires replacement, then that system IS counted as diseased.<br><br>IF THERE ARE NO DISEASED CORONARY ARTERY SYSTEMS THEN INDICATE 0. |

# ANZSCTS CARDIAC SURGERY DATABASE

Medical Record No. ☐☐☐☐☐☐☐☐☐☐ (First character here)

## SECTION 6: OPERATIVE STATUS / CATEGORY

63.0 Consultant Surgeon ☐☐☐ (code)

64.0 Operating Surgeon ☐ Consultant ☐ Senior Registrar ☐ Trainee Registrar ☐ Overseas Fellow ☐ Oversight

65.0 Status -***please read definition overleaf*** ◯ Elective ◯ Urgent ◯ Emergency ◯ Salvage

**If procedure has been classified as URGENT** → Provide a reason for urgent classification from list below

- ◯ Acute Myocardial Infarction (AMI) stabilised and not requiring emergency operation
- ◯ Pre-op Intra-Aortic Balloon Pump (IABP)
- ◯ Threatening coronary anatomy with acute symptoms
- ◯ Unstable angina requiring IV therapy
- ◯ Severe acute valve dysfunction either native or prosthetic

66.0 Direct transfer from cathlab/ ICU to theatre ***-see definition*** ◯ Yes ◯ No

**Category**

67.0 Coronary Artery Bypass ◯ Yes ◯ No

68.0 Valve Surgery ◯ Yes ◯ No

69.0 Other Cardiac Surgery ◯ Yes ◯ No → **If YES select surgery type from list below, if NO skip to Q 70.0**

| | | | |
|---|---|---|---|
| ◯ LV Aneurysm | ◯ Acquired VSD | ◯ ASD | ◯ Trauma |
| ◯ LVOT Myectomy | ◯ LV Rupture Repair | ◯ Pericardiectomy | ◯ Pulm. Thrombo - Endarterectomy |
| ◯ LV Recontruction | ◯ Pulmonary Embolectomy | ◯ Cardiac Tumour | ◯ Cardiac Transplant |
| ◯ Cardiopulmonary Transplant | ◯ Other Congenital | ◯ Permanent LV Epicardial Lead | ◯ Left Atrial Appendage Closure |
| ◯ Atrial Arrhythmia Surgery | ◯ Other | | |

**If YES to ATRIAL ARRHYTHMIA Surgery** → PREDOMINANT Lesion Set and Technique — Lesion Set ☐ 1 - 9 (see overleaf); Energy Source ☐ 1 - 8 (see overleaf)

**If YES to OTHER Surgery** → **Record the specific procedure that was performed** ☐

70.0 Aortic Procedure ◯ Yes ◯ No → **If YES continue answering below if NO skip to Q 71.0**

70.1 Aortic Pathology/Aeitiology
- ◯ Aortic Aneurysm
- ◯ Aortic Dissection → **If YES** 70.1.1 When ◯ Acute (<=2 weeks) ◯ Non-Acute (>2 weeks)
- ◯ Traumatic Transection (occuring within the last 2 weeks)
- ◯ Calcification
- ◯ Other

70.2- 70.3 Aortic Procedure Type
- ◯ Direct Aortoplasty
- ◯ Endarterectomy
- ◯ Patch Repair
- ◯ Replacement → **If Procedure was a Replacement** Location ◯ Ascending ◯ Arch ◯ Descending ◯ Thoraco-Abdominal

71.0 Other Non-Cardiac ◯ Yes ◯ No → **If YES continue answering below, if NO skip to SECTION 7**

71.1 Carotid Endarterectomy ◯ Yes ◯ No

71.2 Lung Resection ◯ Yes ◯ No

71.3 Other Vascular Surgery ◯ Yes ◯ No

71.4 Other Thoracic Surgery ◯ Yes ◯ No

74.5 Other ◯ Yes ◯ No → **If YES** **Record the specific procedure that was performed** ☐

| SECTION 6: OPERTAIVE STATUS / CATEGORY -*DEFINITIONS* | |
|---|---|
| **65.0 STATUS** | |
| Elective | The procedure could be deferred without increased risk of compromised cardiac outcome. |
| Urgent | Not routine - clinical reasons for operating in this admission -<br>a) Within 72 hours of angiography if index operation was performed in the same admission as angiography (here 'same admission' includes situations where angiography was performed in another hospital prior to direct transfer to current hospital where index operation is to be performed) **OR**<br>b) Within 72 hours of an unplanned admission (in patients who had a previous angiogram and was scheduled for surgery but admitted acutely) **OR**<br>c) Procedure required during same hospitalisation in a **clinically compromised** patient in order to minimise chance of further clinical deterioration. |
| Emergency | Unscheduled surgery required in next available theatre on same day (as admission) due to refractory angina or haemodynamic compromise. |
| Salvage | The patient underwent CPR en route to the operating room, prior to surgical incision. |

| | |
|---|---|
| 65.1 Urgent Reason | 1. Acute Myocardial Infarction (AMI) stabilised and not requiring emergency operation.<br>2. Pre-op Intra-Aortic Balloon Pump (IABP)<br>3. Threatening coronary anatomy with acute symptoms<br>4. Unstable angina requiring IV therapy<br>5. Severe acute valve dysfunction either native or prosthetic |

| | |
|---|---|
| 66.0 Direct Transfer from Cathlab/ICU to theatre | Patient required direct transfer to theatre for ongoing management as a results of a cardiac catheter lab event. Includes transfers directly from cardiac catheter lab, ICU or general ward as well as patients who have been temporarily transferred to a ward, usually CCU or ICU for stablisation while the OR is being prepared. Typically due to indications such as ischaemia, rest angina despite maximal treatment, pulmonary oedema requiring intubation, or shock. |

| **69.0 Other Cardiac Surgery** | |
|---|---|
| VSD (Acquired) | The index operation is for the correction of an acquired (usually ischaemic) ventricular septal defect (VSD). |
| ASD | The index operation is for the correction of an atrial septal defect (excludes closure of incidental PFO) |
| LVOT Myectomy | This procedure is performed for either hypertrophic obstructive cardiomyopathy or left ventricular muscular dynamic LVOT obstruction, or in cases of tunnel stenosis in the left ventricular outflow tract. This procedure involves excision of left ventricular endocardial muscle out of the left ventricular outflow tract. |
| LV Rupture Repair | The index operation is for ischaemic rupture of the free wall of the left ventricle. Does not include traumatic LV rupture repair. |
| Pulm. Thrombo-Endarterectomy | The index operation performed for chronic pulmonary thrombo-embolic disease. It involves cardiopulmonary bypass, and usually hypothermic circulatory arrest, and incisions are made in the right and left (or both) pulmonary arteries, and an endartectomy performed out into the distal branches. |
| LV Reconstruction | The index operation is for reshaping of the left ventricle by lateral excision (Batista). Does not include resection and repair of chronic left ventricular aneurysm, by whatever technique. |
| Permanent LV Epicardial Lead | The index operation includes insertion of a permanent LV Epicardial Lead. |
| Atrial Arrhythmia Surgery | The index operation is for paroxysmal, persistent or permanent atrial tachyarrhythmia. |

| **69.17.1 - 69.17.2 Atrial Arrhythmia Surgery** | |
|---|---|
| **Lesion Set:** | **Technique or Energy Source:** |
| 1=Cox-Maze III<br>2=Radial<br>3=Mini-Maze<br>4=Left Atrial Reduction<br>5=Pulmonary Vein Isolation<br>6=Left Arial Only<br>7=Right Atrial Only<br>8=Other<br>9=Cox-Maze IV | 1=Cut & Sew<br>2=Unipolar Radiofrequency<br>3=Bipolar Radiofrequency<br>4=Cryoablation<br>5=Microwave<br>6=Laser<br>7=Ultrasound<br>8=Other |

## SECTION 7: MINIMALLY INVASIVE

72.0 Minimally Invasive Open Technique Attempted (non-standard incision) ◯ Yes ◯ No

73.0 Robotically Assisted ◯ Yes ◯ No

## SECTION 8: CPB AND SUPPORT

74.0 Cardiopulmonary Bypass Used ◯ Yes ◯ No → **If YES continue answering below if NO skip to Q 75.0**

74.1 Cardioplegia ◯ Yes ◯ No → **If YES** Type ◯ Hyperkalaemic ◯ Bretschneider HTK (Custodial)

74.3 Cumulative Cross-Clamp Time ☐☐☐ min

74.4 Cumulative Cardiopulmonary Bypass Time (Perfusion Time) ☐☐☐ min

74.5 Intra-Operative Haemoglobin ☐☐☐ g/L

75.0 Intra Aortic Balloon Pump (IABP) ◯ Yes ◯ No → **If YES** When ◯ Pre-Operative ◯ Intra-Operative ◯ Post-Operative
Indication ◯ Haemodynamic Instability ◯ CBP Wean ◯ PTCA/PCI Support ◯ Prophylactic ◯ Unstable Angina

76.0 Other Mechanical Support (ECMO) ◯ Yes ◯ No → **If YES** When ◯ Pre-Operative ◯ Intra-Operative ◯ Post-Operative
Indication ◯ Cardiac Failure ◯ Rescue/Salvage ◯ Respiratory Failure ◯ Hypothermia

77.0 Other Mechanical Support (VAD) ◯ Yes ◯ No → **If YES** When ◯ Pre-Operative ◯ Intra-Operative ◯ Post-Operative
Indication ◯ Bridge to Transplantation ◯ Bridge to Recovery ◯ Destination ◯ Postcardiotomy Ventricular Failure ◯ Device Malfunction ◯ End of Life

78.0 Intra-Operative TOE ◯ Yes ◯ No

79.0 Intra-Operative Antifibrinolytic Use ◯ Yes ◯ No → **If YES** 79.1Type ◯ Trasylol ◯ Aminocaproic Acid ◯ Tranexamic Acid ◯ Other
◯ Unknown ◯ ATACAS

## SECTION 9: CORONARY BYPASS

80.0 Intraoperative decision to graft coronary artery ◯ Yes ◯ No ◯ Unknown

81.0 ITA used ◯ Yes ◯ No → **If YES continue answering below if NO skip to Q 82.0**

81.1 LITA used ◯ Yes ◯ No → **If YES** 81.1.1 Skeletonised ◯ Yes ◯ No

81.2 RITA used ◯ Yes ◯ No → **If YES** 81.2.1 Skeletonised ◯ Yes ◯ No

82.0 No. of RA conduits harvested ☐ (0-2)

83.0 No. of distal arterial grafts ☐ (0-9)

84.0 No. of ITA distal anastomoses ☐ (0-6)

85.0 No. of radial distal anastomoses ☐ (0-6)

86.0 No. of vein distal anastomoses ☐ (0-9)

87.0 No. of GEPA distal anastomoses ☐ (0-6)

88.0 Arterial T or Y grafts used ◯ Yes ◯ No

## SECTION 10: VALVE SURGERY

**Valve Procedure**

| | Enter code from list below | | Model No | Serial | Lot No | Size |
|---|---|---|---|---|---|---|
| **Aortic** | | Implant | | | | |
| | | Explant | | | | |
| **Mitral** | | Implant | | | | |
| | | Explant | | | | |
| **Tricuspid** | | Implant | | | | |
| | | Explant | | | | |
| **Pulmonary** | | Implant | | | | |
| | | Explant | | | | |

**PROCEDURE CODES**

1. No
2. Annuloplasty Only
3. Replacement
4. Mitral or Tricuspid: Repair or Reconstruction with Annuloplasty
5. Mitral or Tricuspid: Repair or Reconstruction without Annuloplasty
6. Root Reconstruction with Valved Conduit
7. Root Reconstruction with Valve Sparing
8. Re-suspension of the Aortic Valve
9. Resection of Sub-Aortic Stenosis
10. Commissurotomy or Valvotomy with Annuloplasty Ring
11. Commissurotomy or Valvotomy without Annuloplasty Ring
12. Repair of Paravalvular Leak
13. Valvectomy (no replacement)
15. Ross Procedure
16. Inspection Only
17. Decalcification of Valve Only
18. Aortic Subcommissural Annuloplasty
19. Cusp Modification
20. Thrombus Removal
21. Root Enlargement (Manougian type excludes Nicks)
22. Transcatheter Aortic Valve Replacement (TAVR)
23. Aortic Valvuloplasty with subcommissural annuloplasty
24. Aortic Valvuloplasty without subcommissural annuloplasty
25. Alfieri Suture
26. Removal of tumour valve tissue (e.g. Fibroelastoma)
27. Insertion of a Mitraclip device
28. Transcatheter Mitral Valve Replacement (TMVR)
29. Replacement of Pulmonary Root as part of a Ross Procedure

**Valve Pathophysiology**

| | Aortic | Mitral | Tricuspid | Pulmonary |
|---|---|---|---|---|
| **Stenosis** | ○ Yes ○ No | ○ Yes ○ No | ○ Yes ○ No | ○ Yes ○ No |
| **Regurgitation/ Insufficiency** (0-4) | | | | |
| **Pathology/ Aetiology** (see codes) | | | | |

**INSUFFICIENCY CODES**

| | |
|---|---|
| 0 | None |
| 1 | Trivial |
| 2 | Mild |
| 3 | Moderate |
| 4 | Severe |

**VALVE AETIOLOGY CODES**

1. Rheumatic
2. Congenital
3. Ischaemic
4. Idiopathic Calcific
5. Myxomatous Degen.
6. Failed Prior Repair
7. Prosthetic Valve Failure
8. Peri-Prosthetic Leak
9. Prosthetic Valve Thrombosis
10. Active Infection
11. Previous Infection
12. Marfan's
13. Annuloaortic Ectasia
14. Other Degen. Disease
15. Dissection
16. Tumour
17. Trauma
18. Iatrogenic
20. Functional
21. Carcinoid Syndrome
22. Failed TAVR
23. Failed TMVR
99. Other

# ANZSCTS CARDIAC SURGERY DATABASE

Medical Record No.

First character here

## SECTION 11: BLOOD PRODUCT USE

**THIS SECTION REFERS TO CUMULATIVE (INTRA-OPERATIVE + POST-OPERATIVE) BLOOD PRODUCT USE**

94.0 RBC ○ Yes ○ No → **If YES** 94.1 Number of Bank RBC (units) ☐☐

95.0 Non RBC ○ Yes ○ No → **If YES** 95.1 Number of Platelets (units) ☐☐ 95.2 Number of Novo7(mg) ☐☐.☐

95.3 Number of Cryo (units) ☐☐ *See overleaf for conversion*

## SECTION 12: POST-OPERATIVE DATA

96.0 ICU Admission - Date/Time ☐☐/☐☐/☐☐☐☐ (d d / m m / y y y y) ☐☐:☐☐ (H H : M M)

97.0 Extubation - Date/Time ☐☐/☐☐/☐☐☐☐ ☐☐:☐☐

98.0 ICU Discharge - Date/Time ☐☐/☐☐/☐☐☐☐ ☐☐:☐☐

99.0 Re-admitted to ICU ○ Yes ○ No

100.0 Re-intubation ○ Yes ○ No

100.1 Re-intubation - Date/Time ☐☐/☐☐/☐☐☐☐ ☐☐:☐☐

100.2 Re-extubation- Date/Time ☐☐/☐☐/☐☐☐☐ ☐☐:☐☐

101.0 ICC Loss (First 4 hours post surgery): ☐☐☐☐ mls

### COMPLICATIONS - Must not have been present pre-operatively

102.0 Return to theatre ○ Yes ○ No → **If YES**

| | | |
|---|---|---|
| Re-op Valve Dysfunction | ○ Yes | ○ No |
| Re-op Bleeding or Tamponade | ○ Yes | ○ No |
| Re-op Graft Occlusion | ○ Yes | ○ No |
| Re-op Deep Sternal Wound Infection | ○ Yes | ○ No |
| Re-op Deep Thoracotomy Wound Infection | ○ Yes | ○ No |
| Re-op Insertion of Pacemaker/AICD | ○ Yes | ○ No |
| Re-op Other Cardiac | ○ Yes | ○ No |
| Re-op Other Non-Cardiac | ○ Yes | ○ No |

103.0 New Renal Insufficiency ○ Yes ○ No → **If YES** Haemofiltration ○ Yes ○ No

104.0 Highest Post-Operative Creatinine Level ☐☐☐☐ µmol/l

105.0 Peri-/Post-Operative MI ○ Yes ○ No

106.0 Peri-/Post-Operative Cardiogenic Shock ○ Yes ○ No

107.0 Lowest Post-Operative Haemoglobin ☐☐☐ g/L

| | | | |
|---|---|---|---|
| 108.0 - 110.0 Cardiac Inotrope or Vasopressor use (Mark all that apply) | for longer than 4 hours post-operatively | ○ Yes | ○ No |
| | for Low Cardiac Output Syndrome | ○ Yes | ○ No |
| | for Low SVR Syndrome | ○ Yes | ○ No |
| 111.0 New Cardiac Arrhythmia ○ Yes ○ No → **If YES** | Heart Block (requiring PPM) | ○ Yes | ○ No |
| | Other Brady-Arrhythmia (requiring PPM) | ○ Yes | ○ No |
| | Cardiac Arrest | ○ Yes | ○ No |
| | Atrial Arrhythmia (requiring treatment) | ○ Yes | ○ No |
| | Ventricular Tachycardia | ○ Yes | ○ No |
| 112.0 -114.0 New Neurologic → | Stroke Permanent (>72hrs) | ○ Yes | ○ No |
| | Stroke Transient (<72 hrs) | ○ Yes | ○ No |
| | Continuous Coma (=> 24 hrs) | ○ Yes | ○ No |
| 115.0 -117.0 New Pulmonary → | Ventilation Prolonged (>24 hrs) | ○ Yes | ○ No |
| | Pulmonary Embolism | ○ Yes | ○ No |
| | Pneumonia | ○ Yes | ○ No |
| 118.0 -122.0 New Infection → | Sternal Deep Wound | ○ Yes | ○ No |
| | Superficial Access Wound | ○ Yes | ○ No |
| | Donor Site Deep Wound | ○ Yes | ○ No |
| | Deep Access Wound of Parasternal Site | ○ Yes | ○ No |
| | Septicaemia | ○ Yes | ○ No |
| 123.0 -124.0 New Vascular → | Aortic Dissection | ○ Yes | ○ No |
| | Acute Limb Ischaemia | ○ None | ○ Upper Limb ○ Lower Limb |
| 125.0 -127.0 New Other → | Anticoagulant Complications | ○ Yes | ○ No |
| | GIT Complications | ○ Yes | ○ No |
| | Multi-System Failure | ○ Yes | ○ No |

| SECTION 11: BLOOD PRODUCT USE -*DEFINITIONS* | |
|---|---|
| 94.0 Blood Products: RBC | Allogenic red blood cells (RBC) transfused intra and/or postoperatively. Does not include pre-donated blood, pump residual blood, cellsaver blood or chest tube recirculated blood. |
| 95.0 Blood Products: Non RBC | A transfusion of blood products other than RBC (eg. FFP, Platelets) given intra and/or post-operatively. (Excludes Albumin) |
| 95.2 Novo 7 Conversion Calculation | Novo 7 comes in 1, 2, 4 and 8 mg vials. Dose administered is between 50-90mcg per kg.<br>E.g. 50mcg x 70kg person = 3500mcg=3.5mg |

| SECTION 12: POST-OPERATIVE DATA -*DEFINITIONS* | |
|---|---|
| 102.0 Return to Theatre | Patient returned to the operating theatre. Includes operative procedures done in the ICU that normally would be performed in the operating room. |
| 103.0 New Renal Insufficiency | Acute post-operative renal insufficiency characterised by one of the following:<br>a.) Increased serum creatinine to >0.2mmol/l (>200 µmol/l) AND a doubling or greater increase in creatinine over the baseline pre-operative value AND the patient did not require pre-operative dialysis/haemofiltration;<br>b.) A new post-operative requirement for dialysis/haemofiltration where they did not require this pre-operatively.<br>**Renal insufficiency must not be present pre-operatively.** Pre-operative renal transplant does not count as renal insufficiency if the patient did not have impaired liver function and did not require dialysis/haemofiltration. |
| 103.1 Haemofiltration | Acute institution of haemofiltration (or dialysis) as treatment for new renal failure. Excludes haemofiltration for removal of fluid with normal serum urea and creatinine. |
| 105.0 Peri-/Post-operative MI | Diagnosed by finding at least two of the following criteria:<br>a.) Enzyme level elevation: either 1) CK-MB>30 units; or 2) troponin >20.0 micrograms /L, or established level at own institution (provided operation does not involve myocardial incision); b.) New wall motion abnormalities;<br>c.) Serial ECG (at least two) showing Q waves, duration =>0.03ms in 2 contiguous leads. |
| 106.0 Peri-/Post-operative Cardiogenic Shock | Only select 'yes' if ALL the following criteria apply:<br>a.) Sustained (>30 mins) episode of systolic blood pressure <90mm Hg or the requirement for parenteral inotropic or vasopressor agents or mechanical support (e.g. intra-aortic balloon pump (IABP), extracorporeal circulation, ventricular assist devices to maintain BP>90mm Hg); **AND**<br>b.) Evidence of elevated filing pressures (e.g. pulmonary congestion on examination or chest radiograph); **AND**<br>c.) Evidence of end organ hypoperfusion (e.g. urine output 30mL/hours, or cold/diaphoretic extremities, or obtunded mental status if previously normal, etc.). |
| 109.0 Cardiac Inotrope Use for Low Cardiac Output Syndrome | An inotrope(s) was administered for low cardiac output syndrome longer than four hours post-operatively (with the intent to improve cardiac output, irrespective of the reasons for that decision). Does not include Milrinone. |
| 110.0 Cardiac- Vasopressor Use for Low SVR Syndrome | When a primarily alpha adrenergic agonist is given for low systemic vascular resistance syndrome for longer than four hours post-operatively with the intent to increase SVR (where SVR<800). This is usually in presence of high cardiac output. Does not include Noradrenaline given with Milrinone. |
| 111.1 Heart Block | New heart block requiring implantation of permanent pacemaker prior to discharge. |
| 111.2 Other Bradyarrhythmia | New other bradyarrhythmia, not otherwise specified, requiring implantation of permanent pacemaker prior to discharge. |
| 111.3 Cardiac Arrest | Either a.) VF; b.) Rapid VT with haemodynamic instability; c.) asystole; (d) Pulseless electrical activity (PEA) |
| 111.4 New Atrial Arrhythmia | New onset atrial fibrillation/flutter requiring treatment. Does not include recurrence of AF present pre-operatively. |
| 111.5 New Ventricular Tachycardia | New onset of ventricular tachycardia (> 6 beat run) requiring treatment. |
| 112.0 Stroke Permanent | A stroke or new central neurological deficit (defined as persistant loss of neurological function caused by an ischaemic or haemorrhagic event) persisting for > 72 hours peri or post-operatively. |
| 113.0 Stroke Transient | A transient new central neurological deficit that was completely resolved within 72 hours (TIA, RIND). |
| 114.0 Continuous Coma => 24hrs | New postoperative coma that persists for at least 24 hours. Only applicable to a non-sedated patient. |
| 115.0 Ventilation Prolonged > 24hrs | Pulmonary insufficiency requiring prolonged ventilatory support > 24hrs (cumulative). E.g. Adult Respiratory Distress Syndrome and pulmonary oedema. Cumulative period is used if patient is re-intubated. |
| 116.0 New Pulmonary Embolism | Diagnosed by study such as ventilation/perfusion (V/Q) scan or angiogram. |
| 117.0 Pneumonia | Pneumonia diagnosed post-operatively by one of the following: a.) Postive cultures of sputum or trans-tracheal aspirate; b.) Clinical, including haematological findings consistent with the diagnosis of pneumonia and radiographic evidence |
| 118.0 Infection - Sternal Deep | Infection of sternal bone, muscle and/or mediastinum. Must have wound debridement and **one** of following:<br>a.) Positive culture; b.) Treatment with antibiotics. |
| 119.0 Infection - Superficial Access | Infection involving the skin and subcutaneous tissue of the incision occurring within 30-days after the operative procedure AND<br>Patient must have one of the following:<br>a.) Purulent drainage from the superficial incision;<br>b.) Organisms isolated from an aseptically obtained culture of fluid or tissue from the superficial incision;<br>c.) Superficial incision deliberately opened by surgeon AND is culture-positive or not cultured AND patient has at least **one** of the following signs of infection: pain or tenderness, localised swelling, heat;<br>d.) Diagnosis of Superficial Incisional Surgical Site Infection by operating surgeon or assisting physician. |

| SECTION 12: POST-OPERATIVE DATA -*DEFINITIONS cont* | |
|---|---|
| 120.0 Donor Site Deep Wound Infection | Infection involving deep soft tissues (e.g. fascial and muscle layers and/or organs/spaces opened or manipulated during surgery) occurring within 30 days after the operative procedure if implant not present AND patient exhibits one of the following:<br>a.) Purulent drainage from deep soft tissue but not from the organ/space component of the surgical site;<br>b.) Spontaneous dehiscence at incision site or the wound is deliberately explored by a surgeon with the patient showing evidence of one or more of the following signs or symptoms:<br>- Fever > 38°C, localised pain or tenderness with culture-positive specimen. A culture-negative finding does not meet this criterion unless the patient was on antibiotics immediately prior to the wound being explored and/or the culture being taken.<br>- Organisms isolated from an aseptically obtained culture of fluid or tissue obtained from an organ/space.<br>- An abscess or other evidence of infection involving a deep/organ space found on direct examination, during reoperation, or by histopathologic or radiologic examination.<br>- Diagnosis of or antimicrobial treatment of a deep incisional or organ/space surgical site infection by operating surgeon or assisting physician. |
| 121.0 Deep Access Wound Infection of Parasternal Site -Not of Sternotomy | An infection involving a thoracotomy or parasternal site. Must have **one** of the following conditions:<br>a.) Wound opened with excision of tissue;<br>b.) Positive culture;<br>c.) Treatment with antibiotics |
| 122.0 Infection - Septicaemia | Septicaemia requires positive blood cultures supported by at least two of the following indices of clinical infection:<br>a.) Fever; b.) Elevated granulocyte cell counts; c.) Elevated and increasing CRP; d.) Elevated and increasing ESR, post-operatively. |
| 123.0 Aortic Dissection | Dissection occuring in any part of the aorta. |
| 124.0 Acute Limb Ischaemia | Any evidence of limb ischaemia. |
| 125.0 Anticoagulant complications | Bleeding, hemorrhage, and/or embolic events related to anticoagulant therapy. |
| 126. GIT complications | Postop occurrence of any GIT complication including:<br>a.) GI bleeding requiring transfusion;<br>b.) pancreatitis with abnormal amylase/lipase requiring nasogastric suction therapy;<br>c.) cholecystitis requiring cholecystectomy or drainage;<br>d.) mesenteric ischaemia requiring exploration;<br>e.) hepatitis;<br>f.) other GI complication |
| 127.0 Multi-system failure | Postop multi-system failure involving two or more of the following major organ systems failing concurrently for at least 48 hours:<br>a.) Renal - New renal failure (defined previously);<br>b.) Respiratory - Requires endotracheal intubation for respiratory dysfunction;<br>c.) Cardiac - the use of inotropes and/or IABP to treat low cardiac output;<br>d.) Hepatic failure on the basis of enzymes, and bilirubin estimation |

# ANZSCTS CARDIAC SURGERY DATABASE

Medical Record No.
First character here

## SECTION 13: MORTALITY / DISCHARGE / READMISSION

### DISCHARGE

128.0 Patient discharged to
- ◯ Home
- ◯ Local or Referring Hospital
- ◯ Hospital in the Home
- ◯ Hospital Mortality
- ◯ Rehabilitation Unit/Hospital
- ◯ Other Cardiac Unit

### MORTALITY

128.1 Post Discharge within 30 days of surgery: ◯ Yes ◯ No ◯ Unknown

128.2 Mortality Date: d d / m m / y y y y

Provide date of death in hospital during the index admission at any time after the procedure, or death after discharge from hospital within thirty days of the procedure

128.3 Mortality Location: ◯ Operating Room ◯ Hospital ◯ Home (inc. hospital in the home) ◯ Other Care facility

128.4 Mortality Primary Cause:
(choose one of the following)
- ◯ Cardiac Cause — **If yes →** ◯ Ischaemic ◯ Other Cardiac ◯ Unknown
- ◯ Neurologic Event
- ◯ Renal Failure
- ◯ Vascular Event (peripheral vascular or aortic but not aortic dissection)
- ◯ Infection — **If yes →** ◯ Septicaemia ◯ Endocarditis ◯ Other Infection ◯ Unknown
- ◯ Respiratory Failure
- ◯ Valvular Dysfunction
- ◯ Multisystem Failure (as previously defined in 127.0)
- ◯ Other
- ◯ Unknown
- ◯ Pulmonary Embolism
- ◯ Aortic Dissection

129.0 Cognisant patient elected to withdraw from treatment (see definition below) ◯ Yes ◯ No

### READMISSION

130.0 Readmitted <=30 Days from procedure: ◯ Yes ◯ No
(Does not include planned transfer to rehabilitation facility, short-stay wards or emergency. Date of surgery counts as day zero.)

Readmitted reason:
(choose one of the following)
- ◯ Anticoagulant Complication
- ◯ Arrhythmia
- ◯ Congestive Heart Failure (CHF)
- ◯ Valve Dysfunction
- ◯ Pericardial Effusion
- ◯ Cardiac Tamponade
- ◯ Pleural Effusion
- ◯ Other Complication related to Cardiac Surgery (e.g. renal, hepatic, GI etc)
- ◯ Deep Sternal Infection
- ◯ Incisional Complication
- ◯ Pneumonia or other Respiratory Complication
- ◯ Myocardial Infarction (MI)
- ◯ Recurrent Angina
- ◯ Other readmission unrelated to Cardiac Surgery

| *DEFINITIONS* | |
|---|---|
| 128.0 Discharge | Home: Discharged to home, with no planned contact before routine review<br>Hospital in the home: Discharged to home, with planned visits to home by medical or paramedical staff<br>Rehabilitation Unit/Hospital: Discharged for inpatient rehabilitation<br>Local or referring hospital: Discharged for continuing acute care<br>Hospital Mortality<br>Other Cardiac Unit: Transferred to another hospital for further cardiac surgical intervention (e.g ECMO) |
| 128.1 Mortality Post-discharge | Specify whether the patient died after discharge from hospital but within 30 days of surgery. |
| 128.5 Mortality Cause - Cardiac<br>Mortality Cause - Infection | Specify whether the patient died from cardiac ischaemia or from another cardiac complication.<br>Specify whether the patient died from septicaemia, endocarditis or other infection. |
| 129.0 Cognisant patient withdraws from treatment | Patient who was aware of the consequences to his/her actions, elected to withdraw treatment in circumstances where they would survive if treatment was continued. NOTE: Completing "yes" to this field implies automatic review of patient's hospital file and permission for ANZSCTS personnel to review their case. |
| 130.3 Readmission reason - Congestive heart failure | Readmitted as an inpatient within 30 days from the date of surgery for CHF, evidenced by one or more of following; a.) paroxysmal nocturnal dyspnoea (PND); b.) deteriorating dyspnoea on exertion (DOE) due to HF OR c.) chest x-ray (CXR) showing pulmonary congestion. |
| 130.11 Readmission reason - Pneumonia or other respiratory complication | Readmitted as an in-patient within 30 days from surgery for pneumonia or other respiratory complications. Diagnosed by one of the following; a.) **positive cultures** of sputum or trans-tracheal aspirate OR b.) clinical, including haematological findings consistent with the diagnosis of pneumonia and radiographic evidence. |
| 130.13 Readmission reason - Recurrent angina | Readmitted as an inpatient within 30 days from surgery for recurrent angina. Objective confirmation that chest pain is due to ischaemia by exercise test (ECG, nuclear, echo, exercise test or angiography) is required to meet this diagnosis. |

## Appendix 2.3: Extract from the ANZSCTS data definition manual

| Sl | Variable name (field) | Variable definition | Format /codes |
|---|---|---|---|
| 1 | OPERATION ID | This is an arbitrary number that uniquely and permanently identifies each operation. Once assigned to an operation, this can never be changed or reused. | Numeric |
| 2 | AGE | Age of the patient at surgery (In years). | Numeric |
| 3 | Gender | Gender of the patient. | 1 = Male<br>2 = Female |
| 4 | BMI (BMI) | Body Mass Index calculated by the following equation. [WKG / (HTM/100)2] Calculated automatically where height and weight is<br>available. | NUMERIC |
| 5 | Admission date (DOA) | Date Patient admitted/transferred to hospital where surgery performed. | DD/MM/YY |
| 6 | Elective day of surgery admit (DOSA) | Patient admitted for scheduled elective procedure on same day as procedure | 1 = Yes<br>0 = No |
| 7 | Smoking history (SMO_H) | A history confirming any form of tobacco use in the past | 1 = Yes<br>0 = No |
| 8 | Family history of CAD (FHCAD) | Whether any direct blood relatives have had any of the following at age <55:<br>a. Angina,<br>b. Myocardial infarction (MI),<br>c. Sudden cardiac death presumed to be from ischaemic heart disease.<br>d. Coronary intervention | 1 = Yes<br>0 = No |
| 9 | Diabetes (DB) | A history of diabetes, regardless of duration of disease or need for anti-diabetic agents. | 1 = Yes<br>0 = No |
| 10 | Diabetes – control (DB_CON) | Method of diabetic control, at time of intervention.<br>The most aggressive therapy should be indicated as per the following order: insulin > oral > diet.<br>1. No treatment for diabetes<br>2. Diet treatment only<br>3. Oral agent treatment<br>4. Insulin treatment (includes any combination with insulin) | 1 = None<br>2 = Diet<br>3 = Oral<br>4 = Insulin |
| 11 | Hypercholesterolaemia (HCHOL) | Whether the patient has a history of hypercholesterolaemia diagnosed and/or treated by a physician, and/or Cholesterol > 5.0 mmol/L, HDL <1.0 mmol/L or Triglycerides >2.0 mmol/L. | 1 = Yes<br>0 = No |
| 12 | Preoperative creatinine Level (PRECR) | Last serum creatinine recorded prior to surgery.<br>(≥50 μ mol/L to ≤ 2000 μ mol/L) | Numeric |

| Sl | Variable name (field) | Variable definition | Format /codes |
|---|---|---|---|
| 13 | Estimated glomerular filtration rate (eGFR) | 1. Convert preoperative serum creatinine (mmol/L) into mg/dL:<br>= PRECRE x (1000/88.4)<br>2. eGFR is calculated using the Cockroft Gault formulae:<br>For males: [WKG x (140 – AGE)] / [72 x serum creatinine]<br><br>For females: [WKG x (140 – AGE) x 0.85] / [72 x serum creatinine]<br>Calculated automatically where last preoperative serum creatinine and weight are available. mL/min per 1.73m2 | Numeric |
| 14 | Dialysis (DIAL) | Is the patient on dialysis pre-operatively? | 1 = Yes<br>0 = No |
| 15 | Hypertension (HYT) | Does the patient have a diagnosis of hypertension?<br>a. Documented history of hypertension diagnosed and treated with diet, medication and/or exercise.<br>b. Blood pressure >140 systolic or >90 diastolic on at least 2 occasions.<br>c. Currently on antihypertensive medication. | 1 = Yes<br>0 = No |
| 16 | Cerebrovascular disease (CBVD) | Whether the patient has had Cerebro-Vascular Disease, documented by any one of the following:<br>a. Unresponsive coma >24 hrs,<br>b. CVA (symptoms >72 hrs after onset)<br>c. RIND (recovery within 72 hrs),<br>d. TIA (recovery within 24 hrs)<br>e. Non-invasive carotid test with 50% diameter stenosis (equivalent to 75% cross-sectional area stenosis). | 1 = Yes<br>0 = No |
| 17 | Peripheral vascular disease (PVD) | The patient's history of PVD either aneurysmal or chronic or acute occlusion or narrowing of the arterial lumen of the aorta or extremities. Includes the following:<br>a. Claudication either with exertion or rest,<br>b. Amputation for arterial insufficiency,<br>c. Vascular reconstruction, bypass surgery, or percutaneous intervention to the extremities<br>d. Documented aortic aneurysm,<br>e. Documented renal artery stenosis g.<br>f. Positive non-invasive testing documented | 1 = Yes<br>0 = No |
| 18 | Respiratory disease (LD) | Whether the patient has chronic lung disease, and severity level according to the following classification:<br>• On chronic inhaled or oral bronchodilator therapy,<br>• On chronic oral steroid therapy directed at lung disease,<br>• Room Air p02 < 60 or Room Air pC02 > 50, or mechanical ventilation for chronic lung disease | 1 = Yes<br>0 = No |
| 19 | Infective endocarditis (IE) | A patient presenting with valvular disease of infectious aetiology with past or present positive blood culture, or postoperative pathology confirmation. | 1 = Yes<br>0 = No |

| Sl | Variable name (field) | Variable definition | Format /codes |
|---|---|---|---|
| | | | |
| 20 | Infective endocarditis Type (IE_T) | Type of infective endocarditis<br>• Active: If the patient is currently being treated for endocarditis, the disease is considered active.<br>• Treated: If no antibiotic medication (other than prophylactic medication) is being given at the time of surgery, then the infection is considered treated. | 1 = Active<br>2 = Treated |
| 21 | Immunosuppressive rx (IMSRX) | Use of any form of immunosuppressive therapy, including systemic steroid therapy equivalent to ≥ 5mg prednisolone within 30 days or less preceding the operative procedure. | 1 = Yes<br>0 = No |
| 22 | Myocardial infarction (MI) | Patient hospitalised at any time for a Myocardial Infarction documented in the medical record. | 1 = Yes<br>0 = No |
| 23 | Angina (CCS) | Canadian Cardiovascular Society Classification. The highest class leading to current episode of hospitalisation and/or intervention:<br>0 No angina symptoms.<br>1 Ordinary physical activity, such as walking or climbing the stairs does not cause angina. Angina may occur with strenuous, rapid or prolonged exertion at work or recreation.<br>2 There is slight limitation of ordinary activity. Angina may occur with moderate activity such as walking or climbing stairs rapidly, walking uphill, walking or stair climbing after meals or in the cold, in the wind, or under emotional stress, or walking more than two blocks on the level, and climbing more than one flight of stairs at normal pace under normal conditions.<br>3 There is marked limitation of ordinary physical activity. Angina may occur after walking one or two blocks on the level or climbing one flight of stairs under normal conditions at a normal pace.<br>4 There is inability to carry on any physical activity without discomfort; angina may be present at rest. | Numeric (0-4) |
| 24 | Angina – type (ANG_T) | Indicate the type of angina present at the time of surgery:<br>1. Stable: Angina which is controlled by oral or transcutaneous medication.<br>2. Unstable: The presence of ischemia that requires hospitalisation and use of intravenous nitrate, heparin therapy, s.c. clexane or intravenous Tyrofiban for control. | 1 = Stable<br>2 = Unstable |
| 25 | History of congestive heart failure (CHF) | Whether a physician has ever diagnosed Congestive Heart Failure (CHF) by two of the following:<br>a. Paroxysmal nocturnal dyspnoea (PND);<br>b. Dyspnoea on exertion (DOE) due to heart failure;<br>c. Chest X-ray (CXR) showing pulmonary congestion, OR | 1 = Yes<br>0 = No |

| Sl | Variable name (field) | Variable definition | Format /codes |
|---|---|---|---|
| | | d. Patient has received treatment for this – ACE inhibition, diuretics, Carvedilol or digoxin | |
| 26 | CHF at current admission (CHF_C) | The diagnosis and management of CHF was made this admission, OR The management changed due to deterioration in CHF. | 1 = Yes<br>0 = No |
| 27 | NYHA class - (NYHA) | NYHA: New York Heart Association Class - the highest level leading to current episode of hospitalisation and/or procedure.<br>I. Patients with cardiac disease but without resulting limitation of physical activity. Ordinary physical activity does not cause undue fatigue, palpitation, or dyspnoea.<br>II. Patients with cardiac disease resulting in slight limitation of physical activity. They are comfortable at rest. Ordinary physical activity results in fatigue, palpitations, or dyspnoea.<br>III. Patients with cardiac disease resulting in marked limitation of physical activity. They are comfortable at rest. Less than ordinary physical activity results in fatigue, palpitations, or dyspnoea.<br>IV. Patients with cardiac disease resulting in inability to carry on any physical activity without discomfort. Symptoms of cardiac insufficiency may be present even at rest. If any physical activity is undertaken, discomfort is increased. | Numeric (1-4) |
| 28 | Cardiogenic shock (SHOCK) | Is the patient, at the time of procedure, in a clinical state of hypo perfusion according to either of the following criteria:<br>a. Hypotension (a systolic blood pressure < 90 mmHg) and/or CI <2.0 for at least 30 minutes<br>b. The need for supportive measures to maintain a systolic<br>c. pressure > or = 90 mmHg or a CI > 2.0 | 1 = Yes<br>0 = No |
| 29 | Resuscitation (within one hour pre-op) (RESUS) | The patient required cardiopulmonary resuscitation, or initiation of treatment for cardiogenic shock, within one hour before the start of the operative procedure. | 1 = Yes<br>0 = No |
| 30 | Arrhythmia (ARRT) | Was there a pre-operative arrhythmia present by clinical documentation of any one of the following:<br>a. Atrial fibrillation/flutter requiring Rx;<br>b. Heart block;<br>c. Sustained Ventricular Tachycardia or Ventricular fibrillation requiring cardioversion and/or IV Amiodarone;<br>d. Other arrhythmia (e.g. Sick Sinus Syndrome) | 1 = Yes<br>0 = No |
| 31 | Permanent pacemaker in situ (PACE) | Patient has a permanent pacemaker implanted. | 1 = Yes<br>0 = No |
| 32 | Medications - Anticoagulation therapy (MEDAC) | Patient given warfarin/heparin/low MW heparinoid ≤ 24 hours prior to surgery | 1 = Yes<br>0 = No |

| Sl | Variable name (field) | Variable definition | Format /codes |
|---|---|---|---|
| 33 | Medications – inotropes (MEDIN) | Patient on inotropes prior to surgery, for haemodynamic support excluding renal dose Dopamine. | 1 = Yes<br>0 = No |
| 34 | Medications - iv nitrates (MEDNI) | Patient on IV Nitrates prior to surgery. | 1 = Yes<br>0 = No |
| 35 | Medications – steroids (MEDST) | Patient given systemic steroids prior to surgery. | 1 = Yes<br>0 = No |
| 36 | Previous cardiothoracic intervention (surgical or percutaneous) (POP) | Has the patient undergone any previous cardiovascular intervention, surgical or non-surgical including those done during the current admission? Includes all forms of percutaneous angioplasty and thrombolytic therapy for cardiac indications.<br>If the patient has had for example a PTCA Stent at another hospital and was then transferred to this hospital for surgery; ie. same admission episode. | 1 = Yes<br>0 = No |
| 37 | Cardiac catheterization (Angiogram or Pressure study) (CATH) | Has the patient had a cardiac catheter for angiogram or pressure study. | 1 = Yes<br>0 = No |
| 38 | Date of cardiac catheterization (CATH_W) | The date the patient had a cardiac catheter inserted. | DD/MM/YY |
| 39 | Ejection fraction (EF) | The percentage of the blood emptied from the left ventricle at the end of the contraction. Use the most recent determination prior to intervention. Enter a percentage in the range of 5 -90. | Numeric (5-90) |
| 40 | EF estimate (EF_EST) | If Nuclear scan, echo or angiogram did not yield a digital EF%, provide an estimate from reviewing the study. Choose one of:<br>1. Normal (LV-EF > 60%)<br>2. Mild Impairment (EF 46-60%)<br>3. Moderate (EF 30-45%)<br>4. Severe (EF<30%) | 1 = Normal<br>2 = Mild<br>3 = Moderate<br>4 = Severe |
| 41 | Left main stenosis > 50% (LMD) | Any stenosis that involves any parts of the Left Main. Left Main Coronary stenosis is present when there is > 50% compromise of vessel diameter in any angiographic view. | 1 = Yes<br>0 = No |
| 42 | Number diseased coronary systems (DISVES) | The number of major coronary systems (LAD system, Circumflex system, and/or Right System) with > 50% narrowing in any angiographic view. The number of diseased systems should be the number of systems requiring surgical approach at that operation.<br>NOTE: Left main disease (>50%) is counted as TWO systems (LAD and Circumflex). For example, left main and RCA would count as THREE in total. Dominant circumflex counts as TWO systems. | 0 = None<br>1 = One<br>2 = Two<br>3 = Three |
| 43 | Status (STAT) | 1. Elective: The procedure could be Deferred without increased risk of compromised cardiac outcome.<br>2. Urgent: Not routine – medical reason for operating this admission – a) within 72 hours from angiography if on | 1 = Elective<br>2 = Urgent<br>3 = Emergency |

| Sl | Variable name (field) | Variable definition | Format /codes |
|---|---|---|---|
| | | the same admission that angiography was performed (in this case, "same admission" includes the situation when angiography is performed at another hospital and the patient is transferred directly to the hospital where surgery is to be performed) OR b) within 72 hours after an unplanned admission (in a patient who had a previous angiogram and was scheduled for surgery but was admitted acutely).<br>3. Emergency: Unscheduled surgery required in next available theatre on same day due to refractory angina or cardiac compromise<br>4. Salvage: The patient is undergoing CPR en route to the operating room, that is, prior to surgical incision. | 4 = Salvage |
| 44 | Procedure type (TP) | 1. Isolated CABG<br>2. Valve surgery<br>3. Valve + CABG<br>4. Others | 1 = Isolated CABG<br>2 = Valve<br>3 = Valve+CABG<br>4 = Others |
| 45 | Mortality – date (MORT_D) | Provide date of death in hospital during the index admission at any time after the procedure, or death after discharge from hospital within thirty days of the procedure. (Before system date) | DD/MM/YYYY |
| 46 | MORTALITY – LOCATION (MORT_L) | Specify the patient location at time of death:<br>1. Operating Room: (OR)<br>2. Hospital in which operation performed: (Other than Operating Room)<br>3. Home: (Including Hospital in the Home)<br>4. Other Care Facility | 1 = OR<br>2 = Hospital<br>3 = Home<br>4 = Other Facility |
| 47 | Mortality within 30 days of surgery (MORT30) | Specify whether the patient died within 30 days after the procedure was performed. (Date of surgery counts as day 0; calculated from MORT_DDOP) | 1 = Yes<br>0 = No |

## Appendix 2.4: Exemption from ethical review

**Monash University Human Research Ethics Committee (MUHREC)**
Research Office

10 April 2014

Dear Researchers

| | |
|---|---|
| **Project Number:** | **CF14/1117 - 2014000476** |
| **Project Title:** | **Predicting cardiac surgery outcome in Australia and New Zealand patient cohort** |
| **Chief Investigator:** | **Dr Baki Billah** |

The above application has been reviewed by the Chairs of the Monash University Human Research Ethics Committee (MUHREC) who determined that the proposal satisfies section 5.1.22 of the National Statement on Ethical Conduct in Human Research.

Therefore, the Committee has granted an exemption from ethical review for the research as described in your proposal.

Thank you for your assistance.

[redacted]

Professor Nip Thomson
Chair, MUHREC

cc: Prof Christopher M. Reid, Dr Andrew Cochrane, Ms Lavinia Tran, Dr Md Nazmul Karim

Postal – Monash University, Vic 3800, Australia
Building 3E, Room 111, Clayton Campus, Wellington Road, Clayton
Telephone [redacted] Facsimile +61 3 9905 3831
Email [redacted] www.monash.edu/research/ethics/human/index/html
ABN 12 377 614 012 CRICOS Provider #00008C

## Appendix 2.5: Monash University ethics approval - HREC 2008000065

MONASH University

**Standing Committee on Ethics in Research Involving Humans (SCERH)**
Research Office

### Human Ethics Certificate of Approval

| | |
|---|---|
| **Date:** | **30 January 2008** |
| **Project Number:** | **CF08/0322 - 2008000065** |
| **Project Title:** | **The ASCTS Cardiac Surgery Registry** |
| **Chief Investigator:** | **Assoc Prof Chris Reid** |
| **Approved:** | **From 30 January 2008 to 30 January 2013** |

---

**Terms of approval**

1. Approval is only valid whilst you hold a position at Monash University.
2. It is the responsibility of the Chief Investigator to ensure that all pending information (such as permission letters from organisations) is forwarded to SCERH. Research cannot begin at an organisation until SCERH receives a permission letter from that organisation.
3. It is the responsibility of the Chief Investigator to ensure that all investigators are aware of the terms of approval and to ensure the project is conducted as approved by SCERH.
4. You should notify SCERH immediately of any serious or unexpected adverse effects on participants or unforeseen events affecting the ethical acceptability of the project.
5. The Explanatory Statement must be on Monash University letterhead and the Monash University complaints clause must contain your project number.
6. **Amendments to the approved project:** Requires the submission of a Request for Amendment form to SCERH and must not begin without written approval from SCERH. Substantial variations may require a new application.
7. **Future correspondence:** Please quote the project number and project title above in any further correspondence.
8. **Annual reports:** Continued approval of this project is dependent on the submission of an Annual Report. This is determined by the date of your letter of approval.
9. **Final report:** A Final Report should be provided at the conclusion of the project. SCERH should be notified if the project is discontinued before the expected date of completion.
10. **Monitoring:** Projects may be subject to an audit or any other form of monitoring by SCERH at any time.
11. **Retention and storage of data:** The Chief Investigator is responsible for the storage and retention of original data pertaining to a project for a minimum period of five years.

[redacted]

Dr Souheir Houssami
Executive Officer, Human Research Ethics (on behalf of SCERH)

Cc: Dr Hugh David Wolfunden; Dr Diem Dinh

Postal – Monash University, Vic 3800, Australia
Building 3E, Room 111, Clayton Campus, Wellington Road, Clayton
Telephone [redacted] Facsimile +61 3 9905 1420
Email [redacted] www.monash.edu/research/ethics/human/index/html
ABN 12 377 614 012 CRICOS Provider #00008C

## Appendix 2.6 MUHREC extension

**MRO Human Ethics Team** [redacted] 24/09/2013

to me

Dear Researchers

**CF08/0322 – 2008000065: The ASCTS Cardiac Surgical Registry**

Thank you for the Annual Report / Request for Extension form provided in relation to the above project.

This is to advise that the Monash University Human Research Ethics Committee (MUHREC) has noted the comments that you made on the form and research is approved until 31 January 2018.

Please submit a Final Report by 31 January 2018.
To continue with human data collection after 31 January 2018 you will need a new submission to MUHREC. Please ensure that you use the latest version of the application forms which are available on our website http://www.monash.edu.au/researchoffice/human/reports-extensions.html

Thank you for your assistance.
Professor Ben Canny
Chair, MUHREC
Human Ethics
Monash Research Office
*Our aim is exceptional service*

Monash University
Level 1, Building 3e, Clayton Campus
Wellington Rd
Clayton VIC 3800, Australia

Telephone: [redacted]
Email: [redacted]
Website: http://www.monash.edu.au/researchoffice/human
ABN 12 377 614 012 CRICOS Provider No 00008C

## Appendix 2.7: Alfred hospital Ethics approval

TheAlfred

### ETHICS COMMITTEE CERTIFICATE OF APPROVAL

*This is to certify that*

**Project No:** 262/09

**Project Title:** The Australasian Society of Cardiac and Thoracic Surgeons National Cardiac Surgery Registry

**Principal Researcher:** A/Professor Silvana Marasco

**Participant Information and Consent Form version:** *3* **dated:** *02-Nov-2009* **&**

**Information Sheet for Next of Kin version:** *3* **dated:** *02-Nov-2009*

*was considered by the Ethics Committee on* ***22-Oct-2009*** *and* ***APPROVED*** *on* **10-Nov-2009**

It is the Principal Researcher's responsibility to ensure that all researchers associated with this project are aware of the conditions of approval and which documents have been approved.

***The Principal Researcher is required to notify the Secretary of the Ethics Committee, via amendment or progress report, of***

- Any significant change to the project and the reason for that change, including an indication of ethical implications (if any);
- Serious adverse effects on participants and the action taken to address those effects;
- Any other unforeseen events or unexpected developments that merit notification;
- The inability of the Principal Researcher to continue in that role, or any other change in research personnel involved in the project;
- Any expiry of the insurance coverage provided with respect to sponsored clinical trials and proof of re-insurance;
- A delay of more than 12 months in the commencement of the project; and,
- Termination or closure of the project.

***Additionally, the Principal Researcher is required to submit***

- A Progress Report on the anniversary of approval and on completion of the project *(forms to be provided)*;

The Ethics Committee may conduct an audit at any time.

All research subject to the Alfred Hospital Ethics Committee review must be conducted in accordance with the National Statement on Ethical Conduct in Human Research (2007).

The Alfred Hospital Ethics Committee is a properly constituted Human Research Ethics Committee in accordance with the National Statement on Ethical Conduct in Human Research (2007).

**SPECIAL CONDITIONS**
*The data being transmitted to a centralised database will not contain identifiable information of individual surgeons and referring cardiologists.*

[redacted]

**SIGNED:**
*Chair, Ethics Committee (or delegate)*

**R. FREW**
**SECRETARY**
**ETHICS COMMITTEE**

***Please quote Project No and Title in all correspondence***

## Appendix 3.1: Search Strategy: Ovid MEDLINE(R) 1946 till January 24, 2017

| # | Searches | Results |
|---|---|---|
| 1 | ((operat$ or post operat$ or postoperat$) adj (death* or mortality*)).mp. | 22798 |
| 2 | ((hospital or in hospital or in-hospital or in hospital or short-term or short term) adj (mortality or death*)).mp. | 46090 |
| 3 | (30-day mortality or 30 day mortality).mp. | 8738 |
| 4 | 1 or 2 or 3 | 73392 |
| 5 | Validat$.mp. or Predict$.ti. or Rule$.mp. or (Predict$ and (Outcome$ or Risk$ or Model$)).mp. or ((History or Variable$ or Criteria or Scor$ or Characteristic$ or Finding$ or Factor$) and (Predict$ or Model$ or Decision$ or Identif$ or Prognos$)).mp. or (Decision$.mp. and ((Model$ or Clinical$).mp. or Logistic Models/)) or (Prognostic and (History or Variable$ or Criteria or Scor$ or Characteristic$ or Finding$ or Factor$ or Model$)).mp. | 3289690 |
| 6 | models, statistical/ or risk assessment/ | 287389 |
| 7 | ((risk or clinical) adj (predict* or stratification*) adj (model* or scor* or algorithm* or index* or tool* or rule*)).mp. | 3843 |
| 8 | ((predict* or stratification*) adj (model* or scor* or algorithm* or index* or tool* or rule*)).mp. | 27907 |
| 9 | ((risk or clinical or prognostic) adj (model* or scor* or algorithm* or index* or tool* or rule*)).mp. | 42503 |
| 10 | (risk adj (predict* or stratification*)).mp. | 22338 |
| 11 | 5 or 6 or 8 or 9 or 10 | 3448496 |
| 12 | cardiac surgical procedures/ or myocardial revascularization/ or coronary artery bypass/ | 97477 |
| 13 | ((cardiac* or cardio-thoracic or cardio thoracic or cardiothoracic or heart) adj (operation* or surg*)).mp. | 73805 |
| 14 | (CABG or Coronary Artery Bypass or Coronary Artery Bypass grafting).mp. | 56809 |
| 15 | 12 or 13 or 14 | 127425 |
| 16 | 4 and 11 and 15 | 5433 |
| 17 | limit 16 to (English language and full text and humans and "all adult (19 plus years)") | 563 |

## Appendix 3.2: Search Strategy: EMBASE 1974 till January 24, 2017

| # | Searches | Results |
|---|---|---|
| 1 | ((operat$ or post operat$ or postoperat$) adj (death* or mortality*)).mp. | 31411 |
| 2 | ((hospital or in hospital or in-hospital or in hospital or short-term or short term) adj (mortality or death*)).mp. | 53028 |
| 3 | (30-day mortality or 30 day mortality).mp. | 16998 |
| 4 | 1 or 2 or 3 | 97082 |
| 5 | Validat$.mp. or Predict$.ti. or Rule$.mp. or (Predict$ and (Outcome$ or Risk$ or Model$)).mp. or ((History or Variable$ or Criteria or Scor$ or Characteristic$ or Finding$ or Factor$) and (Predict$ or Model$ or Decision$ or Identif$ or Prognos$)).mp. or (Decision$.mp. and ((Model$ or Clinical$).mp. or Logistic Models/)) or (Prognostic and (History or Variable$ or Criteria or Scor$ or Characteristic$ or Finding$ or Factor$ or Model$)).mp. | 4783472 |
| 6 | models, statistical/ or risk assessment/ | 560042 |
| 7 | ((risk or clinical) adj (predict* or stratification*) adj (model* or scor* or algorithm* or index* or tool* or rule*)).mp. | 7737 |
| 8 | ((predict* or stratification*) adj (model* or scor* or algorithm* or index* or tool* or rule*)).mp. | 48867 |
| 9 | ((risk or clinical or prognostic) adj (model* or scor* or algorithm* or index* or tool* or rule*)).mp. | 82948 |
| 10 | (risk adj (predict* or stratification*)).mp. | 45782 |
| 11 | 5 or 6 or 7 or 8 or 9 or 10 | 5105414 |
| 12 | cardiac surgical procedures/ or myocardial revascularization/ or coronary artery bypass/ | 157741 |
| 13 | ((cardiac* or cardio-thoracic or cardio thoracic or cardiothoracic or heart) adj (operation* or surg*)).mp. | 117768 |
| 14 | (CABG or Coronary Artery Bypass or Coronary Artery Bypass grafting).mp. | 92011 |
| 15 | 12 or 13 or 14 | 208058 |
| 16 | 4 and 11 and 15 | 7189 |
| 17 | limit 16 to (full text and human and English language) | 965 |
| 18 | limit 17 to (adult <18 to 64 years> or aged <65+ years>) | 560 |

## Appendix 3.3: CHARMS checklist for Systematic Reviews of Prediction Modelling Studies

CHARMS 2014 Relevant items to extract from individual studies in a systematic review of prediction models

| Domain | Key items | Reported on page # |
|---|---|---|
| SOURCE OF DATA | Source of data (e.g., cohort, case-control, randomized trial participants, or registry data) | |
| PARTICIPANTS | Participant eligibility and recruitment method (e.g., consecutive participants, location, number of centers, setting, inclusion and exclusion criteria) | |
| | Participant description | |
| | Details of treatments received, if relevant | |
| | Study dates | |
| OUTCOME(S) TO BE PREDICTED | Definition and method for measurement of outcome | |
| | Was the same outcome definition (and method for measurement) used in all patients? | |
| | Type of outcome (e.g., single or combined endpoints) | |
| | Was the outcome assessed without knowledge of the candidate predictors (i.e., blinded)? | |
| | Were candidate predictors part of the outcome (e.g., in panel or consensus diagnosis)? | |
| | Time of outcome occurrence or summary of duration of follow-up | |
| CANDIDATE PREDICTORS (OR INDEX TESTS) | Number and type of predictors (e.g., demographics, patient history, physical examination, additional testing, disease characteristics) | |
| | Definition and method for measurement of candidate predictors | |
| | Timing of predictor measurement (e.g., at patient presentation, at diagnosis, at treatment initiation) | |
| | Were predictors assessed blinded for outcome, and for each other (if relevant)? | |
| | Handling of predictors in the modelling (e.g., continuous, linear, non-linear transformations or categorised) | |
| SAMPLE SIZE | Number of participants and number of outcomes/events | |
| | Number of outcomes/events in relation to the number of candidate predictors (Events Per Variable) | |
| MISSING DATA | Number of participants with any missing value (include predictors and outcomes) | |
| | Number of participants with missing data for each predictor | |
| | Handling of missing data (e.g., complete-case analysis, imputation, or other methods) | |
| MODEL DEVELOPMENT | Modelling method (e.g., logistic, survival, neural network, or machine learning techniques) | |
| | Modelling assumptions satisfied | |
| | Method for selection of predictors for inclusion in multivariable modelling (e.g., all candidate predictors, pre-selection based on unadjusted association with the outcome) | |
| | Method for selection of predictors during multivariable modelling (e.g., full model approach, backward or forward selection) and criteria used (e.g., p-value, Akaike Information Criterion) | |
| | Shrinkage of predictor weights or regression coefficients (e.g., no shrinkage, uniform shrinkage, penalized estimation) | |
| MODEL PERFORMANCE | Calibration (calibration plot, calibration slope, Hosmer-Lemeshow test) and Discrimination (C-statistic, D-statistic, log-rank) measures with confidence intervals | |
| | Classification measures (e.g., sensitivity, specificity, predictive values, net reclassification improvement) and whether a-priori cut points were used | |
| MODEL EVALUATION | Method used for testing model performance: development dataset only (random split of data, resampling methods e.g. bootstrap or cross-validation, none) or separate external validation (e.g. temporal, geographical, different setting, different investigators) | |
| | In case of poor validation, whether model was adjusted or updated (e.g., intercept recalibrated, predictor effects adjusted, or new predictors added) | |
| RESULTS | Final and other multivariable models (e.g., basic, extended, simplified) presented, including predictor weights or regression coefficients, intercept, baseline survival, model performance measures (with standard errors or confidence intervals) | |
| | Any alternative presentation of the final prediction models, e.g., sum score, nomogram, score chart, predictions for specific risk subgroups with performance | |
| | Comparison of the distribution of predictors (including missing data) for development and validation datasets | |
| INTERPRETATION AND DISCUSSION | Interpretation of presented models (confirmatory, i.e., model useful for practice versus exploratory, i.e., more research needed) | |
| | Comparison with other studies, discussion of generalizability, strengths and limitations. | |

## Appendix 3.4: Data extraction checklist

| Domain and items | Finding | Location |
|---|---|---|
| 1 . Source of data | | |
| Research setting (1. Administrative 2. Study 3. Registry) | | |
| Duration of data collection (year-year) | | |
| 2. Participants | | |
| Population (Country) | | |
| Number of centers | | |
| Recruitment method Inclusion and exclusion criteria reported (1. Yes 2. No) | | |
| Participant description adequate (1. Yes 2. No) | | |
| 3. Outcome(s) to be predicted | | |
| Endpoints (1. Single, 2. Multiple, 3. Combined) | | |
| Definition of outcome (1. Mortality 2. Morbidity 3. Adverse event 4. Others) | | |
| Method for measurement (1. Prospective follow-up 2. Data linkage) | | |
| Same measurement of outcome for all patients (1. Yes 2. No) | | |
| Blinding patients of outcome (1. Yes 2. No) | | |
| Time of outcome occurrence (1. Short term 2. Long-term) | | |
| 4. Candidate predictors | | |
| Type of predictors (1. Demography 2. preoperative 3. Interoperative) | | |
| Number of predictors | | |
| Timing of measurement (1. preoperative 2. Interoperative 3. perioperative) | | |
| Detailed definition of predictors reported (1. Yes 2. No) | | |
| Transformations of continuous predictors in the modelling (1. Yes 2. No) | | |
| 5. Sample size | | |
| Derivation sample | | |
| Validation sample | | |
| Events per participant | | |
| Events Per predictor | | |
| 6. Missing data | | |
| Percentage of participants with any missing value | | |
| Imputation (1. None, 2. Single imputation 3. Multiple imputation) | | |
| 7. Model development | | |
| Predictor pooling method for model (1. All predictors 2. Association with outcome) | | |
| Modelling method (1. Logistic 2. Survival 3. Bayes 4. Machine learning techniques) | | |
| Modelling assumptions satisfied (1. Yes 2. No) | | |
| Multivariable model building (1. Full model approach 2. Automated selection) | | |
| Criteria used for model building (1.p-value 2. Information criteria 3. others) | | |
| 8. Model performance | | |
| Calibration reported (1. None 2. Derivation sample 3. Validation sample 4. Others) | | |
| Calibration type (1. Calibration plot 2. Calibration slope 3. Hosmer Lemeshow test) | | |
| Discrimination reported (1. None 2. Derivation sample 3. Validation sample) | | |
| Discrimination type (1. C-statistic 2.D-statistic 3.log-rank 4. Others) | | |
| 9. Model evaluation | | |
| Testing model performance (1. Internal validation, 2. External validation 3. Both) | | |
| Internal validation (1. Random data split, 2. Resampling methods 3. Others) | | |
| External validation (1. Temporal 2. Geographical 3. Different setting) | | |
| 10. Results | | |
| Presentation of final model (1. Basic 2. Extended 2. Simplified) | | |
| Presented Regression coefficients and CI/SE (1. Yes 2. No) | | |
| Alternative presentation (1. None 2. Nomogram 3. Score chart 4. Calculator) | | |
| Compared predictor distribution in development vs validation data (1. Yes 2. No) | | |
| 11. Interpretation and Discussion | | |
| Interpretation of presented models (1. Confirmatory 2. Exploratory) | | |
| Discussion of generalizability (1. Yes 2. No) | | |
| Discussion of strengths and limitations (1. Yes 2. No) | | |

## Appendix 4.1: Percentage of missing values in ANZSCTS database

| Variables | Missing data | |
|---|---|---|
| | Frequency | Percent |
| Gender | 0 | 0 |
| Previous cardiac surgery | 0 | 0 |
| Clinical status | 0 | 0 |
| Pacemaker in situ | 0 | 0 |
| Age | 8 | 0.02 |
| BMI | 35 | 0.08 |
| Previous MI | 40 | 0.09 |
| Hypertension | 47 | 0.10 |
| Respiratory disease | 47 | 0.10 |
| Cardiogenic shock | 48 | 0.10 |
| Resuscitation | 47 | 0.10 |
| Immunosuppressant use | 47 | 0.10 |
| Hypercholesterolemia | 50 | 0.11 |
| Cerebrovascular disease | 49 | 0.11 |
| Peripheral vascular disease | 49 | 0.11 |
| Inotrope use | 50 | 0.11 |
| IV-Nitrates | 50 | 0.11 |
| Steroids use | 52 | 0.11 |
| Arrhythmia | 58 | 0.12 |
| Anticoagulant use | 55 | 0.12 |
| Congestive Heart Failure | 59 | 0.13 |
| Number of diseased vessels | 68 | 0.15 |
| Left main disease | 74 | 0.16 |
| Smoking | 91 | 0.20 |
| Diabetes mellitus | 93 | 0.20 |
| 30-day mortality | 162 | 0.35 |
| Angina | 179 | 0.38 |
| Renal dysfunction | 618 | 1.33 |
| Ejection fraction | 1,015 | 2.18 |
| NYHA class | 1,747 | 3.75 |
| Family History of CAD | 5,014 | 10.77 |

## Appendix 4.2: Bootstrap coverage of the proposed predictors

| Variables | Bootstrap coverage |
|---|---|
| Age | 100.00 |
| Peripheral vascular disease | 100.00 |
| Respiratory disease | 100.00 |
| Ejection fraction | 100.00 |
| Renal dysfunction | 100.00 |
| Smoking | 100.00 |
| Arrhythmia | 99.85 |
| Diabetes mellitus | 99.54 |
| Hypercholesterolemia | 99.48 |
| Cerebrovascular disease | 99.22 |
| Hypertension | 99.00 |
| Congestive Heart Failure | 96.38 |
| Steroids use | 89.00 |
| Angina | 87.03 |
| Family History of CAD | 80.27 |
| NYHA class | 77.82 |
| Previous MI | 76.27 |
| Left main disease | 62.19 |
| BMI | 54.97 |
| Inotrope use | 36.46 |
| Previous cardiac surgery | 23.75 |
| Resuscitation | 19.97 |
| Immunosuppressant use | 16.55 |
| Number of diseased vessels | 15.60 |
| Pacemaker in situ | 12.98 |
| IV-Nitrates | 11.76 |
| Anticoagulant use | 8.60 |
| Cardiogenic shock | 7.41 |
| Clinical status | 7.22 |
| Gender | 6.39 |

## Appendix 4.3: Descriptive statistics of study population (n=46,573)

| Variables | n | % |
|---|---|---|
| Age (Mean ± sd year) | 65.9 (10.4) | |
| Gender | | |
| Male | 36,960 | 79.36 |
| Female | 9,613 | 20.64 |
| BMI (Kg/$m^2$) | | |
| 18.5- 25 | 10,654 | 22.89 |
| < 18.5 | 181 | 0.39 |
| 25 - 30 | 19,524 | 41.95 |
| > 30 | 16,179 | 34.77 |
| Family history of CAD | | |
| No | 24,059 | 51.66 |
| Yes | 17,500 | 37.58 |
| Smoking | | |
| No | 15,888 | 34.11 |
| Yes | 30,594 | 65.69 |
| Diabetes mellitus (DM) | | |
| No DM | 30,299 | 65.06 |
| DM no drug | 2,590 | 5.56 |
| DM on oral drug | 9,057 | 19.45 |
| DM on insulin | 4,534 | 9.74 |
| Hypercholesterolemia | | |
| No | 8,846 | 18.99 |
| Yes | 37,677 | 80.9 |
| Hypertension | | |
| No | 9,457 | 20.31 |
| Yes | 37,069 | 79.59 |
| Cerebrovascular disease | | |
| No | 41,757 | 89.66 |
| Yes | 4,767 | 10.24 |
| Peripheral vascular disease | | |
| No | 41,161 | 88.38 |
| Yes | 5,363 | 11.52 |
| Respiratory disease | | |
| No | 40,846 | 87.7 |
| Yes | 5,680 | 12.2 |
| Previous MI | | |
| No | 21,675 | 46.54 |
| Yes | 24,858 | 53.37 |
| Angina | | |
| No angina | 6,132 | 13.17 |
| Stable angina | 28,708 | 61.64 |
| Unstable angina | 11,554 | 24.81 |
| CHF | | |
| No CHF | 40,220 | 86.36 |
| Past CHF | 3,529 | 7.58 |
| Current CHF | 2,765 | 5.94 |
| NYHA class | | |
| Class I & II | 35,935 | 77.16 |
| Class III | 6,967 | 14.96 |
| Class IV | 1,924 | 4.13 |
| Cardiogenic shock | | |
| No | 45,792 | 98.32 |
| Yes | 733 | 1.57 |
| Resuscitation | | |
| No | 46,110 | 99.01 |
| Yes | 416 | 0.89 |
| Arrhythmia | | |

| Variables | n | % |
|---|---|---|
| No | 42,310 | 90.85 |
| Yes | 4,205 | 9.03 |
| Pacemaker | | |
| No | 46,070 | 98.92 |
| Yes | 503 | 1.08 |
| Previous cardiac surgery | | |
| No | 44,754 | 96.09 |
| Yes | 1,819 | 3.91 |
| Ejection fraction | | |
| >60% | 22,624 | 48.58 |
| 46% - 60% | 14,345 | 30.8 |
| 30% - 45% | 6,767 | 14.53 |
| <30% | 1,822 | 3.91 |
| Left main disease | | |
| No | 34,115 | 73.25 |
| Yes | 12,384 | 26.59 |
| Diseased vessels | | |
| One | 2,527 | 5.42 |
| Two | 10,917 | 23.44 |
| Three | 33,061 | 70.99 |
| Clinical status | | |
| Elective | 28,604 | 61.42 |
| Urgent | 16,135 | 34.64 |
| Emergency | 1,741 | 3.74 |
| Salvage | 93 | 0.2 |
| Immunosuppressant | | |
| No | 45,671 | 98.06 |
| yes | 855 | 1.84 |
| Inotrope use | | |
| No | 45,664 | 98.05 |
| Yes | 859 | 1.84 |
| IV-Nitrates | | |
| No | 43,469 | 93.34 |
| Yes | 3,054 | 6.56 |
| Anticoagulant use | | |
| No | 35,938 | 77.16 |
| Yes | 10,580 | 22.72 |
| Steroids use | | |
| No | 45,813 | 98.37 |
| Yes | 708 | 1.52 |
| Renal dysfunction | | |
| None | 10,766 | 23.12 |
| Mild | 23,464 | 50.38 |
| Moderate | 9,981 | 21.43 |
| Severe | 1,036 | 2.22 |
| On-dialysis | 708 | 1.52 |
| 30-day mortality | | |
| no | 45,746 | 98.22 |
| yes | 665 | 1.43 |

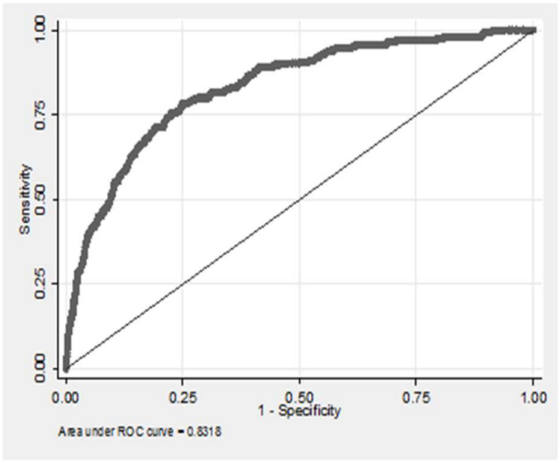
## Appendix 4.4: Univariate association of predictors with mortality in specific time interval

| Preoperative variables | 0-30 days | 31-90 days | 91-365 days | 1-3 years | > 3 years |
|---|---|---|---|---|---|
| | HR (95% CI) | HR (95% CI) | HR (95% CI) | HR (95% CI) | HR (95% CI) |
| Age | 1.06 (1.05, 1.07) | 1.07 (1.05, 1.08) | 1.07 (1.06, 1.08) | 1.05 (1.05, 1.07) | 1.07 (1.07, 1.08) |
| Peripheral vascular disease | 2.51 (2.05, 3.08) | 2.69 (2.02, 3.59) | 3.11 (2.54, 3.82) | 2.71 (2.38, 3.11) | 2.36 (2.17, 2.57) |
| Respiratory disease | 1.53 (1.22, 1.92) | 3.12 (2.38, 4.10) | 1.91 (1.52, 2.40) | 2.03 (1.76, 2.33) | 1.77 (1.62, 1.94) |
| Ejection fraction: 46-60% | 1.55 (1.21, 2.00) | 1.65 (1.17, 2.33) | 1.82 (1.43, 2.31) | 1.67 (1.44, 1.93) | 1.32 (1.22, 1.44) |
| Ejection fraction: 30-45% | 3.37 (2.64, 4.31) | 3.61 (2.57, 5.08) | 3.69 (2.89, 4.72) | 2.66 (2.27, 3.11) | 1.86 (1.70, 2.05) |
| Ejection fraction: <30% | 11.2 (8.70, 14,5) | 7.82 (5.24, 11.6) | 5.17 (3.70, 7.24) | 3.55 (2.82, 4.47) | 2.51 (2.18, 2.88) |
| Renal dysfunction: mild | 2.66 (6.15, 17.3) | 2.56 (1.48, 4.43) | 1.64 (1.18, 2.29) | 1.74 (1.44, 2.11) | 1.83 (1.62, 2.07) |
| Renal dysfunction: | 6.08 (4.27, 8.67) | 6.91 (4.01, 11.9) | 4.44 (3.20, 6.16) | 3.20 (2.62, 3.91) | 3.70 (3.27, 4.19) |
| Renal dysfunction: severe | 11.0 (6.95, 17.4) | 21.4 (11.5, 39.8) | 7.31 (4.58, 11.7) | 6.31 (4.71, 8.45) | 6.07 (5.01, 7.35) |
| On dialysis | 10.3 (6.15, 17.2) | 23.4 (12.2, 45.2) | 11.4 (7.20, 18.0) | 8.68 (6.42, 11.7) | 6.95 (5.48, 8.81) |
| Smoking | 0.88 (0.74, 1.06) | 1.36 (1.03, 1.80) | 1.52 (1.23, 1.88) | 1.26 (1.11, 1,43) | 1.20 (1.12, 1.30) |
| Arrhythmia | 3.26 (2.66, 4.00) | 4.46 (3.39, 5.86) | 2.97 (2.38, 3.71) | 2.13 (1.82, 2.50) | 1.93 (1.75, 2.14) |
| Diabetes: no treatment | 1.14 (0.78, 1.69) | 1.00 (0.54, 1.85) | 1.36 (0.92, 1.99) | 0.96 (0.72, 1.26) | 1.20 (1.04, 1.38) |
| Diabetes: oral drug | 1.37 (1.11, 1.70) | 1.56 (1.15, 2.12) | 1.52 (1.22, 1.90) | 1.33 (1.15, 1.54) | 1.32 (1.21, 1.45) |
| Diabetes: on insulin | 1.76 (1.36, 2.27) | 2.36 (1.68, 3.30) | 1.68 (1.27, 2.22) | 1.77 (1.50, 2.10) | 1.84 (1.65, 2.05) |
| Hypercholesterolemia | 0.81 (0.66, 1.00) | 1.14 (0.82, 1.59) | 0.92 (0.73,1.16) | 0.86 (0.75, 0.99) | 0.76 (0.70, 0.83) |
| Cerebrovascular disease | 1.93 (1.53, 2.42) | 3.13 (2.35, 4.16) | 1.86 (1.45, 2.38) | 2.30 (1.98, 2.66) | 1.97 (1.80, 2.16) |
| Hypertension | 1.57 (1.23, 2.03) | 2.07 (1.40, 3.09) | 1.66 (1.26, 2.17) | 1.43 (1.22, 1.69) | 1.48 (1.35, 1.62) |
| CHF: old | 2.11 (1.59, 2.80) | 2.39 (1.65, 3.46) | 3.04 (2.37, 3.89) | 1.85 (1.54, 2.22) | 1.85 (1.69, 2.04) |
| CHF: Current | 6.75 (5.52, 8.23) | 4.51 (3.27, 6.22) | 3.41 (2.62, 4.45) | 2.56 (2.14, 3.07) | 2.33 (2.07, 2.61) |
| Steroids use | 1.57 (0.88, 2.79) | 3.67 (2.01, 6.42) | 2.11 (1.24, 3.59) | 3.53 (2.69, 4.63) | 1.91 (1.51, 2.40) |
| Stable angina | 0.75 (0.55, 1.01) | 0.53 (0.37 0.75) | 0.68 (0.52, 0.88) | 0.72 (0.61, 0.85) | 0.81 (0.71, 0.91) |
| Unstable angina | 2.56 (1.91, 3.43) | 1.11 (0.78, 1.60) | 0.97 (0.73, 1.28) | 0.96 (0.80, 1.15) | 0.91 (0.80, 1.03) |
| Family History of CAD | 0.80 (0.65, 0.97) | 0.63 (0.46, 0.84) | 0.56 (0.45, 0.69) | 0.67 (0.59, 0.76) | 0.69 (0.64, 0.75) |
| NYHA class III | 2.37 (1.91, 2.95) | 2.32 (1.73, 3.12) | 1.88 (1.51, 2.36) | 1.69 (1.46, 1.96) | 1.54 (1.41, 1.67) |
| NYHA class IV | 8.42 (6.75, 10.5) | 3.94 (2.63, 5.90) | 2.31 (1.62, 3.30) | 2.48 (1.99, 3.09) | 1.61 (1.42, 1.83) |
| Previous MI | 2.63 (2.15, 3.22) | 2.01 (1.54, 2.64) | 1.76 (1.45, 2.14) | 1.45 (1.29, 1.64) | 1.41 (1.31, 1.51) |
| Left main disease | 1.65 (1.38, 1.97) | 1.36 (1.05, 1.78) | 1.42 (1.17, 1,72) | 1.18 (1.04, 1.34) | 1.25 (1.16, 1.35) |
| Underweight | 1.90 (0.78, 4.63) | 1.50 (0.37, 6.12) | 2.47 (1.09, 5.58) | 2.58 (1.48, 4.50) | 1.80 (1.13, 2.87) |
| Overweight | 0.61 (0.49, 0.75) | 0.62 (0.46, 0.76) | 0.59 (0.47, 0.73) | 0.69 (0.59, 0 .80) | 0.77 (0.70, 0.84) |
| Obese | 0.69 (0.56, 0.86) | 0.55 (0.40, 0.76) | 0.56 (0.44, 0.71) | 0.77 (0.66, 0.88) | 0.81 (0.74, 0.88) |
| Inotrope use | 9.52 (7.45, 12.2) | 4.93 (3.09, 7.88) | 1.90 (1.11, 3.23) | 1.42 (0.96, 2.10) | 1.29 (0.97, 1.71) |
| Previous cardiac surgery | 2.11 (1.53, 2.92) | 1.86 (1.14, 3.04) | 2.04 (1.44, 2.91) | 1.18 (0.88, 1.56) | 1.16 (0.99, 1.35) |
| Resuscitation | 10.8 (7.91, 14.8) | 4.21 (2.08, 8.52) | 2.87 (1.53, 5.38) | 1.03 (0.53, 1.98) | 1.06 (0.69, 1.62) |
| Immunosuppressant use | 1.75 (1.06, 2.88) | 2.53 (1.38, 4.64) | 1.99 (1.21, 3.28) | 2.59 (1.95, 3.44) | 1.87 (1.51, 2.31) |
| One diseased vessels | 0.32 (0.11, 0.98) | 0.37 (0.04, 3.28) | 0.45 (0.13, 1.58) | 3.29 (0.45, 24.0) | 0.74 (0.23, 2.33) |
| Two diseased vessels | 0.43 (0.16, 1.17) | 0.99 (0.14, 7.19) | 0.51 (0.16, 1.62) | 3.91 (0.54, 27.9) | 1.18 (0.38,3.69) |
| Three Diseased vessel | 0.43 (0.24, 1.75) | 1.23 (0.17, 8.78) | 0.77 (0.24, 2.38) | 5.81 (0.81, 41.3) | 1.53 (0.49, 4.75) |
| Pacemaker in situ | 2.82 (1.69, 4.72) | 2.77 (1.31, 5.88) | 1.93 (1.00, 3.74) | 1.74 (1.12, 2.71) | 1.91 (1.08, 3.37) |
| IV-Nitrates | 4.24 (3.44, 5.21) | 2.29 (1.59, 3.30) | 1.12 (0.78, 1.61) | 1.22 (0.97, 1.52) | 1.08 (0.96, 1,22) |
| Anticoagulant use | 2.68 (2.24,3.19) | 1.76 (1.35, 2.29) | 1.25 (1.00, 1.54) | 1.22(1.07, 1.40) | 1.13 (1.05, 1.23) |
| Cardiogenic shock | 12.4 (9.73, 15.7) | 7.13 (4.60, 11.0) | 2.17 (1.25, 3.78) | 2.26 (1.60, 3.19) | 1.21 (0.91, 1.62) |
| Clinical status: urgent | 2.16 (1.77, 2.64) | 1.31 (1.00, 1.72) | 1.30 (1.07, 1.57) | 1.18 (1.05, 1.34) | 1.10 (1.02, 1.18) |
| Clinical status: emergency | 10.7 (8.42, 13.6) | 4.31 (2.89, 6.43) | 1.43 (0.90, 2.25) | 1.43 (1.08, 1.90) | 1.30 (1.10, 1.53) |
| Clinical status: salvage | 25.9 (14.8, 45.7) | 9.40 (2.99, 29.5) | 6.46 (2.40, 17.4) | 1.92 (0.61, 5.98) | 0.50 (0.16, 1.55) |
| Gender | 1.80 (1.49, 2.17) | 1.80 (1.38, 2.36) | 1.15 (0.93, 1.44) | 1.09 (0.95, 1.26) | 1.18 (1.09, 1.28) |

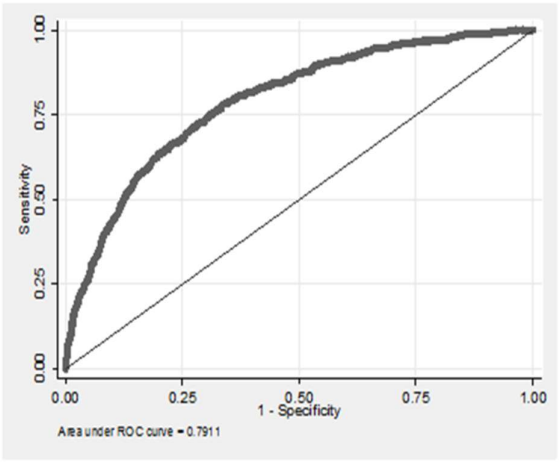
## Appendix 4.5: Baseline hazard of coefficient and Standard error of for models

| Splines | Coefficient | Standard Error |
|---|---|---|
| 3-month model | | |
| rcs1 | 0.0989 | 0.0069 |
| rcs2 | 0.0105 | 0.0041 |
| rcs3 | -0.0003 | 0.0019 |
| 1-year model | | |
| rcs1 | 0.1398 | 0.0073 |
| rcs2 | 0.0092 | 0.0051 |
| rcs3 | -0.0020 | 0.0019 |
| 3-year model | | |
| rcs1 | 0.2231 | 0.0075 |
| rcs2 | -0.0042 | 0.0053 |
| rcs3 | 0.0003 | 0.0023 |
| > 3-year model | | |
| rcs1 | 1.3198 | 0.0246 |
| rcs2 | -0.0936 | 0.0201 |
| rcs3 | -0.0633 | 0.0095 |

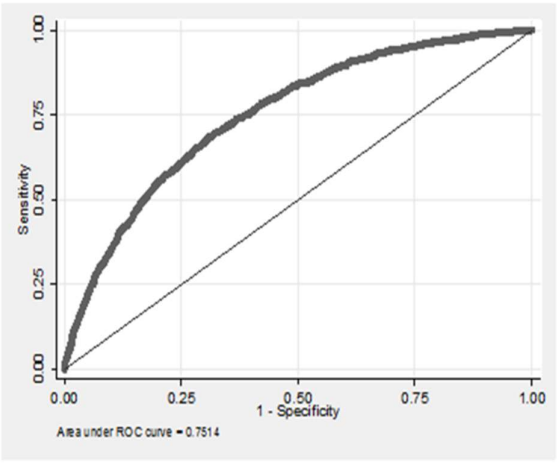
## Appendix 4.6: Model discrimination curves for the four time intervals in the combined dataset.



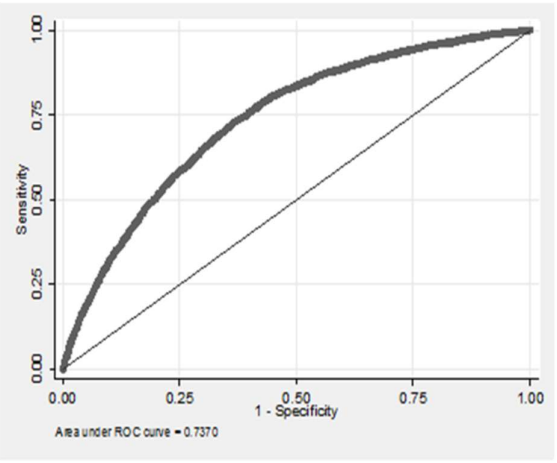

3a. Discrimination of 90-day model in the combined dataset




3b. Discrimination of 1-year model in the combined dataset




3c. Discrimination of 3-year model in the combined dataset




3d. Discrimination of >3-year model in the combined dataset

## Appendix 4.7: TRIPOD Checklist: Prediction Model Development

| Section/Topic | | Checklist Item | Page |
|---|---|---|---|
| **Title and abstract** | | | |
| Title | 1 | Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted. | 1 |
| Abstract | 2 | Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions. | 2 |
| **Introduction** | | | |
| Background and objectives | 3a | Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models. | 3 |
| | 3b | Specify the objectives, including whether the study describes the development or validation of the model or both. | 3 |
| **Methods** | | | |
| Source of data | 4a | 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. | 4 |
| | 4b | Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up. | 4 |
| Participants | 5a | Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres. | 4 |
| | 5b | Describe eligibility criteria for participants. | 4 |
| | 5c | Give details of treatments received, if relevant. | 4 |
| Outcome | 6a | Clearly define the outcome that is predicted by the prediction model, including how and when assessed. | 4 |
| | 6b | Report any actions to blind assessment of the outcome to be predicted. | - |
| Predictors | 7a | Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured. | Sup Table 1 |
| | 7b | Report any actions to blind assessment of predictors for the outcome and other predictors. | - |
| Sample size | 8 | Explain how the study size was arrived at. | 4 |
| Missing data | 9 | Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method. | 4 |
| Statistical analysis methods | 10a | Describe how predictors were handled in the analyses. | 5-6 |
| | 10b | Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation. | 6 |
| | 10d | Specify all measures used to assess model performance and, if relevant, to compare multiple models. | 6-7 |
| Risk groups | 11 | Provide details on how risk groups were created, if done. | - |
| **Results** | | | |
| Participants | 13a | 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. | 8 |
| | 13b | Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome. | 8, 16 |
| Model development | 14a | Specify the number of participants and outcome events in each analysis. | Table 2 |
| | 14b | If done, report the unadjusted association between each candidate predictor and outcome. | 5 |
| Model specification | 15a | 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). | 9, Table 2& Sup Table 5 |
| | 15b | Explain how to the use the prediction model. | 13 |
| Model performance | 16 | Report performance measures (with CIs) for the prediction model. | 9-10 |
| **Discussion** | | | |
| Limitations | 18 | Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data). | 13 |
| Interpretation | 19b | Give an overall interpretation of the results, considering objectives, limitations, and results from similar studies, and other relevant evidence. | 12-13 |
| Implications | 20 | Discuss the potential clinical use of the model and implications for future research. | 11 |
| **Other information** | | | |
| Supplementary information | 21 | Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets. | - |
| Funding | 22 | Give the source of funding and the role of the funders for the present study. | 13 |