Risk prediction for cardiac surgery and interventions

Tom Kai Ming Wang

MBCHB, FRACP

Thesis submitted for the degree Doctor of Medicine (MD)

University of Auckland, 2019

Acknowledgements

Abstract

Background:

Cardiac surgery and interventions have rapidly evolved and advanced significantly over the last two decades with improving outcomes. Risk models play a critical role in the decision-making for all cardiac procedures. Despite this, they are under-utilised and have not been assessed and validated in New Zealand cardiac surgery cohorts or for predicting morbidities and long-term mortality. There is sparse literature regarding the application of risk models in some clinically important settings such as infective endocarditis surgery and transcatheter aortic valve implantation. Finally the prognostic utility of recently developed high-sensitivity troponin assays in cardiac surgery have not been well evaluated. The aims of this thesis are to address each of these aforementioned issues, in assessing performance of risk scores and troponins in New Zealand cohorts, for predicting mortality and morbidities, and by means of meta-analyses where appropriate.

Methods:

A literature review of risk modelling and surgical risk scores for cardiac surgery and interventions was conducted. Eight studies were then performed, including six Auckland City Hospital based cohort studies and two meta-analyses. These studies:

1. Compared surgical risk scores for outcomes after isolated coronary artery bypass grafting.

2. Assessed the mortality and morbidities prediction of risk scores for isolated aortic valve replacement

3. Evaluated the utility of risk scores for mitral valve repair and replacement surgery.

4. Compared risk scores for combined aortic valve replacement and coronary bypass grafting surgery

5. Assessed surgical and endocarditis-specific risk scores for infective endocarditis operations.

6. Performed meta-analysis of surgical risk scores for infective endocarditis surgery.

7. Pooled performance of contemporary surgical risk scores when applied to transcatheter aortic valve implantation outcomes.

8. Reviewed the prognostic utility of high-sensitivity troponin T with ECG and/or echocardiographic changes for coronary artery bypass grafting and evaluating the universal definition for type 5 perioperative myocardial infarction.

Results:

1. The newer EuroSCORE II, STS and AusSCOREs had improved calibration, but only similar discrimination to EuroSCORE (c-statistic 0.64-0.68) for coronary artery surgery.

2. In isolated aortic valve replacement, all scores had moderate discrimination for operative mortality (c-statistic 0.68-0.75), however the STS Score performed best in the highest surgical risk quintile including for calibration, and also for composite and individual post-operative complications.

3. For isolated mitral valve repair or replacement, all scores had high discrimination (c-statistic 0.82-0.85) for operative mortality, and the STS Score performed the best for morbidities.

4. In patients undergoing combined aortic valve and coronary surgery, EuroSCORE II and STS Scores had superior discrimination and calibration to EuroSCORE for operative mortality.

5. Endocarditis-specific scores, especially the De Feo-Cotrufo Score performed better than EuroSCOREs at predicting mortality and morbidities after infective endocarditis surgery.

6. Despite our findings above, other studies combined in our meta-analysis found moderate discrimination of EuroSCOREs for predicting operative mortality after infective endocarditis surgery.

7. Surgical risk scores modestly discriminated operative and 1-year mortality with c-statistic 0.62 after transcatheter aortic valve implantation, although EuroSCORE II and STS had better calibration than EuroSCORE which significantly over-estimates risk.

8. Dual criteria of high sensitivity troponin T rise >140ng/L (10 times 99th percentile upper reference limit) with ECG and/or echocardiographic abnormalities, but not other criteria, was independently associated with 30-day and long-term mortality after coronary bypass surgery.

Conclusions:

Across various types of cardiac surgery, the EuroSCORE, EuroSCORE II and STS scores had similar discrimination, but EuroSCORE significantly over-estimated operative mortality, while the STS Score usually best predicted post-operative complications. Endocarditis-specific scores were superior to EuroSCOREs for endocarditis surgery, while transcatheter aortic valve implantation-specific models and validation are awaited due to modest performance of surgical risk scores in that setting. We also validated the prognostic utility of the Universal Definition's criteria for type 5 myocardial infarction and high-sensitivity troponins after coronary bypass grafting.

Preface

Cardiovascular disease remains the number one cause of death worldwide. Since the invention of cardiopulmonary bypass, cardiac surgery and the first heart transplant performed over 50 years ago, New Zealand has been one of the worldwide pioneers in this field. Whether for the treatment of severe coronary artery disease or valvular heart disease, cardiac surgery has been the gold standard procedure with large body of evidence for efficacy in improving survival and outcomes for these patients. More recently, percutaneous coronary and then structural heart interventions have been rapidly developed and expanded around the world. It is an exciting time to be working as a cardiology clinician and researcher in the current era.

Despite the significant proven benefits, cardiac procedures have one of the highest risk for adverse outcomes amongst all medical interventions. Careful evaluation to weigh the risks and benefits is critical to making an informed clinical decision in the management of individual patients, whether for medical, interventional, surgical or other therapy. Risk modelling plays a central role in diagnosis, management and prognosis across all areas of medicine and is particularly relevant in cardiology when there are competing modalities to manage cardiovascular diseases with high prevalence and particularly poor prognosis if untreated.

A body of literature exists for conventional cardiac surgery risk models and outcomes, but there some notable unaddressed gaps. Firstly, it is unknown whether these risk models accurately perform in New Zealand cardiac surgery cohorts. Secondly, whether the operative mortality risk scores can also predict post-operative complications and long-term mortality is unclear. Thirdly, can these models be appropriately used in clinically important scenarios such as infective endocarditis and transcatheter aortic valve implantation, or should procedural-specific scores be developed. Lastly, what is the prognostic utility of contemporary cardiac biomarkers for predicting adverse outcomes and diagnosing myocardial infarction after cardiac surgery. These are the central issues aimed to be addressed, investigated and reported in this thesis.

List of tables

List of figures

List of publications and presentations

Chapter 3

Wang TKM, Li AY, Ramanathan T, Stewart RAH, Gamble GD, White HW. Comparison of four risk models for contemporary isolated coronary artery bypass grafting. Heart, Lung and Circulation 2014;23:469-474(1)

Presented at Cardiac Society of Australia and New Zealand Annual Scientific Meeting 2013 (oral), and European Society of Cardiology Congress 2013 (moderated poster).

Chapter 4:

Wang TKM, Choi DH, Stewart R, Gamble GD, Haydock DA, Ruygrok P. Comparison of four contemporary risk models at predicting mortality after aortic valve replacement. Journal of Thoracic and Cardiovascular Surgery 2015;149:443-448.(2)

Wang TKM, Choi DH, Haydock DA, Gamble GD, Stewart R, Ruygrok P. Comparison of risk scores for prediction of complications following aortic valve replacement. Heart, Lung and Circulation 2015;24:595-601.(3)

Presented at Cardiac Society of Australia and New Zealand Annual Scientific Meeting 2014 (oral), and World Congress of Cardiology 2014 (oral).

Recipient Royal Australasian College of Physicians Trainee Research award 2015.

Chapter 5:

Wang TKM, Choi D, Ramanathan T, Ruygrok P. Comparing performance of risk scores for combined aortic valve replacement and coronary bypass grafting surgery. Heart, Lung and Circulation 2016;25:1118-1123.(4)

Presented at Cardiac Society of Australia and New Zealand Annual Scientific Meeting 2015 (poster), and European Society of Cardiology Congress 2015 (poster).

Chapter 6:

Wang TKM, Harmos S, Gamble G, Ramanathan T, Ruygrok P. Performance of contemporary surgical risk scores for mitral valve surgery. Journal of Cardiac Surgery 2017;32:172-176.(5)

Presented at Cardiac Society of Australia and New Zealand Annual Scientific Meeting 2016 (oral), and European Society of Cardiology Congress 2016 (oral).

Chapter 7:

Wang TKM, Oh T, Voss J, Kang N, Pemberton J. Comparison of contemporary risk scores for predicting outcomes after surgery for active infective endocarditis. Heart and Vessels 2015;30:227-234.(6)

Presented at Cardiac Society of Australia and New Zealand Annual Scientific Meeting 2013 (poster), and European Society of Cardiology Congress 2013 (oral).

Referenced by the 2015 European Society of Cardiology Guidelines on Infective Endocarditis(7) and the 2016 American Association of Thoracic Surgery Guidelines on Surgery for Infective Endocarditis(8).

Chapter 8

Wang TKM, Wang MT, Pemberton J. Risk scores and surgery for infective endocarditis: a meta-analysis. International Journal of Cardiology 2016;222:1001-1002.

Presented at Cardiac Society of Australia and New Zealand Annual Scientific Meeting 2015 (poster), and European Society of Cardiology Congress 2016 (moderated poster).

Chapter 9:

Wang TKM, Wang MT, Gamble G, Webster M, Ruygrok P. Performance of contemporary surgical risk scores for transcatheter aortic valve implantation: a meta-analysis. International Journal of Cardiology 2017;236:350-355.(9)

Presented at Cardiac Society of Australia and New Zealand Annual Scientific Meeting 2016 (poster), and European Society of Cardiology Congress 2016 (best poster in aortic valve interventions).

Chapter 10:

Wang TKM, Stewart RAH, Ramanthan T, Kang N, Gamble GD, White HW. Diagnosis of myocardial infarction after coronary artery bypass grafting with high-sensitivity troponin and relationship with mortality. European Heart Journal – Acute Cardiovascular Care 2013;2:323-333.(10)

Presented at Cardiac Society of Australia and New Zealand Annual Scientific Meeting 2013 (oral, young investigator finalist), and European Society of Cardiology Congress 2013 (poster).

Recipient Royal Australasian College of Surgeons Young Investigator Prize Cardiothoracic Section 2013.

Referenced by the 2018 Fourth Universal Definition of Myocardial Infarction Guidelines, Task Force from the European Society of Cardiology, American Heart Association, American College of Cardiology and World Heart Federation(11).

Abbreviations list

95%CI	95% confidence interval
AF	atrial fibrillation
ASCTS	Australasian Society of Cardiac and Thoracic Surgeons
AUC	area under curve
AVR	aortic valve replacement
BMI	body mass index
CABG	coronary artery bypass grafting
CCS	Canadian Cardiovascular Society
ECG	electrocardiogram
EuroSCORE	European system for cardiac operative risk evaluation
HR	hazards ratio
Hs-TnT	high sensitivity troponin T
LBBB	left bundle branch block
LVEF	left ventricular ejection fraction
MI	myocardial infarction
MVR	mitral valve surgery
NYHA	New York Heart Association
OR	odds ratio
ROC	receiver operating characteristic
STS	Society of Thoracic Surgeons
TAVI	transcatheter aortic valve implantation

Table of contents

1   Statistics of risk modelling

1.1   Developing a risk model

Risk models play an important role in all aspects of medicine for diagnosis, risk stratification, prognosis, management and clinical research(12-14). Every risk model consists of independent variable(s), a dependent variable, and a formula to relate these variables. Independent variables are otherwise known as predictors, parameters, risk factors, and in most cases when there are two or more in the formula it is called a multivariable model(13, 14). Independent variables can be either continuous or categorical. In cardiology and cardiac surgery, these can include demographics (such as age, sex and ethnicity), presentation (such as urgent surgery and critical pre-operative state), past history (risk factors like diabetes, heart disease like myocardial infarction and co-morbidities like peripheral vascular disease and stroke), investigation results (such as ejection fraction and renal function) and even operative variables (such as prosthetic valve and bypass time).

The dependent variable can take many forms. It usually refers to a disease in a diagnostic model, and an outcome in a prognostic model(14). It could be continuous or categorical but most commonly is categorical and binary. Furthermore, it could be a cross-sectional outcome at a certain time-point such as in-hospital outcomes, but can also be longitudinal, where the model is usually predicts the time to first event. Examples of binary categorical outcomes that could either be cross-sectional or longitudinal include mortality, stroke and so on, whereas a continuous variable could be, for example, number of days in the intensive care unit.

The formula or equation developed is based predominantly on the type of dependent variable. The commonest multivariable analysis techniques, applied to binary categorical dependent variables, are logistic regression for cross-sectional outcomes and Cox proportional hazards regression for longitudinal longer-term outcomes(14). Linear regression is an example of a multivariable analysis technique for a continuous dependent variable.

It is preferable that all and only variables with biologically plausible explanation to be associated with the dependent variable are analysed. The quality of data collected such as tight definitions, subjective or objective measurements, timing and interpretation, variability of observation as well as proportion of and handling of missing data are examples which may influence the validity and applicability of the model constructed(14, 15).

There is no standardised method for the selection of independent variables for and construction of a multivariable risk model. Univariable analysis can help identify factors significantly associated with the dependent variable, but other non-significant factors should not automatically be excluded. Also, a number of factors which are inter-related should not all be included in the multivariable model, as each of them would not add significant prognostic value in addition to each other. For example, variables such as dyspnoea on exertion, New York Heart Association (NYHA) class and history of heart failure should not all be used(15, 16). Existing strategies include the full model approach (including all collected independent variables), backward elimination approach (rather than forward selection building model from the strongest predictor) and univariable significance testing(15). The last of these although commonly performed may cause selection bias and overfitting of the model reducing its accuracy(15, 16).

After a risk model is created, it should undergo validation, and analysis of discrimination, calibration and other novel techniques are discussed below. The first test is always to internally validate the model with the cohort from which it was developed(15). The model generally performs best in the derivation cohort, so it is worrisome if the accuracy for this is suboptimal. The next step would be to assess its performance in other cohorts. Some studies randomly divide their cohort of patients into a derivation cohort and validation cohort, but this has the limitations of reducing the power of the derivation cohort and that the validation cohort is not a true external validation as it is fundamentally similar to the derivation cohort(14). Risk models should be formally externally validated before being utilised outside the derivation cohort institution.

1.2 Logistic regression

Logistic regression deserves special mention as the commonest method for developing a multivariable risk model. It is generally used to predict a binary, and usually short-term, dependent variable, such as operative mortality. The selection of parameters into the model can be difficult as mentioned before, but a number of independent variables are added to the model alongside the dependent variable. The logistic regression model will then calculate for each independent variable X_i whether they are a significant predictor with a P-value, a coefficient β_i, standard error of the confidence interval, and a constant for the overall equation β_0. Predictors that are not statistically or clinically significant can be excluded from the model and subsequent calculations with minimal effect on its accuracy. The model will then be calculated as predicted risk of dependent variable $= e^{\wedge}(\beta_0 + \Sigma \beta_i X_i)/(1 + e^{\wedge}(\beta_0 + \Sigma \beta_i X_i))$, or $1/(1 + e^{\wedge} - (\beta_0 + \Sigma \beta_i X_i))$ where e is the mathematical constant or Euler's number 2.718 (3 decimal places).

There are a few features of note for this formula. The calculation always gives a probability estimate r between 0 and 1. How each independent variable is incorporated in the model will depend on how it is calculated, for example age could the actual number of years, or age group like >60 years old, or per 10-year increase. The odds ratio for each variable X_i is equivalent to $e^{\wedge}\beta_i$ and the 95% confidence interval boundaries are the odds ratio multiply and divided by 1.96 x the standard error for each variable. The β_0 can be considered the calibrating factor of the formula which can scale the estimated risk up if more positive and down if more negative. Finally, the calculation may be complex for non-statisticians and most clinicians, therefore utilising a calculator which allows entering the parameters to automatically calculate the risk makes the formula more user-friendly.

Paradoxically, additive scores are commonly derived from logistic regression models. An additive score is one where the score is calculated by the sum of multiples of its constituent independent variables, which can be positive or negative and may not even be whole numbers. Widely used additive scores include the CHA2DS2-VASc Score for anticoagulation and stroke risk in atrial fibrillation(17) and the Well's Score for venous thromboembolism risk(18). Usually a logistic model of the predicted outcome is first constructed, and then the odds ratios of statistically significant individual independent variables are used as their respective coefficients in the sum. Similar to logistic scores, higher additive scores are meant to suggest higher risk of the dependent variable, however the actual value of the score does not equal the predicted risk. Therefore, although additive scores are easier to calculate and use, this important difference means that discrimination but not calibration can be used to assess the performance of additive scores. Calculators remain helpful for the utility of additive scores especially if there are many independent variables involved.

1.3 Discrimination

The basic measures of accuracy of tests include sensitivity and specificity, which are the probability of the test being positive when a disease or outcome is present, and being negative when a disease or outcome is absent, respectively. In contrast, the positive and negative predictive values are the probability of a disease or outcome being present or absent when the test is positive or negative respectively. The sensitivity and specificity are fixed for each test, however predictive values are influenced by the prevalence of the disease or outcome. These measures also apply to risk models when thresholds of the score are set to be positive or negative, but don't capture all the features of the risk model which is commonly a continuous parameter.

Discrimination is an important method of assessing whether a higher risk model score correlates with a higher risk of the outcome it predicts. The main measure for discrimination is the area (AUC) under the

receiver-operating characteristic (ROC) curve. This curve is a plot of sensitivity (true positive) against 1-specificity (false positive) of a range of score thresholds and probability of the outcome, therefore taking into account of both of these features. The concordance statistic (c-statistic) is equivalent to the AUC when the outcome is categorical and binary which is the majority of scenarios, however methods of calculating c-statistics of time-dependent outcomes have been published(19, 20). The value of AUC lies between 0 to 1, where 0.5 means lack of discrimination, while the closer to 0 or 1 means stronger discrimination. The ROC curve has further use in providing the sensitivity and specificity of risk models at various thresholds to best choose the one with the best clinical utility.

Another way to assess discrimination is the discrimination slope, analysed by the difference of the average risk scores of subjects with and without an event(19). Boxplots and histograms can depict this information visually, and a smaller overlap between those with or without the outcome suggest higher discrimination which is easily visualised. This measure does not apply to longitudinal outcomes.

1.4 Calibration

In contrast to discrimination, calibration assesses whether the estimated risk score reflects the observed risk of the patient. A 10% predicted risk of death should mean that in 100 of these similar or identical patients 10 of them should die. In the simplest form, taking the mean score and the actual outcome rate, the observed/expected ratio can be calculated. A ratio <1.0 means overestimation by the score and >1.0 means underestimation of the score. Note that calibration can only be assessed for logistic but not additive scores.

A calibration plot involves graphing outcome against predictions, with a 45-degee line of outcomes equalling predictions indicating perfect fit(13). For binary or categorical outcomes however the y-axis only takes a few values for the plot, making interpretation more challenging. The two important parameters derived from this plot are the intercept, which is related to systematic biases, and the calibration slope which determines the strength of association.

Another method for assessing calibration is the Hosmer-Lemeshow test looking at goodness-of-fit. This test compares deciles of predicted risk arbitrarily with the observed rate(21). A P-value <0.05 suggests significant discordance in calibration between the score and the outcome.

1.5 Other measures of risk model performance

There are a number of other measures of risk model performance, including traditional measures such as correlation R^2 and the Brier Score, while others have been proposed more recently, including

reclassification and clinical usefulness(13). Many performance measures of risk models are centred around the difference between predicted and observed outcomes, which if binary will be 0 or 1, and continuous an actual number. A smaller difference suggest a higher "goodness-to-fit" of the risk model to the observed data.

R^2 is a measure for continuous outcomes, otherwise known as explained variation or coefficient of determination, which is the square of correlation coefficient. It assesses the strength of correlation between the score and the continuous outcome, and the square ensures a positive number although the range remains between 0 to 1, with a higher number meaning stronger correlation. The commonest way of calculating R is the Pearson (bivariate) correlation coefficient, which is the covariance of the two continuous variables over the product of their standard deviations, and is a parametric measure of linear correlation. Non-parametric rank measures of correlation include the Spearman's rank correlation coefficient and Kendall rank correlation.

The Brier score on the other hand is the mean squared of difference between the score and binary categorical outcome 0 or 1, although it can be calculated for a greater number of categories of outcomes or longitudinal outcomes but is more complex. The lower the score indicates a smaller difference and stronger fit, ranging from 0 which is perfect correlation to 1. For the same cohort and outcomes, the Brier scores for different risk models can be directly compared, but it is trickier to compare Brier scores of risk models used in different cohorts. The maximum Brier score is 0.25 when the incidence of outcomes is 50%, and gets smaller as that figure moves further from 50%(13). The scaled Brier score, which is 1-Brier score/maximum Brier score, is sometimes used instead for comparison, and has a similar value to the Pearson R^2(22).

Reclassification is a newer technique for assessing risk model performance(13). The simplest form is the reclassification table, which evaluates how many subjects are reclassified, and appropriately so, when another marker or variable is added to a model(23). Within cross-classified categories, comparing observed outcomes with predicted risks is called the reclassification statistic(24). Another important measure in this area is the net reclassification index. It aims to quantify correct or incorrect reclassifications of a new compared to existing risk model(13). Finally there is the integrated discrimination index, which combines the net reclassification index across all thresholds.

2 Cardiac Surgery Risk Scores

2.1 EuroSCORE

Table 2.1 displays the most widely used cardiac surgery risk models. The first of these was the original EuroSCORE, derived from a consecutive cardiac surgery cohort under cardiopulmonary bypass from September to November 1995 at 128 surgical centres in eight European countries, totalling 13,302 patients(25). Isolated coronary artery bypass grafting (CABG) made up of 63.6% of patients, valve surgery 29.8% with an operative mortality of 4.8%. Operative mortality is defined as death within 30 days and/or during the same hospital admission as the operation. The operative mortality risk model was developed using multiple stepwise logistic regression of the 68 pre-operative and 29 operative risk factors with strict definitions felt to affect mortality, from which 17 independent predictors were identified. Internal calibration was described as "satisfactory" and discriminative power "very good" with c-statistic of 0.79.

The EuroSCORE was then published as an additive model in 1999(26). The score allocated to each parameter was essentially the odds ratio in the regression model rounded to the nearest whole number(25). The discriminative c-statistic for a separate validation cohort of 1,497 patients remained satisfactory at 0.76. The validation and developmental subsets were randomly divided from the same original database of patients.

Subsequently in a letter to the editor, the logistic EuroSCORE with an online calculator was published in 2003(27). Based on the identical developmental cohort and logistic model, the beta-coefficients and the constant β_0 were reported. Indeed the beta-coefficients could be calculated from odds ratio the original paper(25), but the constant completes the logistic model calculation based on the aforementioned formula. Such a logistic model provides an estimate of the operative risk which as described previously, allows assessment of calibration as well as being more intuitive for general clinicians and patients.

For many years the logistic EuroSCORE was the main cardiac surgery risk model utilised in clinical practice. Over time however it became apparent from many studies including two meta-analyses that the score tended to over-estimate operative mortality(28, 29). The observed/expected ratios overall were 0.43-0.60, and the over-estimation was most pronounced for isolated valve surgery, followed by isolated CABG, and least by the combined valve and CABG operations which carry a higher risk. The calibration discrepancy which increased with time is consistent with improving surgical outcomes, which may be a combination of improved patient selection, surgical technique and peri-operative care. Despite this, the pooled discrimination of operative mortality for EuroSCORE remained satisfactory with c-statistic 0.73-

0.77. It seems feasible to improve calibration without affecting discrimination of the EuroSCORE so that its logistic model could be retained but the constant β_0 adjusted to optimise calibration. An alternative was to develop a new risk model.

Table 2-1 Contemporary cardiac surgery risk models (USA=United States of America, NZ=New Zealand, N/A=not available)

Score	Surgery	Cohort date	Recruitment countries	Cohort size	Operative mortality	Parameters	C-statistic internal	C-statistic validation
EuroSCOREs								
EuroSCORE (additive)(26)	all cardiac surgery	1995 Sep-Nov	Europe (8)	13,302	4.8%	17	0.79	N/A
EuroSCORE (logistic)(27)	all cardiac surgery	1995 Sep-Nov	Europe (8)	13,302	4.8%	17	0.79	0.76
EuroSCORE II(30)	all cardiac surgery	2010 May-July	World (43)	22,381	3.9%	18	N/A	0.81
STS Scores								
Coronary(31)	CABG	2002-2006	USA	774,881	2.3%	22	0.81	0.81
Valve surgery(32)	AVR	2002-2007	USA	67,292	3.2%	11	0.78	0.76
	MV repair	2002-2008	USA	21,238	1.6%	11	0.86	0.84
	MV replacement	2002-2009	USA	21,229	5.7%	11	0.79	0.80
Valve+coronary surgery(33)	AVR+CABG	2002-2010	USA	66,074	5.6%	13	0.75	0.75
	MV repair+CABG	2002-2011	USA	21,924	7.4%	13	0.75	0.76
	MV replacement+CABG	2002-2012	USA	13,663	11.6%	13	0.76	0.74
Australasian Scores								
AusSCORE(34)	CABG	2001 July-2005 June	Australia	7,709	1.7%	8	0.84	0.84
Aus-AVR(35)	AVR	2001 July-2008 June	Australia	3,544	4.2%	8	0.78	0.73
AusSCORE II(36)	CABG	2001-2011	Australia/NZ	53,681	1.6%%	14	0.82	0.85

2.2 EuroSCORE II

Due to the calibration issues with the logistic EuroSCORE over time, EuroSCORE II was developed and published in 2012 with its online calculator(30). This was based on a cohort operated 15 years later than the logistic EuroSCORE during May-July 2010 with 154 hospitals in 43 countries not restricted to Europe including 22,381 consecutive cardiac surgery patients. Operative mortality of 3.9% was lower than the EuroSCORE cohort of 4.8%(26, 30). New risk factors were added to the model. There were less isolated CABGs at 46.7%, similar to valve operations at 46.3%.

Following multivariable logistic regression analysis, a new risk model made up of 18 parameters was created(30). The constant β_0 was -5.32, more negative than the $\beta0$ for EuroSCORE reflecting the difference in calibration. For the same patient, the calculated EuroSCORE II is commonly significantly lower than the EuroSCORE. The c-statistic for the validation cohort of 5,553 patients from the same paper by EuroSCORE II was 0.80.

Since its publication EuroSCORE II has been assessed by many other studies including several meta-analyses(37, 38). These report similar c-statistics of 0.77-0.79 to the logistic EuroSCORE, but with improved calibration in terms of observed/expected ratio being around 1. Thus the EuroSCORE II fitted better with contemporary outcomes. The EuroSCORE II has replaced EuroSCORE in the clinical setting and is an alternative to the Society of Thoracic Surgeon's Score, which together are recommended by guidelines for risk stratification prior to cardiac surgery(39).

2.3 Society of Thoracic Surgeons' (STS) Scores

The Society of Thoracic Surgeons (STS) convened in 1988 to develop a national cardiothoracic surgery database for the United States and opened to its members in 1990(40). Over time it became one of the world's largest registry, with 950 (nearly 90% of all) cardiac surgery providers of the United States enrolled by 2008(31). Risk models with online calculators have been created and updated on a regular basis derived using multivariable logistic models of the endpoints of interest. The STS Scores are now, along with the EuroSCOREs, the most widely used cardiac surgery risk scores.

The first unique feature about the STS Scores is that there are separate models for different types of cardiac surgery. These were most recently published in 2009 and based on the 2002-2006 United States cardiac surgery cohort in the STS database for isolated CABG(31), isolated aortic or mitral valve (MV) surgery, the latter separately as repair or replacement(32), and aortic or mitral valve combined with CABG surgery(33). The number of patients and operative mortality for each type of surgery were isolated CABG 774,881 and 2.3%, isolated aortic valve replacement (AVR) 67,292 and 3.2%, isolated

MV repair 21,238 and 1.6%, isolated MV replacement 21,229 and 5.7%, AVR+CABG 66,074 5.6%, MV repair+CABG 21,924 and 7.4% and MV replacement and CABG 13,663 and 11.6% respectively(31-33). The implication is that some more complex types of cardiac surgery such as double-valve surgery and aortic surgery cannot be directly assessed by a STS risk score.

The second important characteristic is that there are different models for different clinically relevant endpoints(31-33), unlike the EuroSCOREs which were designed to only predict operative mortality(26, 30). The primary endpoint of operative mortality, in-hospital or within 30 days is identical. Eight other in-hospital clinical endpoints with separate risk models for each type of cardiac surgery include permanent stroke (central neurological deficit >72 hours), renal failure (increase serum creatinine to more than 2.0mg/dL, a doubling of most recent pre-operative creatinine level or new dialysis requirement), prolonged ventilation >24 hours, deep sternal wound infection, reoperation (all-cause), composite morbidity or mortality (composite of the above endpoints), prolonged hospital stay >14 days post-operatively, and short hospital stay <6 days post-operatively(31). Raw rates vary significantly depending on the endpoint and type of surgery.

Another factor to consider for the STS Score is the number of parameters, both recorded in the database and then used to develop the logistic model. The online calculator has 42 overall parameters, from demographics, medical history, presentation, investigations and surgical factors(31-33). The complexity of calculation may make the STS Score less preferred for some, especially if the accuracy is not superior to other scores, although it remains the most widely used risk score in the United States. To be precise however, the large number of variables account for all the possible predictors for the 7 cardiac surgery types and 9 different outcomes, and the actual number for each scenario is significantly less. For example, only 22 variables are required for the operative mortality model for isolated CABG which is only slightly more than that required for EuroSCOREs calculations.

2.4 Australasian Society of Cardiac and Thoracic Surgeons' Scores

The Australasian Society of Cardiac and Thoracic Surgeons (ASCTS) database was created in 2001 from a group of Australian cardiac surgical centres and later New Zealand also(34). The logistic EuroSCORE had been the most widely used risk model in Australasia for some time, but subsequently an Australasian study published in 2006 found the EuroSCORE to overestimate operative mortality of 8,331 ASCTS database patients(41), suggesting that new models needed to be developed locally.

The first model called the AusSCORE was based on 7,709 isolated CABG patients operated between July 2001 and June 2005, with a 30-day mortality rate of 1.74%(34). Multiple logistic regression was performed to create a risk score for predicting 30-day mortality, based on 8 parameters, published as both

a logistic model with beta-coefficients and the constant, and as an additive score approximating odds ratios. The c-statistic for the AusSCORE at 0.84 was higher than the logistic EuroSCORE of 0.76-0.79, for both the cderivation cohort of 5,151 patients and the validation cohort of the remaining 2,558 patients. For the same patient, the AusSCORE is generally significantly lower than the EuroSCORE.

Soon after, a second risk score was published based on 3,544 AVR patients operation from July 2001 to June 2008, with 4.15% early mortality, defined as in-hospital or within 30-days of surgery(35). Of note, the development cohort included patients undergoing AVR with or without CABG and/or MV surgery. Similar methods again produced another logistic and additive model with 8 variables. C-statistics were 0.78 for the development cohort and 0.73 for a different 1,268 patient cohort.

The number of patients collected in the database increased to 53,681 cardiac surgery patients from 2001-2011, which now included New Zealand patients, with 31,250 undergoing isolated CABG with operative mortality 1.6%. A new model called AusSCORE II was developed using multivariable logistic regression model and 13 identified predictors of 30-day mortality and was published in 2014 as a logistic regression model only but without a constant(36). When assessed against the developmental cohort and validation cohort here, both AusSCOREs had high c-statistic 0.82-0.85, but the AusSCORE significantly underestimated 30-day mortality resulting in poorer calibration.

2.5 Special scenario: infective endocarditis

Infective endocarditis is a heterogeneous disease with high mortality. Surgery is required in approximately half of patients for treatment of severe valvular regurgitation with heart failure, uncontrolled infection or embolism(7). The decision for surgery is sometimes complicated because of high risk of mortality both with and without surgery. Guidelines recommend risk stratification and involvement of a multidisciplinary endocarditis or heart team for decision-making.

The EuroSCOREs, based on the general cardiac surgery population including active endocarditis as a parameter, can be used in the setting of infective endocarditis(26, 30). Notably in the developmental cohorts endocarditis patients only made up 1.5-2.2% of all patients. My meta-analysis in chapter 8 found the pooled c-statistics were 0.76 for EuroSCORE in seven studies and 0.79 for EuroSCORE II in three studies(42). However, calibration was somewhat suboptimal with observed/expected ratios of 0.76 and 1.25 respectively for the EuroSCORE and EuroSCORE II respectively. On the other hand, the STS Score can sometimes be used for endocarditis surgery if the model for the type of surgery is published, but cannot be used in endocarditis surgery involving the aortic root or multiple valves, which may limit its use(31-33). Similarly the Australasian Scores can only be applied to surgeries involving CABG and/or AVR(34, 35).

A few endocarditis-specific risk models have been developed recently, summarised in table 2.2. Two of the risk models by Costa et al(43) and the PALSUSE(44) models were derived from endocarditis patients who may have had surgery or medical therapy, and therefore is less specific to the endocarditis surgery group. This is because the risk of medically treated endocarditis patients is bimodal, either not severe enough to require surgery, or too high risk to undertake surgery, and possibly because of the latter, in-hospital mortalities of these two groups were high at 24-26%.

The other scores are based on endocarditis surgery only cohorts. The STS Endocarditis score by Gaca et al.(45) is based on endocarditis patients from the STS database which has the largest derivation cohort, and based on parameters in that. Therefore endocarditis-specific variables proven to be important in other studies such as blood culture results and intracardiac abscess are not included. The De Feo-Cotrufo Score(46) was developed from an endocarditis surgery cohort and is the one recommended for use in current endocarditis guidelines(7), referencing my study in Chapter 7 to finding it performing best in this setting(6). The two newest scores Risk-E(47) and AEPEI(48) published in 2017 have not been widely evaluated, but are the first to have a validation cohort in the original publication, though the c-statistic does drop from the developmental to the validation cohort by 0.06-0.11. Further research is required in this clinically important and high risk setting.

Table 2-2 Endocarditis surgery specific risk scores (USA=United States of America, N/A=not available)

Score	Cohort	Cohort date	Countries	Derivation cohort	Operative mortality	Parameters	Model features	C-statistic internal	C-statistic validation
Costa(43)	Endocarditis (surgery+ medical therapy groups)	1988-1998	USA	186	26.3%	7	Logistic model (but constant not provided), additive model	0.84	N/A
Gaca (STS)(45)	Endocarditis surgery	2002-2008	USA	19,543	8.2%	11	Logistic model (but constant not provided), additive model	0.76	N/A
De Feo-Cotrufo(46)	Endocarditis surgery	1980-2009	Italy	440	9.1%	6	Logistic model (but constant not provided), additive model	0.88	N/A
PALSUSE(44)	Endocarditis (surgery+ medical therapy groups)	2008-2010	Spain	1,000	24.3% (in-hospital)	7	Logistic model (but constant not provided), additive model	0.84	N/A
RISK-E(47)	Endocarditis surgery	1996-2014	Spain	424	29.2%	8	Logistic model, additive model	0.87	0.76
AEPEI(48)	Endocarditis surgery	2000-2015	Italy	361	15.5%%	5	Logistic model, additive model	0.78	0.72

2.6   Special scenario: TAVI

Transcatheter aortic valve implantation (TAVI) is an alternative modality for treating severe aortic valve disease to cardiac surgery(39, 49). Randomised multicentre trials investigating the efficacy and safety of TAVI are summarised in table 2.3, with the Sapien prosthesis being balloon-expandable and CoreValve and Evolut R being self-expandable(50-54). The initial PARTNER 1 randomised trials found TAVI to have similar outcomes to AVR in high-risk operable patients and superior to medical therapy in inoperable patients, although there were concerns about higher TAVI stroke or transient iscahemic attack rates than AVR and medical therapy(50, 51). Next study published was CoreValve which found higher 1-year survival for TAVI than AVR in high-risk patients(52). More recently, PARTNER 2 reported that in intermediate risk patients, which include STS Score of 4-8% as one of its inclusion criteria, TAVI is equivalent to AVR, and in the transfemoral TAVI subgroup they had higher survival than AVR, further expanding their indication(53). The latest SURTAVI trial also found similar outcomes between TAVI and AVR in intermediate risk patients(54). Thus TAVI indication and utility have greatly expanded in recent years in developed countries, and long-term outcomes beyond 10-years are awaited, as well as its role in low risk patients.

Table 2-3 Randomised trials of TAVI versus medical therapy or AVR

Score	Published	Prosthesis	Cohort date	Group	Size	Criteria	STS	Mortality 30-days	Mortality 1-year	Conclusion
PARTNER 1B(50)	2010	Sapien	2007/5/11-2009/3/16	TAVI	179	STS>10%+inoperable	11.2%	5.0%	30.7%	Lower 1-year mortality for TAVI
				Medical	179		12.1%	2.8%	50.7%	Higher stroke or transient ischamiec attack for TAVI
PARTNER 1A(51)	2011	Sapien	2007/5/11-2009/8/28	TAVI	348	STS>10%+high risk operable	11.8%	3.4%	24.2%	Similar 1-year mortality
				AVR	351		11.7%	6.5%	26.8%	Higher stroke or transient ischamiec attack for TAVI
CoreValve(52)	2014	CoreValve	2011 Feb-2012 Sep	TAVI	394	consensus risk of death 30 days >15%	7.3%	3.3%	14.2%	Lower 1-year mortality for TAVI
				AVR	401	complications risk 30 days <50%	7.5%	4.5%	19.1%	Similar stroke rates
PARTNER 2(53)	2016	Sapien XT	2011 Dec-2013 Nov	TAVI	1011	STS 4-8%	5.8%	3.9%	12.3%	Similar 2-year mortality and stroke
				AVR	1021		5.8%	4.1%	12.9%	Transfemoral TAVI lower events than AVR
SURTAVI(54)	2017	CoreValve+ Evolut R	2012/6/19-2016/6/30	TAVI	864	STS 3-15%	4.4%	2.2%	6.7%	Similar 2-year mortality and stroke
				AVR	796		4.5%	1.7%	6.8%	

Risk models play in important role in stratification and selection of treatment modality for aortic valve disease according to guidelines, in clinical trials and practice(39, 49). As seen from above they may constitute the inclusions into randomised trials and influence guidelines. Despite this, they were not validated for TAVI outcomes before these trials were designed, and the STS score were not calibrated with the 30-day mortality after TAVI. This is one of the aims of my thesis, to meta-analye the performance of surgical risk scores at predicting outcomes after TAVI(9). We found that EuroSCORE, EuroSCORE II and STS Scores all only modestly predict operative mortality after TAVI c-statistic 0.62. Calibration was also variable, with EuroSCORE over-estimating, and EuroSCORE II and STS Score fitting better though suboptimally with TAVI outcomes. Reasons why surgical risk scores suboptimally predict TAVI outcomes include that they were developed from AVR cohorts, and only a minority of surgical patients have intermediate to high risk seen in TAVI patients. There is a clear and unmet need to design TAVI-specific risk models.

Table 2.4 lists recently developed risk models from TAVI cohorts. These have been developed from TAVI trials or registries, aimed to predict either 30-day and/or 1-year mortality(55-60). The cohort sizes are less than that for surgical risk scores because of the relative novelty of TAVI. It can be seen that these scores only perform moderately at best, with c-statistic of 0.67-0.79 in internal validation and more importantly 0.66-0.71 in external validation, which is somewhat better than surgical risk scores at predicting TAVI though still less than surgical risk scores predicting cardiac surgery(9). There is therefore large room for improvement in the risk prediction for TAVI patients, where in elderly cohorts, other factors such as frailty may play an important role.

Table 2-4 TAVI-specific risk models (USA=United States, N/A=not available)

Score	Cohort date	Countries	Cohort Size	Dependent variable	30-day mortality	1-year mortality	Parameters	C-statistic internal	C-statistic validation
OBSERVANT(55)	2010 Dec- 2012 Jun	Italy	1,256	30-day mortality	6.1%	N/A	7	0.73	0.71
posT TAVI(56)	2007-2012	Italy	1,064	1-year mortality	7.0%	15%	3	0.68	0.67
FRANCE-2(57)	2010-2011	France	2,552	30-day or in-hospital mortality	10.0%	N/A	9	0.67	0.59
TAVI2-SCORe(58)	2007 Nov- 2012 Nov	Netherlands	511	1-year mortality	5.7%	17%	8	0.72	N/A
CoreValve(59)	2011 Feb- 2012 Sep	USA	2,482	30-day and 1-year mortality	5.8%	23%	4, 5	0.75, 0.79	N/A
STS/ACC/TVT(60)	2011 Nov- 2014 Feb	USA	13,718	in-hospital mortality	5.3%	N/A	7	0.67	0.66

3 Comparison of four risk models for contemporary isolated coronary artery bypass grafting.

Coronary artery bypass grafting (CABG) is the commonest form of cardiac surgery, and gold standard for the treatment of severe multivessel and/or left main coronary artery disease(61). It constitutes the greatest proportion of the derivation cohort of cardiac surgery risk models(27, 31), and therefore is the best starting point for risk model evaluation, which has not previously been examined in New Zealand cohorts. This chapter compared the prognostic utility of contemporary risk models at predicting adverse outcomes after CABG.

This manuscript was published in 2014 in Heart, Lung and Circulation volume 23 pages 469-474. As of September 2018, it had 16 citations on google scholar, including two meta-analyses(37, 38). It was an oral presentation at the Cardiac Society of Australia and New Zealand Annual Scientific Meeting 2013 and moderated poster presentation at the European Society of Cardiology Congress 2013.

Author list and affiliations:

Tom Kai Ming Wang[1,2], Acrane Yihe Li[1], Tharumenthiran Ramanathan[1], Ralph Stewart[1,2], Greg Gamble[2], Harvey White[1,2]

1. Green Lane Cardiovascular Service, Auckland City Hospital, Auckland, New Zealand

2. Department of Medicine, University of Auckland, Auckland, New Zealand

Contribution of candidate: study planning and design, data collection, analysis and interpretation, writing manuscript, revision in response to co-authors and editors, conference presentations, journal submission

3.1 Abstract

Background: EuroSCORE I and the Society of Thoracic Surgeons' (STS) Score have been the most widely used risk scores for cardiac surgery. The revised EuroSCORE II and the AusSCORE, based on an Australasian population, were recently developed. We compared the prognostic utility of these four scores for mortality as well as morbidity in patients undergoing isolated coronary artery bypass grafting (CABG).

Methods: The scores were retrospectively calculated for isolated CABG patients at Auckland City Hospital during July 2010-June 2012. Discrimination and calibration of outcomes were assessed.

Results: 818 patients were followed for 1.6+/-0.6 years. Mortality at 30 days was 1.6% and 2.9% on follow up. Medians (Interquartile range) for EuroSCORE I were 2.8% (1.6%, 5.2%), EuroSCORE II 1.6% (1.0%, 2.8%), STS Score 2.3% (1.3%, 4.5%) and AusSCORE 0.5% (0.2%, 1.1%). C-statistics and Hosmer-Lemeshow test p-values for these scores for 30-day mortality were Euro score I 0.675 (95%CI 0.531-0.819)/0.061, EuroSCORE II 0.642 (0.503-0.780)/0.150, STS Score 0.641 (0.507-0.775)/0.243 and AusSCORE 0.661 (0.516-0.807)/0.420. Only EuroSCORE I and STS scores predicted mortality at follow-up (c=0.639 and 0.666). All scores predicted composite morbidity. C-statistics were EuroSCORE I 0.678, EuroSCORE II 0.634, STS score 0.584 and AusSCORE 0.645).

Conclusion: EuroSCORE II, STS Score and AusSCORE had slightly improved calibration but similar discrimination for 30-day mortality compared to EuroSCORE I. Revision of risk models to fit contemporary surgical outcomes is important, but there may only be modest room for improvement in discrimination.

3.2 Introduction

Several operative risk scores for cardiac surgery have been developed in the last few decades including the Parsonnet Score(62), EuroSCOREs(26, 27) and Society of Thoracic Surgeon's (STS) score(31). EuroSCORE I was developed from a European cohort of 14,781 patients having cardiac surgery during 1995 for 30-day mortality, and published as an additive model in 1999(26) and logistic model in 2003(27). The STS score was developed to predict operative morbidity and was derived from an American cohort of 774,881 isolated coronary artery bypass grafting (CABG) patients during 2002-2006 and published in 2008(31).

Despite the early validation of EuroSCORE I in large international populations(63, 64), more recent studies, found the score over-estimated operative mortality, probably because of improving operative and

peri-operative management(28, 65). In Australasian populations characterised by significant ethnic diversity, the EuroSCORE I also over-estimated operative mortality(41). The AusSCORE was published in 2009 from 11,823 patients undergoing isolated CABG during 2001-2005 from the Australasian Society of Cardiac and Thoracic Surgeon's (ASCTS) database to predict 30-day mortality(34). To date, there are no studies assessing its external validity.

More recently as a project to revise the original EuroSCORE to fit contemporary cohorts, the EuroSCORE II was developed from an international cohort of 22,381 patients undergoing cardiac surgery during 2010 and published in 2012(30). Studies which have assessed the external validity of this new score have reported mixed results with EuroSCORE II performing better(66-69) or similar(70, 71) to EuroSCORE I.

EuroSCORE II, STS Score and AusSCORE have not been directly compared for CABG, or assessed in Australasian cohorts. In addition the comparative value of the different scores for predicting mortality beyond 30 days is uncertain. Our objective was to compare the predictive efficacy of logistic EuroSCORE I, EuroSCORE II, STS Score and AusSCORE for morbidity and mortality at 30 days and longer follow-up after isolated CABG.

3.3 Methods

Patient selection and data collection

Ethics approval of this study was obtained from our institution's ethics review committee. Consecutive patients undergoing isolated CABG without concomitant valve surgery at Auckland City Hospital were included from July 2010 to June 2012. Relevant clinical characteristics were collected from computerised records. Logistic EuroSCORE I(27), EuroSCORE II(30), STS Score(31) and AusSCORE(34), all risk models for predicting 30-day operative mortality after cardiac surgery, were retrospectively calculated from all patients using available data.

The EuroSCORE II definitions were used for pre-operative characteristics, including extracardiac arteriopathy, chronic lung disease, critical pre-operative state, poor mobility and categories for renal impairment using creatinine clearance or dialysis, left ventricular ejection fraction and pulmonary hypertension(30). Angina was graded using the Canadian Cardiovascular Society Classification (CCS) and dyspnoea by the New York Heart Association Functional Classification (NYHA). Hypertension was defined as prescribed medications for lowering blood pressure, any measurement of over 140/90mmHg prior to operation and/or a previous formal diagnosis. Hypercholesterolaemia referred to total cholesterol >5.0mmol/L, on treatment to lower cholesterol before admission and/or a previous formal diagnosis.

Stroke included any previous history of a neurological deficit that persisted over 24 hours and caused by disturbance of cerebral blood supply. Number of grafts and durations of cardiopulmonary bypass and aortic cross-clamp were collected.

Post-operatively, high-sensitivity troponin T (hs-TnT) was measured at 12-24 hours routinely. Development of new Q-waves or left bundle branch block (LBBB) on post-operative ECG or new regional wall motion abnormalities on post-operative echocardiograms were independently interpreted by two authors (TKMW and HDW). Peri-operative myocardial infarction was defined as post-operative hs-TnT>140ng/L (10 times 99% upper reference limit) and the ECG and/or echocardiographic criteria above, as per the universal definition(10, 72). Five other post-operative complications (stroke, renal failure, ventilation >24 hours, deep sternal wound infection and return to theatre) as defined by the Society of Thoracic Surgeon's (STS) score and their composite were determined(31). Mortality data were checked against New Zealand's national registry up till 31 December 2012. The three pre-specified outcomes of interest were operative mortality (death in-hospital or within 30 days of operation), medium-term mortality (death during follow-up) and composite morbidity.

Statistical analyses

SPSS (Version 17.0, SPSS Inc., Chicago, IL, USA) and Prism (Version 5, GraphPad Software, San Diego, CA, USA) were used for statistical analysis. Continuous and categorical variables are presented as mean (standard deviation) and percentages (frequency) respectively. Discriminative powers of post-operative outcomes were assessed using the area under receiver-operative characteristics curves (c-statistic) and 95% confidence interval (95%CI) reported. Calibration was assessed by the Hosmer-Lemshow goodness-of-fit test. Kaplan-Meier curves and log-rank (Mantel-Cox) test were used for longitudinal survival analysis, stratifying each risk score into quintiles. P-values less than 0.05 were deemed statistically significant and all statistical tests were two-tailed.

3.4   Results

Table 3.1 presents the baseline characteristics of the study population. Mean age was 64.5+/-10.0 years and 20.2% (168) were female. Mean follow-up was 1.6+/-0.6 years. The median predicted 30-day mortality (interquartile range IQR) for EuroSCORE I was 2.8% (1.6%, 5.2%), EuroSCORE II 1.6% (1.0%, 2.8%), STS Score 2.3% (1.3%, 4.5%) and AusSCORE 0.5% (0.2%, 1.1%). Table 3.2 lists the operative variables and post-operative outcomes.

Table 3-1 Baseline characteristics

Characteristics	Study population n=818
Demographics	
Age (years)	64.5 (10.0)
Female	20.5% (168)
Ethnicity	
New Zealand Europeans	52.9% (433)
Maori/Pacific Islander	24.7% (202)
Other	22.4% (183)
Body mass index (kg/m^2)	29.1 (5.3)
Presentation	
Canadian Cardiovascular Society angina class IV	37.2% (304)
New York Heart Association dyspnoea class	
I	82.0% (671)
II	9.7% (79)
III	4.3% (35)
IV	4.0% (33)
Recent myocardial infarction within three months	53.4% (437)
Urgent operation as inpatient	79.6% (651)
Critical pre-operative state	9.5% (78)
Past Medical History	
Myocardial infarction	66.4% (543)
Percutaneous coronary intervention	11.0% (90)
Cardiac surgery	1.3% (11)
Congestive heart failure	5.4% (44)
Atrial fibrillation	7.5% (61)
Diabetes	38.3% (313)
Diabetes on insulin	10.9% (89)
Hypercholesterolaemia	91.7% (750)
Hypertension	70.0% (573)

Current smoker	14.5% (119)
Family history of coronary artery disease	15.5% (127)
Stroke	6.1% (50)
Extracardiac arteriopathy	11.1% (91)
Chronic lung disease	16.9% (138)
Dialysis	2.9% (24)
Poor mobility	0.1% (1)
Investigations	
Left main stem stenosis $\geq$50%	43.9% (359)
Three-vessel disease	81.4% (666)
Ejection fraction	
$\geq$50%	70.8% (579)
30-49%	23.1% (189)
20-29%	5.3% (43)
<20%	0.9% (7)
Pulmonary hypertension	
None ($\leq$30mmHg)	94.6% (774)
Moderate (31-55mmHg)	4.5% (37)
Severe (>55mmHg)	0.9% (7)
Creatinine clearance (mL/min)	87 (38)
Scores	**Median (lower quartile, upper quartile)**
EuroSCORE I (logistic)	2.8% (1.6%, 5.2%)
EuroSCORE II	1.6% (1.0%, 2.8%)
STS Score	2.3% (1.3%, 4.5%)
AusSCORE	0.5% (0.2%, 1.1%)

Table 3-2 Operative variables and post-operative outcomes

Operation Details	
Off-pump	2.7% (22)
Number of bypassed vessels	3.2 (0.8)
Left internal mammary artery graft	97.9% (801)
Right internal mammary artery graft	6.0% (49)
Radial artery graft	23.1% (189)
Saphenous vein graft	93.8% (767)
Operation time (minutes)	205 (51)
Cardiopulmonary bypass time (minutes)	91 (26)
Cross-clamp time (minutes)	59 (21)
Post-operative outcomes	
Composite morbidity	17.8% (146)
Stroke	1.1% (9)
Renal failure	2.2% (18)
Ventilation >24 hours	13.2% (108)
Deep sternal wound infection	0.4% (3)
Return to theatre	5.0% (41)
Myocardial infarction	13.9% (78)
Operation to discharge (days)	8.1 (5.7)
30-day mortality	1.6% (13)
Re-admission to hospital within 30 days	18.6% (152)

Thirty day mortality

Thirty day mortality was 1.6% (n=13). Results of discrimination and calibration analysis are presented in table 3.3. C-statistics (95%CI) for all 4 scores were significantly higher than chance: EuroSCORE I 0.675 (0.531-0.819), EuroSCORE II 0.642 (0.503-0.780), STS Score 0.641 (0.507-0.775) and AusSCORE 0.661 (0.516-0.807).

Table 3-3 Discrimination (area under curve) and calibration analyses

Outcomes	EuroSCORE I	EuroSCORE II	STS Score	AusSCORE
30-day mortality	0.675 (0.531-0.819)	0.642 (0.503-0.780)	0.641 (0.507-0.775)	0.661 (0.516-0.807)
Hosmer-Lemeshow test	χ^2=13.5, p=0.061	χ^2=12.0, p=0.150	χ^2=10.3, p=0.243	χ^2=8.14, p=0.420
Brier Score	0.0174	0.0156	0.0163	0.0154
Mortality during follow-up	0.639 (0.525-0.752)	0.604 (0.483-0.752)	0.666 (0.564-0.769)	0.593 (0.480-0.705)
Myocardial infarction	0.476 (0.407-0.546)	0.509 (0.444-0.574)	0.530 (0.464-0.595)	0.490 (0.421-0.559)
Composite morbidity	0.678 (0.631-0.726)	0.634 (0.582-0.686)	0.584 (0.532-0.635)	0.645 (0.593-0.698)
Stroke	0.736 (0.582-0.889)	0.532 (0.405-0.658)	0.468 (0.257-0.680)	0.694 (0.537-0.851)
Renal failure	0.656 (0.569-0.744)	0.635 (0.512-0.758)	0.707 (0.589 (0.826)	0.663 (0.547-0.779)
Ventilation >24 hours	0.712 (0.658-0.765)	0.655 (0.595-0.715)	0.561 (0.501-0.622)	0.669 (0.609-0.728)
Deep sternal wound infection	0.517 (0.221-0.813)	0.720 (0.547-0.893)	0.441 (0.289-0.594)	0.470 (0.371-0.568)
Return to theatre	0.605 (0.523-0.687)	0.626 (0.534-0.718)	0.566 (0.480-0.652)	0.573 (0.482-0.664)

In terms of calibration, the Hosmer-Lemeshow test for predicting 30-day mortality approached statistical significance for EuroSCORE 1 (p=0.061) and was non-significant, for EuroSCORE II (p=0.15), STS Score (p=0.243), and AusSCORE (p=0.42).

Surgical morbidity

Composite surgical morbidity was 17.8% (n=146), mainly driven by prolonged ventilation >24 hours at 13.2% (108). C-statistics (95%CI) for all 4 scores were again significantly better than chance: EuroSCORE I 0.678 (0.631-0.726), EuroSCORE II 0.634 (0.582-0.686), STS Score 0.584 (0.532-0.635) and AusSCORE 0.645 (0.593-0.698).

For individual post-operative complications, EuroSCORE I had the greatest c-statistics for stroke 0.736 (0.582-0.889) and ventilation > 24hours 0.712 (0.658-0.765). EuroSCORE II had the greatest c-statistics for deep sternal wound infection 0.720 (0.547-0.893) and return to theatre 0.626 (0.534-0.718). STS Score had the highest c-statistic for renal failure 0.707 (0.589-0.826). No scores reliably predicted peri-operative myocardial infarction.

Survival

Mortality during follow-up was 2.9% (n=24). There were 11 late deaths, 9 of which occurred between 30 days and 1-year. The c-statistic (95%CI) for mortality during follow-up for EuroSCORE I was 0.639 (0.525-0.752), STS Score 0.666 (0.564-0.769), EuroSCORE II 0.604 (0.483-0.752) and AusSCORE 0.593 (0.480-0.705).

Kaplan-Meier survival curves are displayed in Figure 3.1 for quintiles (cutpoints shown adjacent to each curve) of EuroSCORE I (A), EuroSCORE II (B), STS Score (C) and AusSCORE (D). There were significant differences in survival by EuroSCORE I (p=0.047) and STS Score (p=0.013) quintiles, but not by EuroSCORE II (p=0.124) or AusSCORE (p=0.667) quintiles.

Figure 3-1 Kaplan-Meier survival curves stratified by quintiles of A) EuroSCORE I, B) EuroSCORE II, C) STS Score and D) AusSCORE
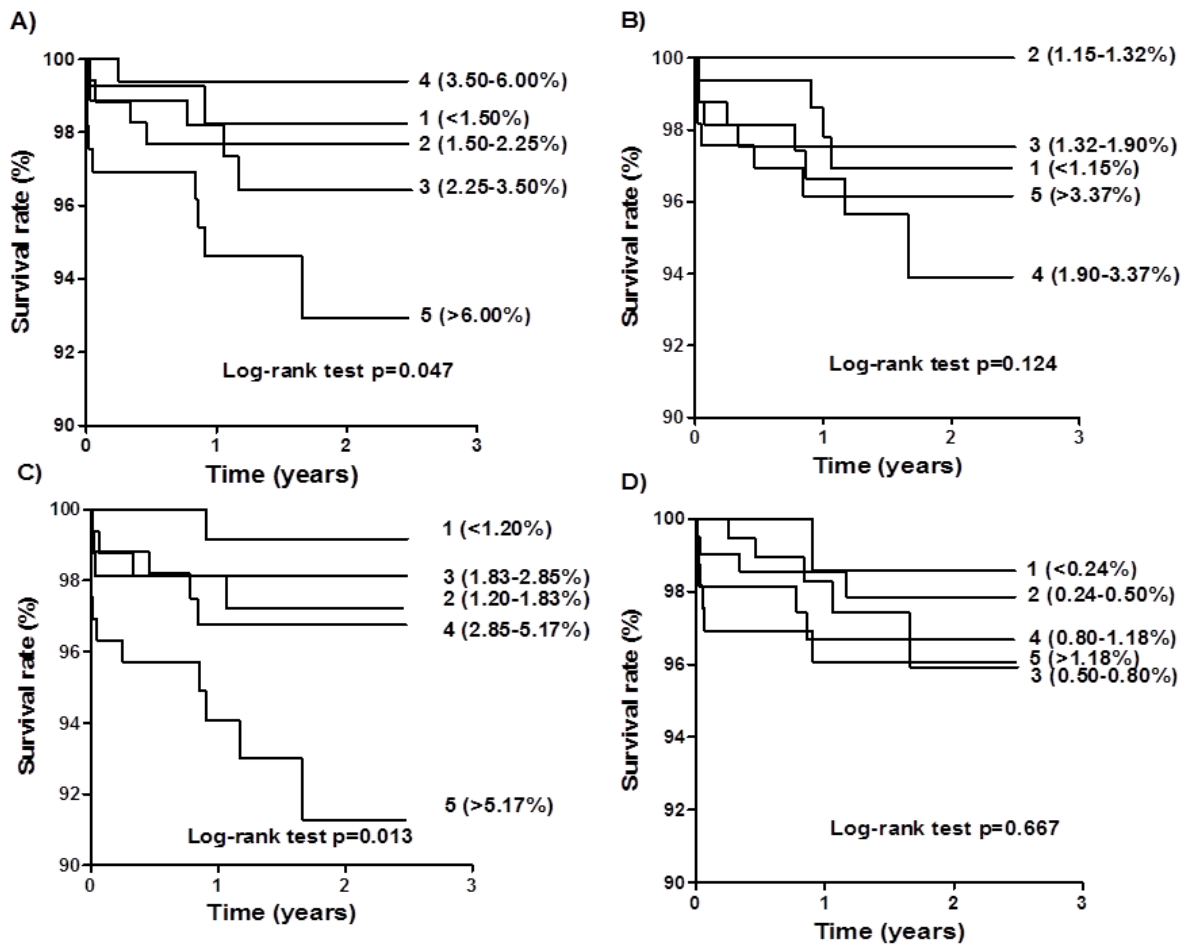
3.5 Discussion

Our study shows that EuroSCORE I, EuroSCORE II, STS Score and AusSCORE were all able to discriminate outcomes after CABG with modest accuracy and varying strengths. EuroSCORE II, STS Score and AusSCORE had slightly better calibration than EuroSCORE I for 30-day mortality, but discrimination for outcomes were not superior to EuroSCORE I.

Mortality

EuroSCORE I had been a success for over a decade since its introduction as the primary international risk score for operative mortality prediction, but recent studies have a common trend that EuroSCORE I over-predicts mortality in contemporary cohorts, although outcomes discrimination remain satisfactory(28, 65). Our study also showed a trend for EuroSCORE I to over-estimate 30-day mortality while the other scores did not. The calibration drift is likely due to changing patient demographics, risk factor profiles and improving techniques in anaesthesia, surgery and post-operative intensive care(71).

Therefore, newer models such as the AusSCORE and EuroSCORE II were developed to capture trends in contemporary outcomes(30, 34).

Although EuroSCORE I, EuroSCORE II and STS Scores were statistically significant in detecting 30-day mortality in our study, their c-statistics of 0.675, 0.642 and 0.641 respectively were relatively poor compared to other studies reporting c-statistics of 0.77-0.85(66-69, 71). A first possible explanation is the ethnic differences between our cohort with the international cohort of EuroSCOREs and American cohort of STS Score, given that ethnicity may affect outcomes(31, 73). Secondly, the EuroSCOREs were derived from various cardiac operations and not strictly isolated CABG only like our cohort, which may skew the risk estimates even if modality of cardiac surgery is a parameter in the risk scores.

However in contrast to the EuroSCOREs, and similarly for the STS Score for CABG, the AusSCORE derived from an Australasian population was based on patients undergoing isolated CABG(34). This is the first study to assess its external validity. We found that the AusSCORE had similar C-statistics to the EuroSCOREs and STS Score for mortality and composite morbidity after isolated CABG but did not have improved accuracy despite being based on a population demographically closer to our cohort than the derivation populations of the other scores. The AusSCORE also gave the lowest risk prediction amongst the four scores.

Previous studies have shown that EuroSCORE I(74-76), EuroSCORE II(66) and STS Score(77) predict long-term mortality. We found that EuroSCORE I and STS Score as continuous parameters or in quintiles, but not EuroSCORE II or AusSCORE, predicted mortality during follow-up. Studies generally have found the c-statistics for these risk models for 30-day mortality to be higher than for medium or long-term mortality(66), which may partially explain our results. The relatively short follow-up time in our study and therefore the limited number of events may have also reduced the power of our analyses. Nevertheless, our results show that the STS Score was better than the EuroSCORE II and the AusSCORE for predicting mortality during a follow-up of 1-2 years after CABG, however the c-statistic was modest at 0.666, and development of risk models for long-term mortality are needed.

Morbidity

EuroSCORE I predicted the occurrence of a number of post-operative morbidities, including the duration of intensive care stay and major cardiac events(74, 77). For EuroSCORE II only one study has investigated this and found it to predict the occurrence of stroke, inotrope requirement, new dialysis, re-operation for mediastinitis and prolonged intensive care stay(66). The STS Score is unique in that it was developed to predict post-operative complications and it has separate risk models for each post-operative complication, although they have not been individually validated externally. We found similar findings

for all four scores predicting some but not all post-operative complications, as well as their composite, despite the EuroSCORE I, EuroSCORE II, and AusSCORE scores being developed for predicting 30-day mortality(27, 30, 34), which may contribute as to why their c-statistics were modest.

Further studies are required to investigate the utility of these risk scores in predicting post-operative complications, which are important problems with significant cost implications, or whether complication-specific models should be separately developed.

Limitations

This was a single-centred retrospective observational study. The population size was moderate therefore the number of adverse events especially mortality is a limitation of the accuracy of the discriminative analyses. The risk scores we assessed were derived for different reasons. The EuroSCORE I, EuroSCORE II, and AusSCORE scores were derived for the prediction of 30-day mortality and the STS score was derived for the prediction of post-operative complications. We wanted to assess their utility beyond their derivation and in particular their value in predicting longer term mortality beyond 30-days. We also thought it would be of interest to compare the scores for predicting post-operative morbidities.

3.6 Conclusion

In a contemporary cohort of patients undergoing isolated CABG the EuroSCORE II, STS Score and AusSCORE had modest improvements in calibration for 30-day mortality, as compared with EuroSCORE I. All four scores predicted some but not all post-operative complications. Revision of risk models to fit contemporary surgical outcomes is important for calibration, but the room for improvement for discriminating adverse outcomes may be limited. Given the modest c-statistics found in our analysis, there is a need to develop risk models for predicting long-term mortality and post-operative complications.

4 Comparison of four contemporary risk models at predicting mortality and morbidities after aortic valve replacement

Aortic valve replacement (AVR) is the commonest form of valve surgery, and indicated for patients with severe aortic valve disease with symptoms, impaired cardiac function or positive exercise test(39). Due to the blossoming of transcatheter aortic valve implantation over the last decade, accurate risk stratification and modelling has become critical in the decision-making of these patients for subsequent interventions. This chapter compared the performance of cardiac surgery risk scores for mortality and morbidities after AVR to add to the growing literature in this field.

Two manuscripts constitute this section, analysing the prediction of mortality and morbidities after AVR. They were published in 2015 in the Journal of Thoracic and Cardiovascular Surgery volume 149 pages 443-448 and Heart, Lung and Circulation volume 24 pages 595-601, respectively. As of September 2018, there were 12 and eight citations on google scholar respectively, including two meta-analyses(37, 38). The work in this chapter was also an oral presentation 2014 at the Cardiac Society of Australia and New Zealand Annual Scientific Meeting and the World Congress of Cardiology, and won the Royal Australasian College of Physicians Trainee Research Award in 2015.

Author list and affiliations:

Tom Kai Ming Wang[1,2], David Choi[1], Ralph Stewart[1,2], Greg Gamble[2], David Haydock[1], Peter Ruygrok[1,2]

1. Green Lane Cardiovascular Service, Auckland City Hospital, Auckland, New Zealand

2. Department of Medicine, University of Auckland, Auckland, New Zealand

Contribution of candidate: study planning and design, data collection, analysis and interpretation, writing manuscript, revision in response to co-authors and editors, conference presentations, journal submission

4.1 Mortality prediction manuscript

4.1.1 Abstract

Background: Risk stratification for aortic valve replacement (AVR) is desirable given the increased demand for intervention and the introduction of transcatheter aortic valve implantation. We compared the prognostic utility of EuroSCORE, EuroSCORE II, Society of Thoracic Surgeon's (STS) Score and an Australasian model (Aus-AVR Score) for AVR.

Methods: We retrospectively calculated the four risk scores for patients undergoing isolated AVR at Auckland City Hospital during 2005-2012, and assessed their discrimination and calibration for short and long-term mortality.

Results: A total of 620 patients were followed-up for 3.8+/-2.4 years, with operative mortality of 2.9% (18). The mean EuroSCORE, EuroSCORE II, STS Score and Aus-AVR Scores were 8.7%+/-8.3%, 3.8%+/-4.7%, 2.8%+/-2.7%, 3.2%+/-4.8%. C-statistics and 95% confidence intervals for operative mortality were 0.752 (0.652-0.852), 0.711 (0.607-0.815), 0.716 (0.593-0.837) and 0.684 (0.557-0.811). Hosmer-Lemeshow test P-values (χ^2) for calibration were 0.007 (21.1), 0.125 (12.6), 0.753 (5.0) and 0.468 (7.7), while the Brier Scores were 0.0348, 0.0278, 0.0276 and 0.0294. Independent predictors of operative mortality included critical pre-operative state, atrial fibrillation, extracardiac arteriopathy and mitral stenosis. Log-rank test P-values were all <0.001 for mortality during follow-up for all four scores by quintiles.

Conclusions: All four risk scores discriminated operative mortality after isolated AVR. The EuroSCORE had poor calibration over-estimating operative mortality, whilst the other three scores fitted well with contemporary outcomes. The STS score was the best calibrated in the highest quintile of operative risk.

4.1.2 Introduction

Surgical aortic valve replacement (AVR) is the recommended treatment for severe symptomatic aortic valve disease(78-80). The introduction of transcatheter aortic valve implantation (TAVI) makes accurate risk stratification even more important in selecting the optimal treatment modality for high risk patients with aortic stenosis(51, 78). The European System for Cardiac Operative Risk Evaluation (EuroSCORE) published in 1999 and Society of Thoracic Surgeons' (STS) Score published in 2008 are the two most widely used risk models for cardiac surgery(26, 32). Despite this, recent studies have shown that the original EuroSCORE over-estimates operative mortality for valve surgery(28), and performs less well than the STS Score(81-83). The revised EuroSCORE II was published in 2012 in order to better fit contemporary outcomes(30). Other risk models specific to AVR have been developed, including one

based on an Australasian population (Aus-AVR Score)(35). We aimed to compare the prognostic utility of EuroSCORE, EuroSCORE II, STS Score and Aus-AVR Score for detecting mortality in a contemporary isolated AVR cohort.

4.1.3   Methods

All patients undergoing isolated AVR without concomitant valve or coronary surgery from January 2005 to December 2012 at Auckland City Hospital were included. Clinical characteristics and outcomes were extracted from a prospectively collected database. The EuroSCORE II definitions were used for pre-operative characteristics(30). Angina and dyspnoea were graded using the Canadian Cardiovascular Society Classification (CCS) and the New York Heart Association Functional Classification (NYHA) respectively. Valvular stenosis or regurgitation were counted only if they were at least moderate in severity. Estimated glomerular filtration rate was calculated using the Modification of Diet and Renal Disease equation and the last serum creatinine measurement pre-operatively.

The EuroSCORE(26), EuroSCORE II(30), STS Score(32) and Aus-AVR Score(35) were calculated for all patients retrospectively following separate data collection of parameters for each score where definitions may differ. There were no significant changes in operative techniques and peri-operative care over the study period. Mortality data were checked against New Zealand's national registry up to 30 June 2013. Both operative mortality (in-hospital or within 30 days) and mortality during follow-up were pre-specified outcomes for analyses.

Mann-Whitney U test and Fisher's exact test were used for univariate analyses for continuous (presented as mean/standard deviation) and categorical variables (percentage/frequency) respectively. Kaplan-Meier curves and the log-rank (Mantel-Cox) test was used for univariate survival analyses for mortality during follow-up. Variables with $P<0.10$ in univarate analyses as well as age, gender and ethnicity are incorporated into multivariate analyses, using logistic regression to identify predictors of operative mortality. Receiver-operative characteristics analysis was used to calculate the c-statistic and 95% confidence interval (95%CI) for operative mortality respectively. C-statistics were compared for significant differences using the Hanley and McNeil test. The Fisher's exact test was used to compare the observed and predicted operative mortality for each score, and together with the Hosmer-Lemeshow test and Brier Score, used to assess calibration. Significance level was set at 0.05 and all tests were two-tailed. SPSS (Version 17.0, SPSS Inc., Chicago, IL, USA) and Prism (Version 5, GraphPad Software, San Diego, CA, USA) were used for statistical analyses. Ethics approval was attained from the regional ethics committee.

4.1.4 Results

During the 8-year study period 620 patients underwent isolated AVR at Auckland City Hospital. Baseline characteristics are listed in Table 4.1. Mean age was 64.8+/-15.5 years and 34.5% (214) were female. The mean EuroSCORE, EuroSCORE II, STS Score and Aus-AVR Score were 8.7%+/-8.3%, 3.8%+/-4.7%, 2.8%+/-2.7%, 3.2%+/-4.8%. Mechanical valves were implanted in 30.6% (190) and bioprosthetic valves in 69.4% (430).

Table 4-1 Baseline characteristics – demographics, presentation and past history

Demographics	
Age (years)	64.8 (15.5)
Female	34.5% (214)
Ethnicity	
New Zealand European	70.8% (439)
Maori or Pacific	21.1% (131)
Other	8.1% (50)
Body mass index (kg/m^2)	29.6 (11.5)
Body surface area (m^2)	1.92 (0.26)
Presentation	
New York Heart Association class	
1	21.9% (136)
2	37.3% (231)
3	27.9% (173)
4	12.9% (80)
Unstable angina class 4	2.7% (17)
Syncope	6.1% (38)
Critical pre-operative state	3.1% (19)
Operation status	
Urgent	50.6% (314)
Emergency	0.3% (2)
Past medical history	
Previous cardiac surgery	22.6% (140)
Valve surgery	14.7% (91)
Coronary artery bypass grafting	8.4% (52)
Other cardiac operation	1.8% (11)
Congestive heart failure	20.3% (126)
Myocardial infarction	8.7% (54)
Recent myocardial infarction in 90 days	2.9% (18)

Atrial fibrillation	19.2% (119)
Diabetes	17.3% (107)
Hypertension	49.0% (304)
Hypercholesterolaemia	54.0% (335)
Current smoker	10.8% (67)
Smoking history	54.4% (337)
Active infective endocarditis	10.8% (65)
Cerebrovascular accident	6.1% (38)
Extracardiac arteriopathy	6.0% (37)
Chronic pulmonary disease	19.0% (118)
Dialysis	2.4% (15)
Poor mobility	0.2% (1)
Investigations	
Ejection Fraction	
Normal ($\geq$50%)	65.0% (403)
Mild/moderate impairment (30-49%)	29.8% (185)
Severe Impairment (<30%)	5.2% (32)
Aortic stenosis	
Severe	74.0% (459)
Moderate	1.3% (8)
Aortic regurgitation	
Severe	23.5% (146)
Moderate	12.4% (77)
Mitral stenosis	
Severe	0.0% (0)
Moderate	0.5% (3)
Mitral regurgitation	
Severe	0.5% (3)
Moderate	7.6% (47)
Tricuspid regurgitation	

Severe	0.5% (3)
Moderate	4.2% (26)
Pulmonary hypertension	
Moderate (31-55mmHg)	15.6% (97)
Severe (>55mmHg)	2.7% (15)
Left main artery ≥50% stenosis	2.7% (17)
Main coronary vessels ≥50% stenosis	
0	82.9% (514)
1	8.2% (51)
2	2.4% (15)
3	6.5% (40)
Creatinine clearance (mL/min)	83 (40)
Operative mortality	2.9% (18)

Operative mortality was 2.9% (18/620). Multivariate analyses revealed critical pre-operative state, atrial fibrillation, extracardiac arteriopathy and mitral stenosis to independently predict operative mortality (table 4.2).

Table 4-2 Multivariate analysis of operative mortality (predictors P<0.10 shown)

Outcome/Predictors	Ratio	95% confidence interval	P-value
Critical pre-operative state	7.72	1.49-39.9	0.015
Atrial fibrillation	3.38	1.18-9.69	0.023
Extracardiac arteriopathy	4.11	1.04-16.1	0.043
Chronic pulmonary disease	2.85	0.955-8.53	0.060
Mitral stenosis	6.13	1.39-27.0	0.015

Table 4.3 lists the results of discrimination and calibration analyses for operative mortality. C-statistics and 95% confidence intervals for operative mortality by EuroSCORE, EuroSCORE II, STS Score and Aus-AVR Score were 0.752 (0.652-0.852), 0.711 (0.607-0.815), 0.716 (0.593-0.837) and 0.684 (0.557-0.811). There was no statistically significant difference between these c-statistics (Hanley and McNeill P-values 0.485-0.967).

In terms of calibration, EuroSCORE had the only significant statistically significant Fisher's Exact Test and Hosmer-Lemeshow test P-values of <0.001 and 0.007, while they were 0.433 and 0.125, 1.000 and 0.753, and 0.869 and 0.468 for EuroSCORE II, STS Score and Aus-AVR Score respectively. The EuroSCORE also had the highest Brier Score of 0.0348, while the others were 0.0278, 0.0276 and 0.0294 respectively. Calibration plots by quintiles of scores for operative mortality are illustrated in Figure 4.2. The EuroSCORE over-estimated operative mortality in all quintiles, with increasing discrepancy in higher quintiles. The EuroSCORE II and Aus-AVR Score also over-estimated operative mortality in the highest quintile to a moderate degree, while STS Score had good calibration in all quintiles.

Table 4-3 Discrimination and calibration analyses

Outcomes	EuroSCORE	EuroSCORE II (ESII)	STS Score (STS)	Aus-AVR Score (Aus)
Predicted operative mortality (%)	8.7 (8.3)	3.8 (4.7)	2.8 (2.7)	3.2 (4.8)
Discrimination for operative mortality				
C-statistic (95% confidence interval)	0.752 (0.652-0.852)	0.711 (0.607-0.815)	0.715 (0.593-0.837)	0.684 (0.557-0.811)
Hanley and McNeill P-value	0.671 (ESII)	0.967 (STS)	0.754 (Aus)	
	0.701 (STS)	0.785 (Aus)		
	0.485 (Aus)			
Calibration for operative mortality (2.9%)				
Observed/predicted ratio	0.33	0.77	1.05	0.90
Fisher's Exact Test P-value	<0.001	0.433	1.000	0.869
Brier Score	0.0348	0.0278	0.0276	0.0294
Hosmer-Lemeshow test P-value (χ^2)	0.007 (21.1)	0.125 (12.6)	0.753 (5.0)	0.468 (7.7)

Figure 4-1 Calibration of operative mortality observed and predicted quintiles of each risk model a) EuroSCORE, b) EuroSCORE II, c) Society of Thoracic Surgeon's Score, and d) Aus-AVR Score.
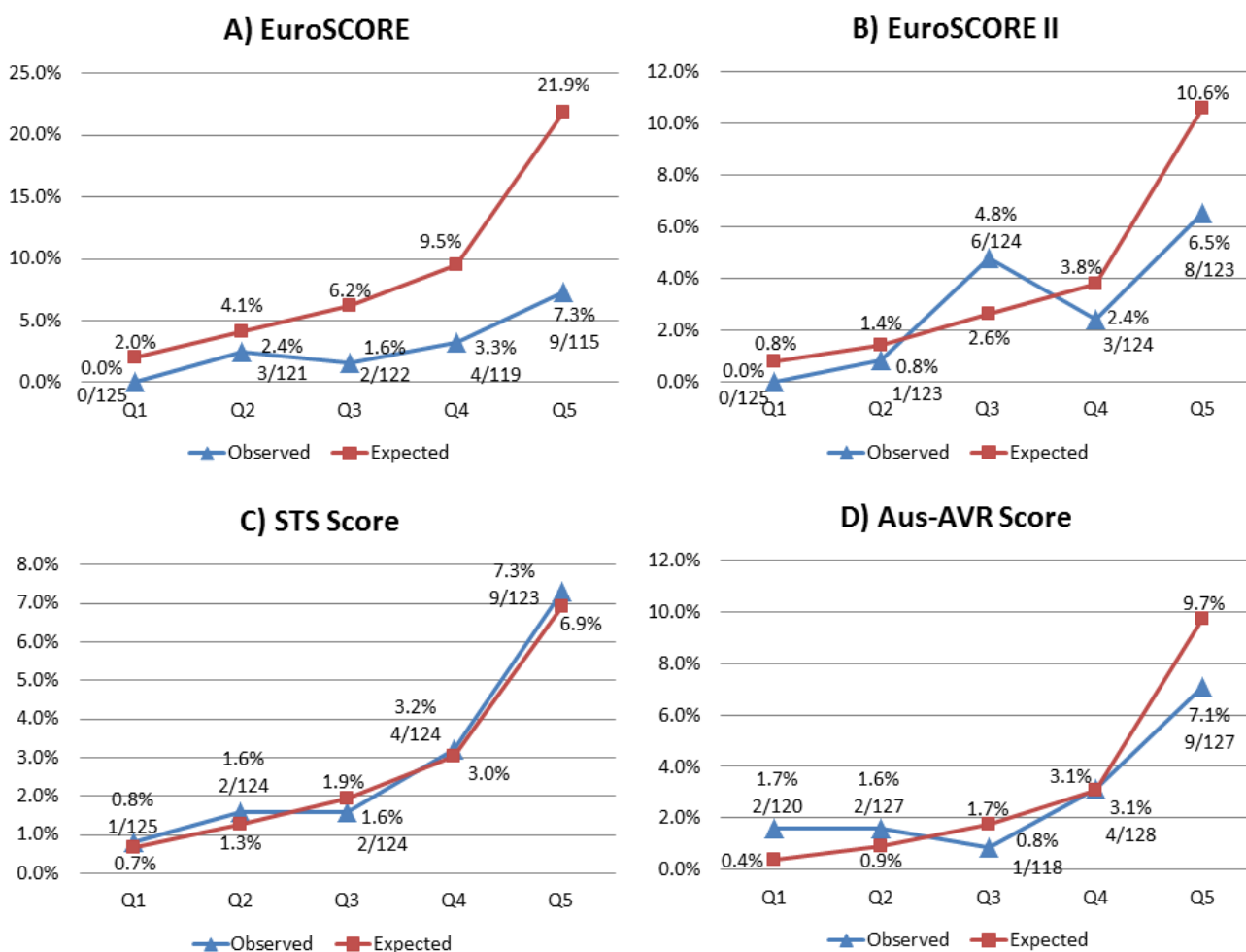
Mean follow-up period was 3.8+/-2.4 years, and 1, 3 and 5- year survivals of the cohort were 94.2%, 89.1% and 82.6%. Kaplan-Meier survival curves by quintiles of all four scores are shown in Figure 4.2. Log-rank tests were p<0.001 for all scores.

Figure 4-2 Survival curves by quintiles of each risk model a) EuroSCORE, b) EuroSCORE II, c) Society of Thoracic Surgeon's Score, and d) Aus-AVR Score.
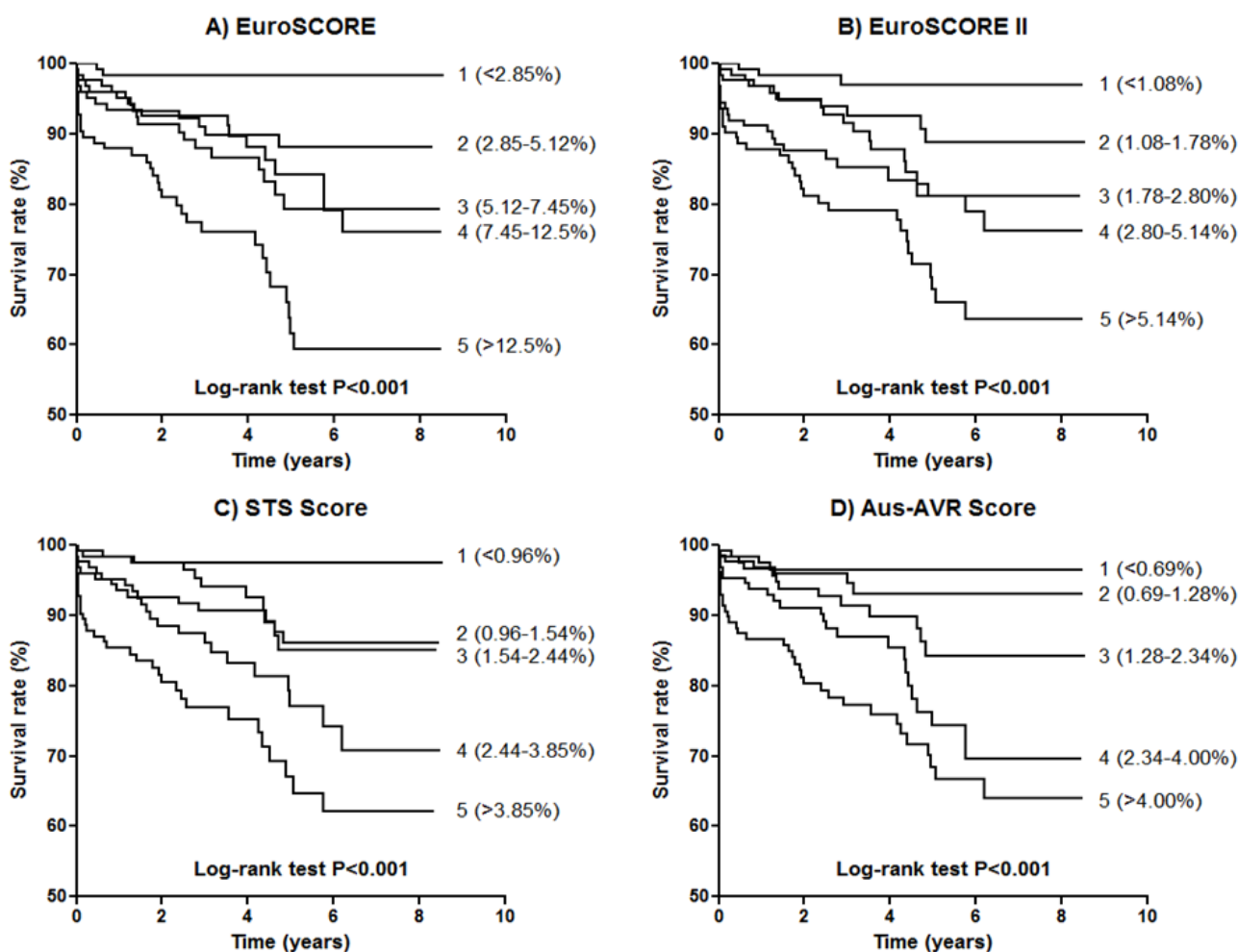
4.1.5 Discussion

There are several important findings in this study regarding risk model utility for isolated AVR. Discrimination of operative mortality was similar for all scores. The three newer scores were better calibrated to contemporary outcomes by not over-estimating operative mortality like the original EuroSCORE. In the highest quintile, the STS Score was the most accurate. We also showed that the accuracy of the Aus-AVR Score was comparable to the EuroSCORE II and STS Score.

The original EuroSCORE has been reported in recent studies to over-estimate operative mortality in cardiac surgery including isolated AVR(28), attributed to changing patient demographics and risk factor profiles and improving peri-operative care in anaesthesia, surgery and intensive care(71). This encouraged revision and development of newer risk models including those evaluated in our study(30, 32, 35). Prior to publishing EuroSCORE II, studies had identified the STS Score to have improved calibration upon EuroSCORE(81-83). More recent studies have shown that the EuroSCORE II provides

more accurate estimates of operative mortality than EuroSCORE in cardiac surgery cohorts with 13.3-18.9% of cases being isolated AVR(67, 68, 71), and one study comparing EuroSCORE, EuroSCORE II and STS Scores for isolated AVR had similar findings(84). Our observations that the EuroSCORE over-estimated operative mortality while the newer scores generally fitted better to contemporary isolated AVR outcomes were consistent with other studies.

Despite improvements in calibration, we did not find the newer scores to have improved discrimination for isolated AVR, and in fact the EuroSCORE had the highest c-statistic of 0.752 for operative mortality. Other studies also found statistically similar c-statistics between EuroSCORE and EuroSCORE II(67-69, 71, 84) and between EuroSCORE and STS Score(81, 82, 84), although like our study, statistical power may have influenced these outcomes. Further studies are required to assess whether incorporation of other potentially important parameters not traditionally incorporated into risk scores may help in this regard, such as the assessment of frailty, porcelain aorta, liver disease and right ventricular dysfunction(85).

The accuracy of risk scores to predict mortality in high risk patients is of great importance, as this is the group of patients where alternative modalities such as medical treatment or TAVI is considered, given that TAVI is currently recommended for high-risk operable or inoperable patients, which is difficult to define(78). Our study showed that in the highest quintile of predicted risk, EuroSCORE immensely over-estimated operatively mortality, EuroSCORE II and Aus-AVR Score also moderately though non-significantly over-estimated operative mortality, while STS Score was the most accurate. We therefore recommend the STS Score is utilised for risk stratification in patients referred for TAVI. Other studies generally found the EuroSCORE to significantly over-estimate in high risk patients for AVR(67, 68, 71, 81-84). Mixed results have been reported for EuroSCORE II to over-(69, 71) and under-estimate(68) risks and also for STS Score to over-(82) and under-estimate(81, 83, 84) operative mortality in high risk patients, though always to smaller degrees than the EuroSCORE.

In the Australasian population with a different ethnic distribution to the rest of the world, the EuroSCORE has also been shown to over-estimate mortality(41). This stimulated the development of regional risk models for isolated coronary artery bypass grafting(34) and isolated AVR (Aus-AVR Score)(35). Our study externally validated the latter score, showing that it has good discrimination and calibration for operative mortality comparable to the EuroSCORE II and STS Score.Of note, although the population from which each score is derived from has a different ethnic breakdown, only the STS Score includes ethnicity as one of the determinants in predicting mortality.

A number of studies have found the EuroSCORE(74-76), EuroSCORE II(66) and STS Score(86) to predict mortality during follow-up for cardiac surgery cohorts although these were mostly made up of

those undergoing coronary artery bypass grafting. Our Kaplan-Meier curve shows similar trends that these risk scores may stratify long-term mortality, with the lowest and highest risk quintiles having the best and worst survival over time. Given our limited numbers however it was underpowered for further statistical analyses to prove the scores predictive power.

Limitations

This is a single-centre observational study, focusing on isolated AVR. Retrospective calculation of risk scores based on clinical records could introduce minor biases to the values obtained. The unique ethnic breakdown of our New Zealand population may restrict the generalisability of study results, however as shown by Aus-AVR Score not being superior to other scores, the influence of ethnicity is likely to be limited. The moderate sample size and numbers of mortality events contributed to important limitations and influenced outcomes of the study in several ways: a) the absolute c-statistics in our study were generally lower than those reported in validation studies of risk scores with larger samples; b) although there were differences c-statistics and calibration parameters between scores these did not reach statistical significance to know which performed best in different aspects if there were inherently small differences between them; c) analyses of how well scores performed in subgroups could not be performed., These meant that larger studies and meta-analyses are required to confirm our findings. Additionally, follow-up time was somewhat restricted given that this was a contemporary cohort.

4.1.6 Conclusion

We found all four risk scores to discriminate operative mortality, with no statistically significant differences though this finding was influenced by the limited sample size. The EuroSCORE significantly over-estimated operative risk while the other three more recent scores had reasonable calibration in contemporary isolated AVR. In the highest quintile of risk which is important when considering other treatment modalities such as TAVI, the STS Score appears to have the best calibration whereas other scores over-estimate risk.

4.2 Morbidities prediction manuscript

4.2.1 Abstract

Background: Risk models play an important role in stratification of patients for cardiac surgery, but their prognostic utilities for post-operative complications are rarely studied. We compared the EuroSCORE, EuroSCORE II, Society of Thoracic Surgeon's (STS) Score and an Australasian model (Aus-AVR Score) for predicting morbidities after aortic valve replacement (AVR), and also evaluated seven STS complications models in this context.

Methods: We retrospectively calculated risk scores for 620 consecutive patients undergoing isolated AVR at Auckland City Hospital during 2005-2012, assessing their discrimination and calibration for post-operative complications.

Results: Amongst mortality scores, the EuroSCORE was the best at discriminating stroke (c-statistic 0.845); the EuroSCORE II at deep sternal wound infection (c=0.748); and the STS Score at composite morbidity or mortality (c=0.666), renal failure (c=0.634), ventilation>24 hours (c=0.732), return to theatre (c=0.577) and prolonged hospital stay >14 days post-operatively (c=0.707). The individual STS complications models had a marginally higher c-statistic (c=0.634-0.846) for all complications except mediastinitis, and had good calibration (Hosmer-Lemeshow test P-value 0.123-0.915) for all complications.

Conclusion: The STS Score was best overall at discriminating post-operative complications and their composite for AVR. All STS complications models except for deep sternal wound infection had good discrimination and calibration for post-operative complications.

4.2.2 Introduction

Aortic valve replacement (AVR) is the recommended treatment for severe symptomatic aortic valve disease as prognosis is significantly improved when compared to medical treatment(78, 79). Risk models play an important role in stratification as well as decision-making for the optimal treatment modality in high-risk patients, whether it be AVR, transcatheter aortic valve implantation (TAVI) or conservative medical therapy(51, 78, 85). Although several studies have compared how well contemporary risk scores predict mortality after AVR(81-84), there is a paucity of literature around the prognostic utility of these risk models for other post-operative complications which adversely impact upon subsequent quality of life.

The most widely used risk models for mortality in cardiac surgery are the European System for Cardiac Operative Risk Evaluation (EuroSCORE)(26), the revised EuroSCORE II(30), and the Society of Thoracic Surgeons' (STS) Score(32). Other scores specific to AVR have been developed including in an Australasian population (Aus-AVR Score)(35). Furthermore, the STS Score is unique in being the only score to provide seven separate risk models for individual and composite post-operative morbidities, but these have not been externally validated(32). We set out to compare the EuroSCORE, EuroSCORE II, STS Score and Aus-AVR Score at predicting post-operative complications after AVR, as well as the utility of the seven STS complication models.

4.2.3 Methods

Isolated AVR patients operated on from January 2005 to December 2012 at Auckland City Hospital were studied. Clinical characteristics and outcomes were extracted from a prospectively recorded database. Pre-operative characteristics were defined in accordance to the EuroSCORE II parameters definitions. The Canadian Cardiovascular Society Classification (CCS) and New York Heart Association Functional Classification (NYHA) were used for grouping angina and dyspnea symptoms respectively on presentation. At least moderate valvular stenosis or regurgitation needed to be present to be counted. Renal function was presented as estimated glomerular filtration rate using the Modification of Diet and Renal Disease equation and the last pre-operative serum creatinine level.

We retrospectively calculated the EuroSCORE(26), EuroSCORE II(30), STS Score(32) and Aus-AVR Score(35) for all patients. The Society of Thoracic Surgeons database definitions(32) were used for post-operative complications including:

a) permanent stroke (acute neurological deficit>24 hours due to cerebral blood supply disturbance)

b) renal failure (new dialysis requirement or increase of creatinine to >4.0mg/dL and >3 times last pre-operative level)

c) prolonged ventilation >24 hours post-operatively

d) deep sternal wound infection

e) return to theatre for any reason

f) the composite of the above five complications and/or operative mortality (in-hospital or within 30 days post-operatively)

g) prolonged hospital stay after operation >14 days.

Each of these complications were pre-specified as outcomes for analyses, as was new onset atrial fibrillation and pacemaker implantation post-operatively before discharge. We also estimated the risk of the STS complications using the specific STS AVR risk models [11].

Statistical analyses

Univariate analyses were performed using Mann-Whitney U test for continuous variables presented as mean (standard deviation) and Fisher's exact test for categorical variables presented as percentage (frequency). Multivariate analyses were conducted using variables with $P < 0.10$ in univarate analyses, by logistic regression to calculate odds ratios (OR) or Cox proportional hazards regression used to calculate hazards ratios (HR). C-statistics (area under the receiver-operative characteristics curve) with 95% confidence interval (95%CI) was used to assess discrimination of outcomes. The Hanley and McNeil test was used to compare c-statistics. The Hosmer-Lemeshow test and Brier Score were used to assess calibration. Significance level was set at 0.05 and all tests were two-tailed. SPSS (Version 17.0, SPSS Inc., Chicago, IL, USA) and Prism (Version 5, GraphPad Software, San Diego, CA, USA) were used for statistical analyses. Appropriate ethical approval was obtained before the commencement of the study

4.2.4   Results

Cohort characteristics and outcomes

A total of 620 patients underwent isolated AVR during the study period, and baseline characteristics are shown in table 4.4. Their mean age was 64.8 (15.5) years and 34.5% (214) were female. Of note, 22.6% (140) had undergone previous cardiac surgery, predominantly valve surgery in 14.7% (91). Significant aortic stenosis was present in 75.3% (467) and aortic regurgitation in 36.0% (223).

Table 4-4 Cohort characteristics

Demographics	
Age (years)	64.8 (15.5)
Female	34.5% (214)
Ethnicity	
New Zealand European	70.8% (439)
Maori or Pacific	21.1% (131)
Other	8.1% (50)
Body mass index (kg/m^2)	29.6 (11.5)
Body surface area (m^2)	1.92 (0.26)
Presentation	
New York Heart Association class	
1	21.9% (136)
2	37.3% (231)
3	27.9% (173)
4	12.9% (80)
Angina class 4	2.7% (17)
Syncope	6.1% (38)
Critical pre-operative state	3.1% (19)
Operation status	
Urgent	50.6% (314)
Emergency	0.3% (2)
Past medical history	
Previous cardiac surgery	22.6% (140)
Valve surgery	14.7% (91)
Coronary artery bypass grafting	8.4% (52)
Other cardiac operation	1.8% (11)
Congestive heart failure	20.3% (126)
Myocardial infarction	8.7% (54)
Recent myocardial infarction in 90 days	2.9% (18)

Atrial fibrillation	19.2% (119)
Diabetes	17.3% (107)
Diabetes on insulin	4.2% (26)
Hypertension	49.0% (304)
Hypercholesterolaemia	54.0% (335)
Current smoker	10.8% (67)
Smoking history	54.4% (337)
Active infective endocarditis	10.8% (45)
Cerebrovascular accident	6.1% (38)
Extracardiac arteriopathy	6.0% (37)
Chronic pulmonary disease	19.0% (118)
Dialysis	2.4% (15)
Poor mobility	0.2% (1)
Investigations	
Ejection Fraction	
Normal ($\geq$50%)	65.0% (403)
Mild/moderate impairment (30-49%)	29.8% (185)
Severe Impairment (<30%)	5.2% (32)
Aortic stenosis	75.3% (467)
Mitral stenosis	0.5% (3)
Aortic regurgitation	36.0% (223)
Mitral regurgitation	8.1% (50)
Tricuspid regurgitation	4.7% (29)
Pulmonary hypertension	
Moderate (31-55mmHg)	15.6% (97)
Severe (>55mmHg)	2.7% (15)
Left main artery $\geq$50% stenosis	2.7% (17)
Main coronary vessels $\geq$50% stenosis	
0	82.9% (514)
1	8.2% (51)

2	2.4% (15)
3	6.5% (40)
Creatinine clearance (mL/min)	83 (40)

Table 4.5 lists the mean risk scores and the STS defined outcomes. Mean EuroSCORE, EuroSCORE II, STS Score and Aus-AVR Score were 8.7+/-8.3%, 3.8+/-4.7%, 2.8+/-2.7% and 3.2+/-4.8% respectively. Composite morbidity and/or mortality occurred in 18.5% (115), and stroke in 1.3% (8).

Table 4-5 Risk scores and operative outcomes

Risk scores	
EuroSCORE (%)	8.7% (8.3%)
EuroSCORE II (%)	3.8% (4.7%)
Society of Thoracic Surgeon (STS) AVR Score (%)	2.8% (2.7%)
Aus-AVR Score (%)	3.2% (4.8%)
STS stroke model	1.3% (0.7%)
STS renal failure model	3.8% (3.4%)
STS ventilation>24 hours model	11.3% (9.1%)
STS deep sternal wound infection model	0.3% (0.2%)
STS return to theatre model	8.6% (3.4%)
STS composite mortality/morbidity model	18.1% (10.7%)
STS prolonged hospital stay after operation >14 days	8.1% (7.0%)
Outcomes	
Operative mortality	2.9% (18)
Stroke	1.3% (8)
Renal failure	4.5% (28)
Ventilation>24 hours	11.1% (69)
Deep sternal wound infection	0.8% (5)
Return to theatre	8.1% (50)
Composite mortality/morbidity	18.5% (115)
Prolonged hospital stay after operation >14 days	9.5% (59)
New atrial fibrillation	23.7% (147)
New pacemaker	4.8% (30)

Multivariate analyses of post-operative complications

Table 4.6 displays multivariate predictors of various post-operative outcomes with P<0.10. Of note, age was an independent predictor of composite morbidity and/or mortality (P=0.021) and stroke (P=0.043);

critical pre-operative state predicted composite morbidity and/or mortality (P<0.001), renal failure (P<0.001), stroke (P<0.001), ventilation >24 hours (P<0.001), deep sternal wound infection (P=0.040) and prolonged hospital stay >14 days (P=0.010); dialysis predicted composite morbidity and/or mortality (P=0.003) and ventilation>24 hours (P=0.035); history of stroke predicted renal failure (P=0.017) and stroke (P=0.031); impaired ejection fraction ventilation>24 hours (P=0.019) and prolonged hospital stay >14 days (P=0.038); and active infective endocarditis predicted ventilation>24 hours (P=0.050) and prolonged hospital stay>14 days (P=0.003) also.

Table 4-6 Multivariate analysis (all predictors P<0.10)

Outcome/Predictors	Odds ratio	95% confidence interval	P-value
Composite morbidity/mortality			
Age (per 1 year)	1.02	1.00-1.04	0.021
Body mass index (per 1 kg/m^2)	0.967	0.931-1.01	0.094
Critical pre-operative state	8.98	2.93-27.5	<0.001
Urgent or emergency surgery	1.62	1.02-2.58	0.043
Atrial fibrillation	1.68	0.994-2.83	0.053
Dialysis	7.89	2.06-30.3	0.003
Renal failure			
Male	3.23	1.06-9.84	0.039
Critical pre-operative state	6.57	2.21-19.5	<0.001
Hypertension	3.98	1.48-10.7	0.006
History of cerebrovascular accident	3.90	1.28-11.9	0.017
Stroke			
Age (per 1 year)	1.10	1.00-1.20	0.043
Syncope	6.07	1.03-35.8	0.047
Critical pre-operative state	7.30	2.08-25.6	<0.001
History of cerebrovascular accident	7.23	1.20-43.6	0.031
Ventilation>24 hours			
Body mass index (per 1 kg/m^2)	0.934	0.884-0.986	0.013
Critical pre-operative state	8.21	2.15-31.4	<0.001
Atrial fibrillation	2.48	1.32-4.69	0.005
Active infective endocarditis	2.61	1.00-6.83	0.050

Chronic pulmonary disease	1.81	0.929-3.52	0.081
Dialysis	4.08	1.11-15.0	0.035
Ejection fraction (per category)	1.47	1.07-2.03	0.019
Deep sternal wound infection			
Critical pre-operative state	12.3	1.13-135	0.040
History of coronary artery bypass grafting	10.5	1.36-81.1	0.024
Return to theatre			
Body mass index	0.856	0.908-1.01	0.090
Angina class 4	5.25	1.25-14.3	0.020
Prolonged hospital stay after operation >14 days			
Critical pre-operative state	4.43	1.43-13.7	0.010
Active infective endocarditis	3.38	1.52-7.51	0.003
Extracardiac arteriopathy	2.23	0.897-5.55	0.084
Ejection fraction (per category)	1.37	1.02-1.85	0.038
New atrial fibrillation			
Age (per 1 year)	1.02	1.00-1.04	0.023
New York Heart Association Class 4	1.79	1.04-3.08	0.036
Hypertension	1.56	1.04-2.35	0.033
Mitral regurgitation	2.30	1.20-4.42	0.012

Comparison of mortality risk models for discriminating complications

Results of receiver-operative characteristics of the EuroSCORE, EuroSCORE II, STS Score and Aus-AVR Score for detecting post-operative complications are presented in Table 4.7. All four scores could detect composite morbidity/mortality, ventilation>24 hours and prolonged hospital stay>14 days. The EuroSCORE was the best at discriminating stroke (c-statistic 0.845); the EuroSCORE II at deep sternal wound infection (c=0.748); and the STS Score at composite morbidity or mortality (c=0.666), renal failure (c=0.634), ventilation>24 hours (c=0.732), return to theatre (c=0.577) and prolonged hospital stay >14 days post-operatively (c=0.707).

Table 4-7 Receiver-operative characteristics analyses: c-statistic (95% confidence interval)

Outcomes	EuroSCORE	EuroSCORE II	STS Score	Aus-AVR Score
Composite morbidity/mortality	0.653 (0.597-0.710)	0.649 (0.592-0.706)	0.666 (0.609-0.722)	0.618 (0.559-0.676)
Stroke	0.845 (0.783-0.907)	0.770 (0.633-0.908)	0.812 (0.771-0.854)	0.642 (0.464-0.821)
Renal failure	0.599 (0.487-0.711)	0.614 (0.501-0.727)	0.634 (0.524-0.743)	0.599 (0.491-0.707)
Ventilation>24 hours	0.727 (0.664-0.790)	0.726 (0.661-0.791)	0.732 (0.668-0.797)	0.675 (0.603-0.746)
Deep sternal wound infection	0.675 (0.428-0.921)	0.748 (0.512-0.984)	0.666 (0.422-0.910)	0.502 (0.206-0.798)
Return to theatre	0.556 (0.477-0.636)	0.566 (0.484-0.649)	0.577 (0.503-0.651)	0.560 (0.474-0.645)
Prolonged hospital stay after operation >14 days	0.672 (0.603-0.741)	0.675 (0.603-0.746)	0.707 (0.640-0.773)	0.678 (0.608-0.747)
New atrial fibrillation	0.571 (0.519-0.624)	0.558 (0.505-0.611)	0.577 (0.526-0.628)	0.603 (0.551-0.655)
New pacemaker	0.449 (0.333-0.564)	0.448 (0.347-0.549)	0.480 (0.365-0.595)	0.528 (0.424-0.632)

Table 4.8 lists the discrimination and calibration of the seven STS AVR risk models for complications. All these risk models except that for deep sternal wound infection are able to discriminate their corresponding complication with slightly higher c-statistic then the STS mortality score (c=0.634-0.846) with good calibration (observed/expected ratio 0.8-1.2, Hosmer-Lemeshow test P-value 0.123-0.915). Figure 4.3 illustrates the calibration plots of these models and their complications in quintiles.

Table 4-8 Discrimination and calibration analyses of STS morbidities risk models

Outcomes	Analyses	Results
Composite morbidity and mortality	Observed/predicted ratio C-statistic (95% confidence interval) Hanley and McNeill P-value Brier Score Hosmer-Lemeshow test P-value (χ^2)	18.5%/18.1%=1.03 0.686 (0.632-0.741) 0.825 (ES), 0.805 (ESII), 0.893 (STS), 0.650 (Aus) 0.1386 0.369 (8.7)
Stroke	Observed/predicted ratio C-statistic (95% confidence interval) Hanley and McNeill P-value Brier Score Hosmer-Lemeshow test P-value (χ^2)	1.3%/1.3%=0.98 0.845 (0.752-0.939) 1.000 (ES), 0.568 (ESII), 0.795 (STS), 0.140 (Aus) 0.0125 0.770 (4.9)
Renal failure	Observed/predicted ratio C-statistic (95% confidence interval) Hanley and McNeill P-value Brier Score Hosmer-Lemeshow test P-value (χ^2)	4.5%/3.8%=1.20 0.695 (0.599-0.791) 0.235 (ES), 0.316 (ESII), 0.450 (STS), 0.235 (Aus) 0.0424 0.666 (5.8)
Ventilation>24 hours	Observed/predicted ratio C-statistic (95% confidence interval) Hanley and McNeill P-value Brier Score Hosmer-Lemeshow test P-value (χ^2)	11.1%/11.3%=0.99 0.747 (0.683-0.812) 0.691 (ES), 0.677 (ESII), 0.765 (STS), 0.160 (Aus) 0.0865 0.811 (4.5)
Deep sternal wound infection	Observed/predicted ratio C-statistic (95% confidence interval) Hanley and McNeill P-value Brier Score Hosmer-Lemeshow test P-value (χ^2)	0.81%/0.26%=3.07 0.605 (0.355-0.855) 0.712 (ES), 0.441 (ESII), 0.748 (STS), 0.582 (Aus) 0.0080 0.778 (4.8)
Return to theatre	Observed/predicted ratio	8.1%/8.6%=0.94

	C-statistic (95% confidence interval)	0.634 (0.553-0.716)
	Hanley and McNeill P-value	0.207 (ES), 0.271 (ESII), 0.357 (STS), 0.231 (Aus)
	Brier Score	0.0730
	Hosmer-Lemeshow test P-value (χ^2)	0.915 (3.3)
Prolonged hospital stay after operation >14 days	Observed/predicted ratio	9.5%/8.1%=1.17
	C-statistic (95% confidence interval)	0.738 (0.672-0.805)
	Hanley and McNeill P-value	0.233 (ES), 0.255 (ESII), 0.572 (STS), 0.278 (Aus)
	Brier Score	0.0812
	Hosmer-Lemeshow test P-value (χ^2)	0.123 (12.7)

Figure 4-3 Calibration plots of post-operative complications observed and predicted by Society of Thoracic Surgeon's morbidity risk score quintiles
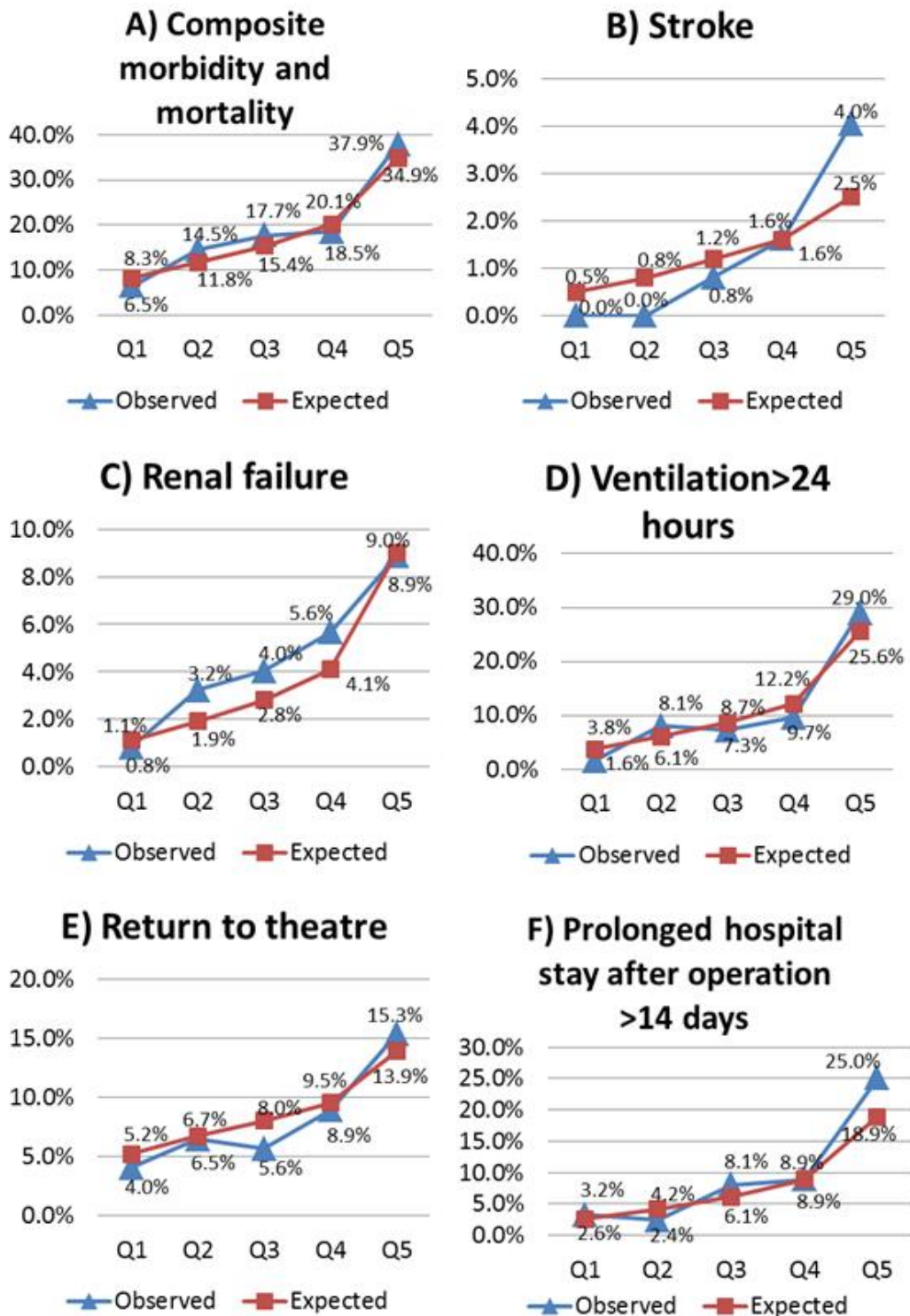
4.2.5 Discussion

Accurate risk prediction of morbidities after AVR is important as complications can be significantly debilitating, impairing quality of life. The rates of various complications differ between AVR and TAVI thereby influencing decision-making in high-risk patients(51, 85). Our study showed that the four mortality risk models were all able to detect several post-operative complications, with the STS Score having the highest c-statistic in most instances. The individual STS complication models however not only have good discrimination but also calibration for their corresponding outcomes.

Prolonged ventilation, intensive care and hospital stay post-operatively are important adverse outcomes because of their associations with increased mortality, costs and reduced quality of life(87), but also higher incidences following isolated AVR than TAVI(51, 85). C-statistics were 0.67-0.73 for the four mortality risk scores at predicting ventilation>24 hours in our study, comparable to 0.66-0.80 reported in the literature(66, 88, 89). The good discrimination of these outcomes reflect common predictors with mortality after cardiac surgery such as age, heart failure, renal impairment, chronic respiratory disease, peripheral vascular disease, critical pre-operative state and urgent surgery(26, 30, 32, 90). The STS prolonged ventilation score was also accurate for the actual incidence of our cohort so should be more widely used.

Stroke is potentially the most debilitating morbidity of cardiac surgery with long-term functional and cost implications, although it occurs at a slightly higher incidence following TAVI than AVR(51, 85). C-statistics in our cohort for the EuroSCORE, EuroSCORE II and STS Score for stroke were the highest at 0.77-0.85 amongst outcomes, and also compared to the 0.63-0.77 reported in other studies(66, 77, 91). Independent predictors of stroke frequently found include a history of cerebrovascular disease, impaired ejection fraction, peripheral vascular disease and cardiopulmonary or cross-clamp times(92, 93). We also demonstrated that the STS score for stroke had good calibration with observed incidence so should be utilised for considering treatment modality and giving a risk figure for patients.

For post-operative renal failure, the EuroSCORE and EuroSCORE II have been shown to have good discrimination c=0.65-0.87(66, 77, 91, 94), which was lower in our cohort at c=0.60-0.63 and only the EuroSCORE II and STS Score reached statistical significance. This is consistent with the fact that most of the renal failure predictors we identified (male, hypertension, history of stroke) are not common parameters of cardiac surgery risk models except the STS score. The presence of these and other independent predictors previously identified such as urgent surgery, heart failure or cardiogenic shock and diabetes(95, 96) should alert clinicians to the higher risk of renal failure and precautions such as maintaining blood pressure peri-operatively should be undertaken.

Deep sternal wound infection, or mediastinitis, is another devastating complication associated with reoperation, prolonged stay, higher mortality and costs(97). Although c-statistics of the EuroSCORE, EuroSCORE II and STS Score of our cohort were reasonable at 0.66-0.75 for mediastinitis and only statistically significant for the EuroSCORE II. The wide confidence intervals for this outcome are in part due to low incidence of 0.8% (5/620) in our cohort. Most studies (c=0.70-0.76)(66, 91, 98), except 1 (c=0.54)(77), however, have found the EuroSCORE, EuroSCORE II and STS to detect mediastinitis with statistical significance, although study numbers were greater at n=800-11,000. We found critical pre-operative state and previous coronary surgery to predict mediastinitis, and other predictors identified in the literature include diabetes, peripheral vascular disease, chronic respiratory disease and dialysis(97-99). Strategies to reduce mediastinitis are very important especially in those at high risk and should be implemented, including antibiotic prophylaxis, strict peri-operative glycaemic control and optimal skin preparation.

Our findings suggest although mortality risk scores can be used to detect many post-operative complications, the STS score had the best performance across the complications assessed. The individual STS complications score however had good calibration with at least equal if not better discrimination for their respective outcomes, so should be used preferentially. This advocates development of complications and operation specific models. The obvious disadvantage is the complexity of all STS risk models with more parameters than other scores, but if this is overcome then they should be routinely used in risk stratification of cardiac surgery candidates.

Limitations

This is a single-centre retrospective observational study. The moderate sample size and power mean small but significant differences may not always be detected, particularly differences in c-statistics. Retrospective calculation of risk scores based on clinical records could introduce minor biases to the values obtained. The demographically unique characteristics of our AVR cohort may impede to some extent the generalisability of our results. Further studies are required to see how well risk scores predict post-operative complications in other forms of cardiac surgery.

4.2.6   Conclusion

All four contemporary mortality risk scores for cardiac surgery were able to detect at least half of the post-operative morbidities, with the STS score having the highest c-statistic for composite end-point and four other complications. All the STS complication models, except that for deep sternal wound infection, fitted well to observed rates with incrementally higher c-statistics to the STS score, therefore we recommend using these to predict risk of complications after AVR.

5 Comparing performance of risk scores for combined aortic valve replacement and coronary bypass grafting surgery

A significant proportion of patients with severe aortic valve disease have significant coronary artery disease and vice versa(39). A combined operation however has additive risk on rates of adverse outcomes, and utility of risk models has not been well evaluated in this context in the past. This chapter is amongst the first to specifically evaluate the accuracy of contemporary risk models at predicting outcomes for the combined aortic valve replacement and coronary artery bypass grafting operation, which is not infrequently performed.

This manuscript was published in 2016 in Heart, Lung and Circulation volume 25 pages 1118-1123. As of September 2018, it had five citations on Google scholar. It was a poster presentation at both the Cardiac Society of Australia and New Zealand Annual Scientific Meeting and the European Society of Cardiology Congress in 2015.

Author list and affiliations:

Tom Kai Ming Wang[1,2], David Choi[1], Tharumenthiran Ramanathan[1], Peter Ruygrok[1,2]

1. Green Lane Cardiovascular Service, Auckland City Hospital, Auckland, New Zealand

2. Department of Medicine, University of Auckland, Auckland, New Zealand

Contribution of candidate: study planning and design, data collection, analysis and interpretation, writing manuscript, revision in response to co-authors and editors, conference presentations, journal submission

5.1 Abstract

Background: Aortic valve replacement (AVR) and/or coronary artery bypass grafting (CABG) make up the majority of cardiac surgery with increasing demand as the population ages. Accuracy of risk stratification is important, especially as interventional aortic valve and coronary procedures continue to emerge, but have been rarely studied for the combined AVR+CABG operation. We compared the prognostic utility of EuroSCORE, EuroSCORE II and Society of Thoracic Surgeon's (STS) Score for AVR+CABG.

Methods: All patients (n=450) undergoing AVR+CABG at Auckland City Hospital during 2005-2012 with mean follow-up of 4.7+/-2.5 years were included. The three risk scores were calculated and their discrimination and calibration for mortality and morbidities assessed.

Results: Operative mortality was 6.4% (29), and mean scores were EuroSCORE 12.5+/-11.1%, EuroSCORE II 6.6+/-6.1% and STS Score 5.5+/-4.4%. C-statistics were 0.587, 0.669 and 0.699 respectively for operative mortality, Hosmer-Lemeshow test P-values were 0.064, 0.718 and 0.567, and Brier Score 0.716, 0.585 and 0.588. Independent predictors of operative mortality were history of myocardial infarction and impaired renal function. STS score also was the most accurate score for predicting mortality during follow-up (c=0.663), composite morbidity (c=0.627), stroke (c=0.642), prolonged ventilation>24 hours (c=0.642), and return to theatre (c=0.612).

Conclusion: The STS score has the best discriminative ability for mortality and the majority of complications after AVR+CABG, while its calibration was similar to EuroSCORE II and superior to EuroSCORE. It should therefore be used for risk stratification and when considering surgical versus percutaneous intervention in those with concurrent aortic valve and coronary artery disease.

5.2 Introduction

Coronary artery disease affects a third of patients undergoing aortic valve replacement (AVR) indicated for severe symptomatic aortic valve disease, with guidelines recommending concurrent coronary artery bypass grafting (CABG) with AVR(61, 78). Demand for this combined operation continues to increase as the population ages, but is associated with increased risk of mortality and morbidities(33, 100-102). Furthermore, the introduction of transcatheter aortic valve implantation (TAVI), which may be performed with percutaneous coronary intervention, makes accurate risk stratification even more important in deciding the optimal treatment modality for high risk patients with aortic stenosis and coronary artery disease(50-52).

The most widely used risk models in cardiac surgery include the European System for Cardiac Operative Risk Evaluation (EuroSCORE) and Society of Thoracic Surgeons' (STS) Score, the latter having a specific model for combined valve replacement and CABG(27, 33). The revised EuroSCORE II was more recently published to better fit contemporary outcomes(30). We compared the prognostic utility of EuroSCORE, EuroSCORE II and STS Score at predicting mortality and morbidity in a contemporary cohort of patients undergoing combined AVR and CABG surgery.

5.3 Methods

Consecutive patients undergoing combined AVR and CABG operation during 2005-2012 at Auckland City Hospital were included. Relevant clinical characteristics were collected from computerised records. Logistic EuroSCORE(27), EuroSCORE II(30) and STS Score(33) were retrospectively calculated for all patients using available data.

Characteristics were defined based on the EuroSCORE II risk model, including critical pre-operative state, extracardiac arteriopathy, chronic lung disease, poor mobility and categories left ventricular ejection fraction, renal impairment and pulmonary hypertension(30). Dyspnoea was graded by the New York Heart Association Functional Classification (NYHA), and angina by the Canadian Cardiovascular Society Classification (CCS). Hypertension was defined as prescribed medications for lowering blood pressure, any measurement of over 140/90mmHg prior to operation and/or a previous formal diagnosis. Stroke included any previous history of a neurological deficit that persisted over 24 hours and caused by disturbance of cerebral blood supply. Operative characteristics collected include valve replacement type, number of grafts and duration of cardiopulmonary bypass and aortic cross-clamp times.

Mortality data was checked against New Zealand's national registry until 31 December 2014. Operative mortality includes deaths within 30 days of operation and/or during the same hospital admission of the operation. Post-operative complications (stroke, renal failure, ventilation >24 hours, deep sternal wound infection, return to theatre and prolonged hospital stay>14 days) followed the Society of Thoracic Surgeon's (STS) score definitions, and their composite was also determined(33).

Quantitative and categorical variables are presented as mean (standard deviation) and percentages (frequency) respectively. Discriminative ability of post-operative outcomes was assessed using the area under receiver-operative characteristics curves (c-statistic) with 95% confidence interval (95%CI) reported. Calibration was assessed by observed/expected ratio, Hosmer-Lemeshow goodness-of-fit test and Brier Score. Kaplan-Meier curves and log-rank (Mantel-Cox) test were used for longitudinal survival analysis, stratifying each risk score into quartiles. Statistical significance was defined as P-value less than 0.05 and all tests were two-tailed. Statistical analyses were performed using SPSS (Version 17.0, SPSS

Inc., Chicago, IL, USA) and Prism (Version 5, GraphPad Software, San Diego, CA, USA). Ethics approval was obtained from our institution's ethics review committee.

5.4 Results

A total of 450 patients underwent AVR+CABG surgery during 2005-2012 comprising the study cohort. Table 5.1 lists the baseline characteristics. Mean age was 73.0 (8.8) years and 25.1% (113) were female. Mean scores were EuroSCORE 12.5 (11.1)%, EuroSCORE II 6.6 (6.1)% and STS Score 5.5 (4.4)%. Table 5.2 shows the operative variables and post-operative outcomes.

Table 5-1 Baseline characteristics

Demographics	
Age (years)	73.0 (8.8)
Female	25.1% (113)
Ethnicity	
New Zealand European	80.7% (363)
Maori or Pacific	14.0% (63)
Other	5.3% (24)
Body mass index (kg/m^2)	28.5 (5.0)
Body surface area (m^2)	1.89 (0.21)
Presentation	
New York Heart Association class	
1	35.6% (160)
2	32.0% (144)
3	21.3% (96)
4	11.1% (50)
Unstable angina class 4	10.4% (47)
Syncope	5.6% (25)
Critical pre-operative state	2.7% (12)
Inpatient urgent operation	66.4% (299)
Past medical history	
Previous cardiac surgery	8.9% (40)

Valve surgery	4.7% (21)
Coronary artery bypass grafting	6.4% (29)
Other cardiac operation	3.3% (15)
Congestive heart failure	16.0% (72)
Myocardial infarction	37.1% (167)
Recent myocardial infarction in 90 days	21.8% (98)
Percutaneous coronary intervention	12.7% (57)
Atrial fibrillation	19.8% (89)
Diabetes	28.0% (126)
Hypertension	68.0% (306)
Hypercholesterolaemia	80.7% (363)
Current smoker	7.3% (33)
Smoking history	58.4% (263)
Active infective endocarditis	1.3% (6)
Cerebrovascular accident	5.8% (26)
Extracardiac arteriopathy	17.1% (77)
Chronic pulmonary disease	20.0% (90)
Dialysis	3.1% (14)
Poor mobility	0.2% (1)
Investigations	
Ejection fraction	
Normal ($\geq$50%)	69.1% (311)
Mild/moderate impairment (30-49%)	14.9% (67)
Severe Impairment (<30%)	16.0% (72)
Aortic stenosis	91.6% (412)
Mitral stenosis	1.1% (5)
Aortic regurgitation	16.0% (72)
Mitral regurgitation	4.9% (22)
Tricuspid regurgitation	1.3% (6)
Pulmonary hypertension	
Moderate (31-55mmHg)	19.8% (89)

62

Severe (>55mmHg)	3.1% (14)
Left main artery ≥50% stenosis	19.8% (89)
Main coronary vessels ≥50% stenosis	
1	27.9% (126)
2	26.6% (120)
3	45.2% (204)
Creatinine clearance (mL/min)	66 (28)

All figures are mean (standard deviation) or percentage (frequency)

Table 5-2 Operative variables and post-operative outcomes

Operative variables	
Mechanical valve	15.8% (71)
Number of grafts	2.3 (1.5)
Cardiopulmonary bypass time (minutes)	154 (45)
Cross-clamp time (minutes)	121 (34)
In-hospital outcomes	
Composite morbidity	33.8% (152)
Stroke	3.3% (15)
Renal Failure	1.3% (6)
Ventilation>24 hours	27.6% (124)
Deep sternal wound infection	1.3% (6)
Return to theatre	12.0% (54)
Prolonged hospital stay >14 days	19.1% (86)
Operative mortality	6.4% (29)

All figures are mean (standard deviation) or percentage (frequency)

Operative mortality was 6.4% (29). Discrimination and calibration analyses for mortality are shown in Table 5.3. C-statistics with 95%CI of 0.669 (0.571-0.767) for EuroSCORE II and 0.699 (0.607-0.791) for STS score were significantly higher than chance, but not 0.587 (0.477-0.698) for EuroSCORE.

In terms of calibration, the Hosmer-Lemeshow test for detecting operative mortality was approaching statistical significance p=0.064 for EuroSCORE, but non-significant p=0.718 for EuroSCORE II and

p=0.567 for STS Score. EuroSCORE also had a high observed/predicted ratio and Brier Score compared to the other two scores.

Table 5-3 Discrimination and calibration analyses for risk scores and mortality

	EuroSCORE	EuroSCORE II	STS Score
Predicted score	12.5 (11.1)%	6.6 (6.1)%	5.5 (4.4)%
Operative mortality			
C-statistic	0.587 (0.477-0.698)	0.669 (0.571-0.767)	0.699 (0.607-0.791)
Observed/predicted ratio	1.9	1.02	0.85
Hosmer-Lemeshow Test	χ^2=11.7, p=0.064	χ^2=5.4, p=0.718	χ^2=6.7, p=0.567
Brier Score	0.716	0.585	0.588
Mortality during follow-up			
C-statistic	0.608 (0.538-0.678)	0.623 (0.555-0.690)	0.663 (0.599-0.727)

Mean (standard deviation) and c-statistic (95% confidence interval)

Composite morbidity occurred in 33.8% (152), including stroke 3.3% (15) and prolonged ventilation >24 hours 27.6% (124). Discrimination analyses for post-operative complications are shown in table 5.4. All three scores detected composite morbidity, prolonged ventilation>24 hours and prolonged hospital stay >14 days, however only the STS score detected stroke (c=0.642) and return to theatre (c=0.612).

Table 5-4 Discrimination analyses for risk scores and post-operative complications

Outcome	EuroSCORE	EuroSCORE II	STS Score
Composite morbidity	0.586 (0.530-0.642)	0.611 (0.555-0.666)	0.627 (0.573-0.682)
Stroke	0.574 (0.413-0.736)	0.623 (0.491-0.755)	0.642 (0.498-0.786)
Renal failure	0.267 (0.164-0.369)	0.160 (0.076-0.244)	0.318 (0.190-0.445)
Ventilation>24 hours	0.618 (0.559-0.678)	0.629 (0.570-0.689)	0.642 (0.584-0.700)
Deep sternal wound infection	0.575 (0.345-0.805)	0.654 (0.455-0.853)	0.631 (0.441-0.820)
Return to theatre	0.534 (0.454-0.615)	0.566 (0.493-0.650)	0.612 (0.534-0.690)
Prolonged hospital stay >14 days	0.583 (0.519-0.646)	0.605 (0.541-0.670)	0.638 (0.576-0.700)

C-statistic (95% confidence interval)

Mean follow-up was 4.7 (2.5) years, and survivals were 90.9%, 85.7% and 75.7% at 1, 3 and 5 years respectively. Figure 5.1 illustrates the overall cohort survival and survival by quartiles of each risk score.

All three scores detected mortality during follow-up with c-statistics (95% CI) and log-rank test p-value of 0.608 (0.538-0.678) and 0.003 respectively for EuroSCORE, 0.623 (0.555-0.690) and 0.003 for EuroSCORE II and 0.663 (0.599-0.727) and <0.001 for STS Score.

Figure 5-1 Kaplan-Meier survival curves for A) entire cohort, B) EuroSCORE quartiles, C) EuroSCORE II quartiles and D) STS Score quartiles (with log-rank test p-values)
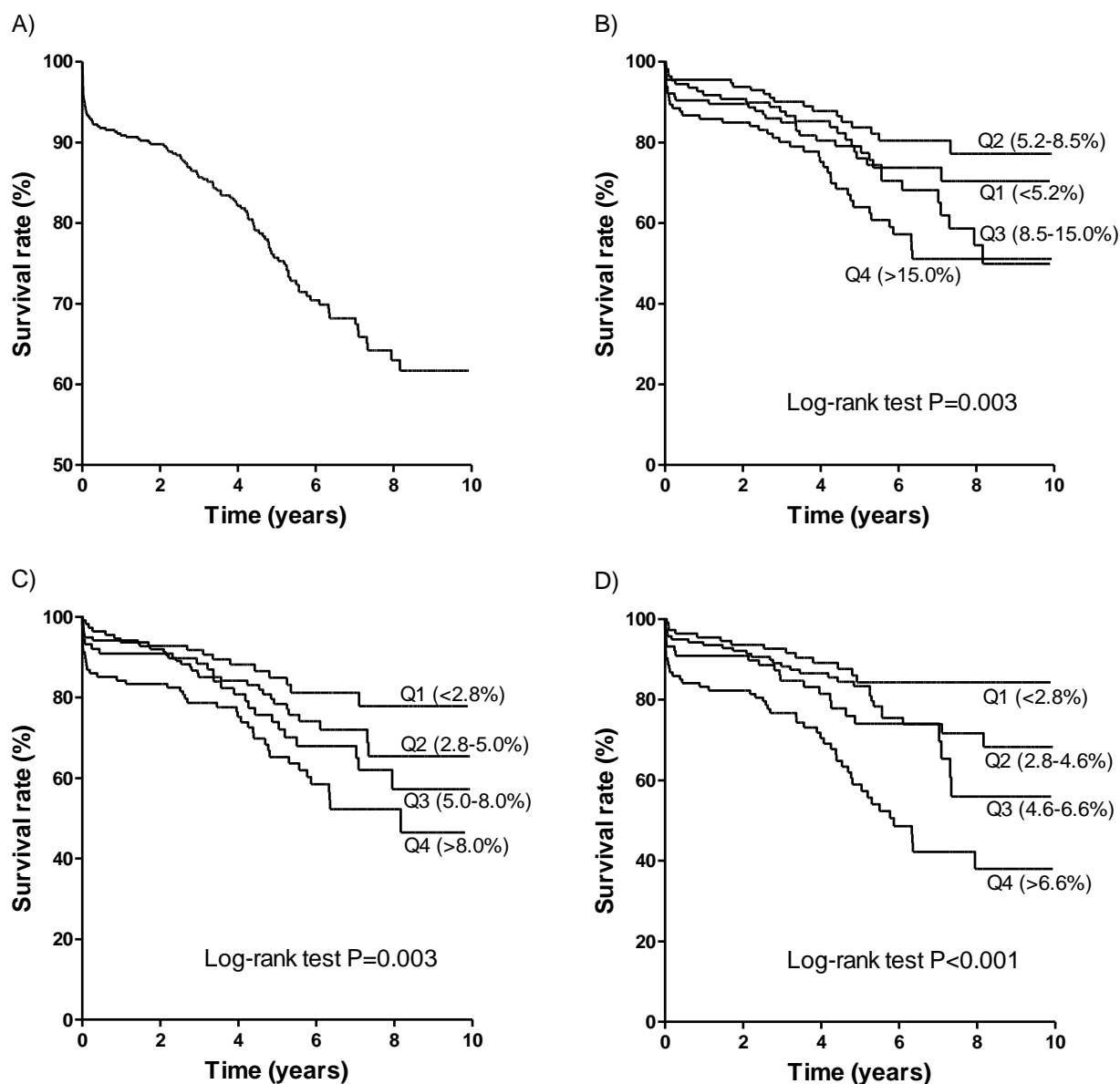
5.5 Discussion

Our study has provided several important findings for the prognostic utility of conventional risk models in AVR+CABG. Firstly, the original EuroSCORE was not able to discriminate and appeared to over-estimate operative mortality, whereas EuroSCORE II and STS Score had improved discrimination and calibration. Secondly, all scores were able to discriminate, albeit with modest accuracy, mortality during

follow-up. Lastly, they also all discriminated composite morbidity, prolonged ventilation>24 hours and return to theatre but were less applicable to other complications. The STS Score had the highest c-statistic discrimination for nearly all outcomes.

Up until recently, the EuroSCORE has been the most widely used risk model for cardiac surgery internationally both clinically and in research. Other studies, including those evaluating AVR+CABG specifically found EuroSCORE to discriminate operative mortality well, c-statistic 0.67-0.79 with statistical significance, however we did not find this(28, 67, 70, 103). On the other hand, recent studies of AVR+CABG including ours, just like other cardiac surgeries, uniformly found EuroSCORE to over-estimate operative mortality, although to varying degree with observed-to-predicted ratio of 1.2-2.4 (67, 70, 103). This reflects changing patient demographics and improving operative peri-operative care and outcomes, and support the need for risk models tobe updated overtime for enhanced calibration(71).

EuroSCORE II also have been reported to have good discrimination of AVR+CABG operative mortality in other studies with c-statistic 0.67-0.77 which are not improved from EuroSCORE(71). The main difference lies in calibration with reported observed-to-predicted ratio of 0.73-1.0, which is generally an improvement over EuroSCORE in contemporary cohorts, though caution with under-estimation was reported in one study(70). As a result, recent myocardial revascularsiation guidelines have advised the use of EuroSCORE II rather than EuroSCORE for risk stratification for CABG, and our results suggest this proposition can be extended to AVR+CABG (61).

Compared to EuroSCORE and EuroSCORE II, STS has been less studied for cardiac surgery and not solely for AVR+CABG, though frequently give similar if not improved predictive value(61). Notably, STS has more parameters in its model, which may lead to better accuracy but makes calculating the score more cumbersome. Similar to our isolated AVR cohort we found STS to have the best discrimination for operative mortality(2). Although it slightly under-estimated operative mortality, this did not reach statistical significance.

Each of the three risk scores have also been found to discriminate long-term mortality in cardiac surgery though again not specifically for AVR+CABG(66, 74, 77). We found all three scores as continuous parameters or in quartiles to discriminate mortality during follow-up, again STS having the highest c-statistic and just like our isolated AVR but not isolated CABG cohort, supporting the use of STS Scores in this context as well.

Accurate risk estimates of complications are also important as adverse outcomes can significantly reduce function and quality of life, however despite this, no studies have investigated this in context of AVR+CABG surgery. A number of studies have found all three scores also to predict all the various STS

defined morbidities in other cardiac operations, though generally not as good as predicting mortality(66, 74, 77, 89). This is expected given that these scores were designed to estimate operative mortality rather than morbidities. Our other studies of isolated CABG or isolated AVR also found EuroSCOREs and STS Score to discriminate composite morbidity, prolonged ventilation>24 hours, and more variably stroke, renal failure and return to theatre just like AVR+CABG(1, 3). These findings suggest the risk factors predicting operative mortality are similar to those predicting these complications.

In terms of limitations, this is a single-centre observational study. Retrospective calculation of risk scores based on clinical records could introduce minor biases to the values obtained. Sample size and number of adverse events were moderate limiting the power of statistical analyses, including for discrimination and calibration to detect significant differences. Follow-up time was also somewhat restricted given that this was a contemporary cohort.

5.6 Conclusion

In conclusion, the EuroSCORE did not discriminate and appeared to over-estimate operative mortality, whereas the EuroSCORE II and STS score could detect this with good calibration. All scores could detect mortality during follow-up and composite morbidity, however STS appeared to have the best discriminative ability. Based on our findings, we suggest the STS score to be the model of choice in the risk stratification of intervention in patients with concurrent severe aortic valve and coronary artery disease.

6   Performance of contemporary surgical risk scores for mitral valve surgery

Mitral valve surgery, in the form of repair or replacement, is the second commonest form of valve surgery, and recommended for patients with severe mitral valve regurgitation or stenosis and symptoms, impaired cardiac function and sometimes new atrial fibrillation and/or pulmonary hypertension(39). Percutaneous mitral valve intervention are under research and development and have started to gain popularity in high risk patients in some overseas centres (104). This chapter evaluated cardiac surgery risk models' prognostic accuracy for outcomes after mitral valve surgery.

This manuscript was published in 2017 in the Journal of Cardiac Surgery volume 32 pages 172-176. As of September 2018, it had three citations on Google scholar. It was an oral presentation at both the Cardiac Society of Australia and New Zealand Annual Scientific Meeting and the European Society of Cardiology Congress in 2016.

Author list and affiliations:

Tom Kai Ming Wang[1,2], Sophie Harmos[1], Greg Gamble[2], Tharumenthiran Ramanathan[1], Peter Ruygrok[1,2]

1. Green Lane Cardiovascular Service, Auckland City Hospital, Auckland, New Zealand

2. Department of Medicine, University of Auckland, Auckland, New Zealand

Contribution of candidate: study planning and design, data collection, analysis and interpretation, writing manuscript, revision in response to co-authors and editors, conference presentations, journal submission

6.1 Abstract

Background: Risk stratification for mitral valve repair or replacement (MVR) is important in the decision-making for treating several mitral valve disease but is rarely studied. We compared the prognostic utility of EuroSCORE, EuroSCORE II and Society of Thoracic Surgeon's (STS) Score for MVR.

Methods: The three scores were retrospectively calculated for consecutive patients undergoing isolated MVR at Auckland City Hospital during 2005-2012 and their discrimination and calibration for mortality and morbidities assessed.

Results: There were 408 patients (mitral valve repair 48.1% and replacement 51.9%) followed-up for 6.0+/-2.6 years. The operative mortality was 2.5%. Mean EuroSCORE, EuroSCORE II and STS Score were 7.6%, 3.4% and 3.5%. C-statistics were 0.844, 0.817 and 0.850 for operative mortality. Hosmer-Lemeshow test P-values were 0.076, 0.541 and 0.306, and Brier scores 0.0246, 0.0035 and 0.0075 respectively for operative mortality. The numerically highest c-statistic for predicting complications include EuroSCORE for return to the operating room (c=0.673), EuroSCORE II for stroke (c=0.669) and mediastinitis (c=0.801), and STS for renal failure (c=0.828), ventilation>24 hours (c=0.789) and composite morbidity (c=0.732). The individual STS complication models for MVR had a numerically higher c-statistic only for stroke (c=0.737).

Conclusions: All scores discriminated mortality and most morbidities after MVR, although EuroSCORE over-estimated operative mortality. The STS Score was the best overall predictor of mortality and morbidity in the MVR cohort.

6.2 Introduction

Although mitral valve surgery (MVR) is the mainstay treatment for severe symptomatic mitral valve disease, techniques such as percutaneous mitral valve interventions are evolving and becoming an alternative in higher risk patients(49, 78, 80, 104, 105). Accurate risk modeling therefore has an increasingly important role in the stratification and selection of the optimal management of patients with mitral valve disease. The European System for Cardiac Operative Risk Evaluation (EuroSCORE) and Society of Thoracic Surgeons' (STS) Score are the most widely used surgical risk scores to assess operative mortality(27, 32). The latter has specific models for both mitral valve repair and replacement, as well as predicting complications. The EuroSCORE II was recently published to better evaluate contemporary outcomes(30). We compared the performance of EuroSCORE, EuroSCORE II and STS Score at predicting mortality and morbidities after MVR.

6.3 Methods

All patients undergoing isolated MVR including repair and replacement, but without concurrent coronary or other valve surgery, at Auckland City Hospital during 2005-2012 were included. We retrospectively calculated the logistic EuroSCORE(27), EuroSCORE II(30) and the STS Score(32) using available data for all patients. The EuroSCORE and EuroSCORE II are general cardiac surgery models, while the STS Score has separate models for mitral valve repair and replacement which we used, as well as separate models for mortality and post-operative complications.

EuroSCORE II risk model parameter definitions were used for baseline characteristics(30), including for critical pre-operative state, extracardiac arteriopathy, left ventricular ejection fraction, pulmonary hypertension and renal function8. The New York Heart Association (NYHA) and Canadian Cardiovascular Society (CCS) classifications were used for dyspnea and angina grading respectively. Hypertension included any blood pressure measurement over 140/90 mmHg or being on medications for lowering blood pressure. Stroke included any history of a neurological deficit persisting >24 hours as assessed by a neurologist. Details on valve repair or replacement, valve prosthesis used, and times for operation, cardiopulmonary bypass and cross-clamp were recorded as operative characteristics.

Operative mortality was defined as deaths in-hospital or within 30 days of surgery. Post-operative complications including stroke, renal failure, ventilation>24 hours, mediastinitis, return to the operating room, prolonged hospital stay>14 days and their composite were defined as per the STS database(32). We focused our analyses on these early and/or in-hospital outcomes only.

Mean+/-standard deviation and percentages (frequency) were used to present continuous and categorical variables respectively. Area under the receiver operative characteristics curves (c-statistics) with 95% confidence intervals (95%CI) were used to assess the discriminative ability of risk scores. Observed/expected ratios, Hosmer-Lemshow goodness-of-fit test and Brier Scores were used to assess calibration. Univariate analyses including logistic regression were not performed. We assessed the performance of mortality risk scores at predicting operative mortality, and both mortality and STS complication-specific scores at predicting post-operative complications. Statistical analyses were performed using SPSS (Version 17.0, SPSS Inc., Chicago, IL, USA) and Prism (Version 5, GraphPad Software, San Diego, CA, USA). All tests were two-tailed and P<0.05 considered statistically significant. Ethical approval was obtained from the Auckland District Health Board Research Office before the commencement of the study, and individual consent was waived.

6.4 Results

MVR was performed in 407 consecutive patients at Auckland City Hospital during 2005-2012. Baseline characteristics are presented in table 6.1. Mean age was 56.6+/-16.2 years and 44.0% (179) patients were female. Mean scores were 7.6+/-8.3% for EuroSCORE, 3.4+/-4.8% for EuroSCORE II and 3.5+/-7.9% for STS Score.

Table 6-1 Baseline characteristics

Demographics	
Age (years)	56.6+/-16.2
Female	44.0% (179)
Ethnicity	
New Zealand European	57.4% (234)
Maori or Pacific	33.1% (135)
Other	9.3% (38)
Body mass index (kg/m^2)	28.2+/-6.5
Body surface area (m^2)	1.92+/-0.31
Presentation	
New York Heart Association class	
1	25.6% (104)
2	31.9% (130)
3	33.4% (136)
4	9.1% (37)
Unstable angina class 4	0.0% (0)
Syncope	1.5% (6)
Critical pre-operative state	8.6% (35)
Inpatient urgent operation	51.4% (202)
Past medical history	
Previous cardiac surgery	19.7% (80)
Valve surgery	18.7% (76)
Coronary artery bypass grafting	2.2% (9)

Congestive heart failure	27.8% (113)
Myocardial infarction	6.6% (27)
Recent myocardial infarction in 90 days	2.5% (10)
Percutaneous coronary intervention	1.7% (7)
Atrial fibrillation	47.7% (194)
Diabetes	8.7% (35)
Hypertension	30.2% (123)
Hypercholesterolaemia	32.5% (132)
Current smoker	13.6% (55)
Smoking history	50.6% (206)
Active infective endocarditis	12.0% (49)
Cerebrovascular accident	9.3% (38)
Extracardiac arteriopathy	2.5% (10)
Chronic pulmonary disease	19.7% (80)
Dialysis	1.0% (4)
Investigations	
Ejection Fraction	
Normal ($\geq$50%)	85.7% (349)
Mild/moderate impairment (30-49%)	13.5% (55)
Severe Impairment (<30%)	0.7% (3)
Mitral regurgitation	90.7% (369)
Mitral stenosis	15.7% (64)
Aortic regurgitation	1.7% (7)
Aortic stenosis	1.5% (6)
Tricuspid regurgitation	13.0% (53)
Pulmonary hypertension	
Moderate (31-55mmHg)	
Severe (>55mmHg)	
Left main artery >50% stenosis	0.2% (1)
Main coronary vessels >50% stenosis	

1	6.4% (26)
2	1.7% (7)
3	2.7% (11)
Creatinine clearance (mL/min)	75+/-25
Risk score	
EuroSCORE	7.6+/-8.3%
EuroSCORE II	3.4+/-4.8%
STS Score	3.5+/-7.9%

All figures are mean+/-standard deviation or percentage (frequency)

Table 6.2 lists the operative characteristics and post-operative outcomes. Valve replacement and repair were performed in 51.8% (211) and 48.2% (196) respectively, and amongst valve replacements the majority 79.1% (167) of patients received mechanical valves. Operative mortality occurred in 2.5% (10) patients, and composite morbidity in 18.9% (77) with stroke in 1.7% (7), renal failure in 2.9% (12) and ventilation>24 hours in 13.3% (54).

Table 6-2 Operative variables and post-operative outcomes

Operative variables	
Valve replacement	51.8% (211)
Mechanical valve	79.1% (167/211)
Valve repair	48.2% (196)
Operation time (minutes)	224+/-75
Cardiopulmonary bypass time (minutes)	129+/-50
Cross-clamp time (minutes)	96+/-37
In-hospital outcomes	
Composite morbidity	18.9% (77)
Stroke	1.7% (7)
Renal Failure	2.9% (12)
Ventilation>24 hours	13.3% (54)
Deep sternal wound infection	1.0% (4)
Return to theatre	8.1% (33)
Prolonged hospital stay >14 days	13.8% (56)

Operative mortality	2.5% (10)	

All figures are mean+/-standard deviation or percentage (frequency)

Table 6.3 shows the discrimination of mortality by scores. All scores discriminated operative mortality (c=0.817-0.850), however EuroSCORE over-estimated operative mortality with O/E ratio 0.32, the highest Brier Score of 0.025 and near significant Hosmer-Lemeshow P value of 0.076.

Table 6-3 Discrimination and calibration analyses for risk scores and mortality

	EuroSCORE	EuroSCORE II	STS
Mean+/-standard deviation	7.6+/-8.3%	3.4+/-4.8%	3.5+/-7.9%
Operative mortality AUC	0.844 (0.745-0.943)	0.817 (0.713-0.920)	0.850 (0.751-0.949)
O/E ratio	0.32	0.72	0.69
Hosmer-Lemeshow P-value	0.076	0.541	0.306
Brier Score	0.0246	0.0035	0.0075

AUC= c-statistic (95% confidence interval)

Discrimination of post-operative complications is shown in table 6.4. All scores were able to detect composite morbidity, renal failure, ventilation>24 hours, return to the operating room and prolonged hospital stay >14 days. The STS Score performed best for composite morbidity c=0.732, renal failure c=0.828 and ventilation>24 hours c=0.789, STS morbidity-specific score was only superior to the others at predicting stroke c=0.737.

Table 6-4 Discrimination analyses for risk scores and post-operative complications

Outcome	EuroSCORE	EuroSCORE II	STS	STS Complications
Composite morbidity	0.716 (0.652-0.781)	0.719 (0.652-0.781)	0.732 (0.669-0.795)	0.732 (0.667-0.797)
Stroke	0.597 (0.380-0.814)	0.669 (0.471-0.867)	0.665 (0.446-0.883)	0.737 (0.569-0.905)
Renal failure	0.745 (0.575-0.914)	0.790 (0.662-0.918)	0.828 (0.711-0.945)	0.813 (0.714-0.912)
Ventilation>24 hours	0.752 (0.682-0.821)	0.769 (0.703-0.836)	0.789 (0.726-0.853)	0.784 (0.716-0.852)
Mediastinitis	0.734 (0.561-0.854)	0.801 (0.637-0.965)	0.721 (0.539-0.904)	0.447 (0.159-0.735)
Return to theatre	0.673 (0.577-0.768)	0.668 (0.569-0.767)	0.668 (0.574-0.762)	0.643 (0.544-0.742)
Length of stay>14 days	0.697 (0.626-0.768)	0.719 (0.651-0.787)	0.696 (0.621-0.771)	0.716 (0.644-0.789)

C-statistic (95% confidence interval), STS Complications=specific morbidity scores of STS at predicting its complications (the other columns are mortality scores)

6.5 Discussion

In our cohort of MVR, the EuroSCORE, EuroSCORE II and STS Score all displayed good discrimination of operative mortality (c=0.81-0.85), however the EuroSCORE over-estimated mortality by three-fold. All three scores had moderate discrimination of post-operative complications except for stroke. Our surgical outcomes are comparable to other contemporary cohorts in the STS registry(32).

Contrary to coronary artery bypass grafting and isolated aortic valve replacement, very few studies have specifically assessed risk scores for MVR. One study of MVR only reported c-statistic for operative mortality of 0.67 and 0.74 for EuroSCORE II and STS Score(106). Two other studies of all cardiac operation reported c-statistics for their isolated MVR subgroup: 0.88 for EuroSCORE and 0.87 for EuroSCORE II in one(67), and 0.89 for EuroSCORE II in the other study(68). We also found high discriminative performance for all the scores with c-statistics over 0.80. These scores might be further improved by incorporating new parameters such as frailty, nutrition and anemia.

Compared to studies of other types of cardiac surgery from our centre, discrimination of risk scores was best in MVR. C-statistics were lowest 0.64-0.68 for isolated coronary bypass grafting, higher 0.71-0.75 for isolated aortic valve replacement and 0.59-0.70 for the combined procedure for the same three risk scores(1, 2, 4). This is somewhat surprising for EuroSCORE and EuroSCORE II which were derived from populations predominantly made up of those undergoing coronary surgery. Similar findings were identified in other studies where c-statistics for MVR were the highest amongst various types of cardiac surgeries(67, 68). This should encourage greater utility of risk scores in the management of mitral valve disease requiring intervention. Our patient cohort has a higher prevalence of rheumatic heart disease as the underlying mitral valve pathology, which can explain some of the differences observed in performance of these scores.

In terms of calibration, we found the original EuroSCORE to over-estimate operative mortality by about 3 times, compared to the two more contemporary scores similar to other studies(1, 2, 4, 66, 67). The only other study reporting calibration of risk scores for MVR found that both EuroSCORE II and STS Score also significantly over-estimated operative mortality by 2-3 times, however in this cohort operative mortality was relatively low at only 1.0% (9). This emphasizes the need for risk scores to be updated regularly in order to better fit the ever improving surgical outcomes.

Previous studies have found all these scores to predict composite morbidity(1, 3, 66), stroke(3, 66), renal failure(1, 3), prolonged ventilation(3, 4), mediastinitis(66) and prolonged hospital stay(1, 3, 4), in coronary artery bypass grafting and/or aortic valve replacement, although they were designed to only predict operative mortality. This is the first study specifically looking at risk scores predicting

complications in MVR, we found moderate discrimination for all risk scores except stroke. Stroke was the only outcome where using the specific STS complication model for stroke provided incremental benefit for predicting this outcome. This differed from the findings of our AVR study where the STS complication model performed better than other scores for all complications while having good calibration(3).

The Mitraclip edge to edge repair is the main percutaneous mitral intervention available currently(104). Only one study has assessed the performance of risk scores at predicting operative mortality in the Mitraclip to date, with c-statistics of 0.67, 0.80 and 0.62 for EuroSCORE, EuroSCORE II and STS Scores respectively(107). These scores are, expectedly, lower than predicting MVR outcomes(27, 30, 32). Development of risk models specific to percutaneous mitral valve intervention is warranted to optimize risk stratification to guide clinical practice.

This study has some limitations. It is a single-centre observational study which may limit the generalizability of our findings to other centres. Risk scores were calculated retrospectively from clinical records that may contain minor biases. The number of participants and adverse events limited the statistical power for analyses, in particular for direct comparison of c-statistics. There was an insufficient number of events to compare the calibration of scores by risk quartiles. Furthermore, small differences in c-statistics may not be clinically meaningful, and large differences are less likely to be observed when it is high, as in our study. Mortality was the only outcome collected at follow-up, and although this is the main endpoint of interest, other important long-term endpoints include, stroke, redo operations, symptoms and quality of life, and cause of death also could not be obtained.

6.6 Conclusion

In conclusion, all three scores discriminated operative mortality, however the EuroSCORE grossly over-estimated operative mortality. All scores moderately discriminated post-operative complications except for stroke. These scores appear to be superior for MVR than other types of cardiac surgery. The STS Score was the best overall predictor of morbidity and mortality in this MVR cohort.

7 Comparison of contemporary risk scores for predicting outcomes after surgery for active infective endocarditis

Beyond individual types of cardiac surgery, there are some important clinical scenarios for cardiac interventions that warrant accurate risk modelling to guide selection and management. Surgery for infective endocarditis is considered in patients with resultant heart failure, uncontrolled infection, haemodynamic instability and embolism prevention(7). About half of endocarditis patients have surgical indications however they are generally high risk and frequently undertreated. This chapter was the first study in the literature to compare the outcomes predictions of cardiac surgery and endocarditis-specific risk models after infective endocarditis surgery.

This manuscript was published in 2013 in Heart and Vessels volume 30 pages 227-234. As of September 2018, it had 18 citations on Google scholar, including being referenced by European Society of Cardiology guidelines of infective endocarditis 2015(7), American Association fort Thoracic Surgery guidelines for surgical treatment of infective endocarditis 2016(8) and one meta-analysis(42). It was a poster presentation at the Cardiac Society of Australia and New Zealand Annual Scientific Meeting and oral presentation at the European Society of Cardiology Congress in 2013.

Author list and affiliations:

Tom Kai Ming Wang[1,2], Timothy Oh[1], Jamie Voss[1], Nicholas Kang[1], Peter Ruygrok[1,2]

1. Green Lane Cardiovascular Service, Auckland City Hospital, Auckland, New Zealand

2. Department of Medicine, University of Auckland, Auckland, New Zealand

Contribution of candidate: study planning and design, data collection, analysis and interpretation, writing manuscript, revision in response to co-authors and editors, conference presentations, journal submission

7.1 Abstract

Background: Decision-making regarding surgery for acute bacterial endocarditis is complex given its heterogeneity and often fatal course. Few studies have investigated the utility of operative risk scores in this setting. Endocarditis-specific scores have recently since been developed. We assessed the prognostic utility of contemporary risk scores for mortality and morbidity after endocarditis surgery.

Methods: Additive and logistic EuroSCORE I, EuroSCORE II, additive Society of Thoracic Surgeon's (STS) Endocarditis Score and additive De Feo-Cotrufo Score were retrospectively calculated for patients undergoing surgery for endocarditis during 2005-2011. Pre-specified primary outcomes were operative mortality, composite morbidity and mortality during follow-up.

Results: A total of 146 patients were included with an operative mortality of 6.8% followed for 4.1+/-2.4 years. Mean scores were additive EuroSCORE I: 8.0+/-2.5, logistic EuroSCORE I: 13.2+/-10.1%, EuroSCORE II: 9.1%+/-9.4%, STS Score: 32.2+/-13.5 and De Feo-Cotrufo Score: 14.6+/-9.2. Corresponding areas under curve (AUC) for operative mortality 0.653, 0.645, 0.656, 0.699 and 0.744; for composite morbidity were 0.623, 0.625, 0.720, 0.714 and 0.774; and long-term mortality 0.588, 0.579, 0.686, 0.735 and 0.751. The best tool for post-operative stroke was EuroSCORE II: AUC 0.837; and ventilation>24 hours and return to theatre were De Feo-Cotrufo Score: AUC 0.821 and 0.712. Pre-operative inotrope or intra-aortic balloon pump treatment, previous coronary bypass grafting and dialysis were independent predictors of operative and long-term mortality.

Conclusion: Risk models developed specifically from endocarditis surgeries and incorporating endocarditis variables have improved prognostic ability of outcomes, and can play an important role in the decision-making towards surgery for endocarditis.

7.2 Introduction

Infective endocarditis remains a heterogeneous disease with high mortality, despite advances in diagnostic and treatment over the last few decades(108). During the active phase of endocarditis when patients are on intravenous antibiotics, surgery is recommended for treatment of resultant heart failure or haemodynamic instability, uncontrolled infection or prevention of systemic embolism(109, 110). Surgery however comes with significant risks, so the decision to operate is often complex. Prognostic scoring tools, if accurate, can be of help to clinicians and researchers.

Few studies have investigated the utility of additive(26) and logistic(27) EuroSCORE I for patients undergoing surgery for endocarditis, and only for detecting operative mortality(46, 111, 112).

EuroSCORE II(30) had since been developed and validated for cardiac surgery, predominantly coronary artery bypass grafting and valve surgery(66, 69). More recently, Gaca et al.(45) developed a risk score specific to endocarditis surgery using 13,617 patients from the Society of Thoracic Surgeon's (STS) database, given that the original STS score(31) cannot be used in endocarditis patients having surgery. De Feo et al.(46) also developed a risk score in their single-centred piloted study of 440 native-valve endocarditis patients undergoing surgery. The external validities of these novel endocarditis-specific scores have not been fully assessed. We aimed to assess the prognostic utility of the EuroSCOREs I and II, STS Endocarditis Score and De Feo-Cotrufo Score for mortality and morbidity after surgery for active endocarditis.

7.3   Methods

Patient selection and data collection

Consecutive patients undergoing cardiac surgery for active endocarditis during 2005-2011 at Auckland City Hospital were identified from the adult cardiothoracic surgical unit database. All surgeries were undertaken after discussion at a multi-disciplinary cardiac conference, taking into account international guidelines, important indications of heart failure, severe sepsis, haemodynamic instability and embolic prevention, patient's co-morbidities and surgical risks. Endocarditis was defined as active if patients were on intravenous antibiotic therapy for endocarditis at the time of surgery with confirmatory intra-operative findings of endocarditis. Relevant clinical characteristics, operative variables and post-operative outcomes were retrospectively collected from computerised hospital records. Additive(26) and logistic(27) EuroSCORE I, EuroSCORE II(30), additive STS Endocarditis Score(45) and additive De Feo-Cotrufo Score(46) were calculated for all patients, blinded to outcomes.

Definitions of presentation with congestive heart failure, unstable angina, urgency of surgery, history of hypertension, cerebrovascular accident, peripheral vascular disease and chronic respiratory disease are identical to corresponding parameters in the STS score(31). Inotrope or intra-aortic balloon pump treatment refers to cardiac support therapies that were initiated pre-operatively in the same admission. Valve regurgitation or stenosis need to be graded moderate or severe to be counted.

The primary outcome of the study was operative mortality, defined as in-hospital death or death within 30 days of operation. Secondary outcomes include mortality during follow-up and composite morbidity, consisting of the five post-operative complications of permanent stroke, renal failure, prolonged ventilation over 24 hours, deep sternal wound infection and return to theatre for any reason as defined by the STS score(31). Mortality data were checked against New Zealand's national registry up till 31 December 2012.

79

Statistical analyses

Continuous and categorical variables are presented as mean (standard deviation) and percentages (frequency) respectively. Mann-Whitney U test and Fisher's exact test were used for univariate analyses. Discriminative powers for post-operative outcomes for all 5 risk scores were assessed using the area under the receiver-operative characteristics curve (AUC). Logistic regression and Cox proportional hazards regression were used to identify predictors of pre-specified end-points, calculating odds ratios (OR) or hazards ratios (HR) and their 95% confidence intervals (95%CI). Only pre-operative variables with p<0.10 in univariate analyses, excluding risk scores, were incorporated in these multivariate models. Statistical analyses were performed using SPSS (Version 17.0, SPSS Inc., Chicago, IL, USA) and Prism (Version 5, GraphPad Software, San Diego, CA, USA). P-values less than 0.05 were deemed statistically significant and all statistical tests were two-tailed. Ethical approval was attained from our institution's research office.

7.4 Results

Patient characteristics

A total of 146 patients had surgery for active endocarditis during the 7-year study period and table 7.1 presents the baseline characteristics including mean risk scores. Mean age was 48.8 (16.0) years and 70.5% (103/146) were male. Mean additive EuroSCORE I was 8.0 (2.5), logistic EuroSCORE I was 13.2 (10.1)%, EuroSCORE II was 9.1 (9.4)%, additive STS Endocarditis Score was 32.2 (13.5) and additive De Feo-Cotrufo Score was 14.6 (9.2).

Table 7-1 Baseline characteristics

Characteristics	All	Death	Alive	P-value
Number	146	10	136	
Demographics				
Age (years)	48.8 (16.0)	55.9 (15.0)	48.2 (16.0)	0.125
Male	70.5% (103)	80.0% (8)	69.9% (95)	0.724
Body mass index (kg/m^2)	27.8 (6.4)	29.2 (5.0)	27.7 (6.4)	0.244
Presentation				
Congestive heart failure	47.3% (69)	60.0% (6)	46.3% (63)	0.517
Unstable angina	1.4% (2)	0.0% (0)	2.9% (4)	1.000
Inotrope or intra-aortic balloon pump treatment	23.3% (34)	60.0% (6)	20.6% (28)	0.011

New complete heart block	6.2% (9)	0.0% (0)	6.6% (9)	1.000
New embolic event	40.4% (59)	50.0% (5)	39.7% (54)	0.526
New cerebral event	27.4% (40)	20.0% (2)	27.9% (38)	0.728
Structures involved				
Aortic valve	64.4% (94)	60.0% (6)	64.7% (88)	0.744
Mitral valve	42.5% (62)	40.0% (4)	42.6% (58)	1.000
Tricuspid valve	7.5% (11)	0.0% (0)	7.5% (11)	1.000
Pulmonary valve	0.7% (1)	0.0% (0)	0.7% (1)	1.000
Device	3.4% (5)	10.0% (1)	2.9% (4)	0.302
More than one of above	14.4% (21)	0.0% (0)	14.4% (21)	0.358
Prosthetic valve	33.6% (49)	30.0% (3)	33.8% (46)	1.000
Intracardiac abscess	27.4% (40)	40.0% (4)	26.5% (36)	0.462
Blood culture				
Staphylococcus aureus	29.5% (43)	50.0% (5)	27.9% (38)	0.160
Streptococcus species	31.5% (46)	20.0% (2)	32.4% (44)	0.506
Enterococcus faecalis	8.9% (13)	10.0% (1)	8.8% (12)	1.000
Other	21.9% (32)	10.0% (1)	22.8% (31)	0.692
Negative	8.2% (12)	10.0% (1)	8.1% (11)	0.588
Operation status				0.171
Emergency	2.0% (3)	10.0% (1)	1.5% (2)	
Urgent	96.6% (141)	90.0% (9)	97.0% (132)	
Elective	1.4% (2)	0.0% (0)	1.5% (2)	
Past medical history				
Previous endocarditis	11.0% (16)	0.0% (0)	11.8% (16)	0.602
Rheumatic heart disease	7.5% (11)	10.0% (1)	7.4% (10)	0.555
Congenital heart disease	6.2% (9)	0.0% (0)	6.2% (9)	1.000
Cardiac valve operation	28.1% (41)	30.03% (3)	27.9% (38)	1.000
Coronary artery bypass grafting	2.7% (4)	20.0% (2)	1.5% (2)	0.024
Myocardial infarction	4.8% (7)	0.0% (0)	5.1% (7)	1.000
Congestive heart failure	28.8% (42)	20.0% (2)	29.4% (40)	0.724
Hypercholesterolaemia	26.0% (38)	30.0% (3)	25.7% (35)	0.720

Hypertension	28.1% (41)	30.0% (3)	27.9% (38)	1.000
Diabetes mellitus	11.6% (17)	30.0% (3)	10.3% (14)	*0.094*
Current smoker	18.5% (27)	40.0% (4)	16.9% (23)	*0.088*
Atrial fibrillation	21.2% (31)	30.0% (3)	20.6% (28)	0.443
Cerebrovascular accident	20.5% (30)	20.0% (2)	20.6% (28)	1.000
Peripheral vascular disease	5.5% (8)	10.0% (1)	5.1% (7)	0.441
Chronic respiratory disease	7.5% (11)	10.0% (1)	7.4% (10)	0.555
Dialysis	8.2% (12)	30.0% (3)	6.6% (9)	0.037
Investigations				
Ejection fraction				0.971
Normal ($\geq$60%)	81.5% (119)	80.0% (8)	81.6% (111)	
Mild (45-59%)	11.0% (16)	10.0% (1)	11.0% (15)	
Moderate (30-44%)	6.8% (10)	10.0% (1)	6.6% (9)	
Severe (<30%)	0.7% (1)	0.0% (0)	0.7% (1)	
Valve regurgitation	72.6% (106)	60.0% (6)	73.5% (100)	0.462
Valve stenosis	13.0% (19)	20.0% (2)	12.5% (17)	0.619
Pulmonary arterial systolic pressure (mmHg)				0.612
Normal (<31)	80.5% (95/118)	80.0% (8/10)	80.6% (87/108)	
Moderate (31-55)	13.6% (16/118)	20.0% (2/10)	13.0% (14/108)	
Severe (>55)	4.8% (7/118)	0.0% (0/10)	6.5% (7/108)	
Creatinine clearance (mL/min)	90 (47)	81 (51)	91 (47)	0.403
Risk scores				
EuroSCORE I additive	8.0 (2.5)	9.3 (2.6)	7.9 (2.5)	0.103
EuroSCORE I logistic (%)	13.2% (10.1%)	17.5% (12.5%)	12.9% (9.9%)	0.126
EuroSCORE II (%)	9.1% (9.4%)	14.1% (11.6%)	8.7% (9.2%)	0.100
Society of Thoracic Surgeon's Score	32.2 (13.5)	41.7 (15.3)	31.5 (13.2)	0.033
De Feo-Cotrufo Score	14.6 (9.2)	23.1 (10.4)	14.0 (8.8)	0.010

In-hospital outcomes

Table 7.2 shows the operative and post-operative outcomes. Operative mortality was 6.8% (10/146). Both logistic scores (EuroSCORE I and EuroSCORE II) significantly over-estimated operative mortality

(p<0.001 and p=0.004). Composite morbidity occurred in 33.6% (49/146), predominantly ventilation>24 hours in 28.8% (42/146) and return to theatre in 14.4% (21/146).

Table 7-2 Operative characteristics and post-operative outcomes

	All	Death	Alive	P-value
Number	146	10	136	
Operation				
Valve repair	29.5% (43)	10.0% (1)	30.9% (42)	0.282
Annuloplasty	13.7% (20)	10.0% (1)	14.0% (19)	1.000
Valve replacement	77.4% (113)	80.0% (8)	77.2% (105)	1.000
Mechanical	45.9% (67)	50.0% (5)	42.5% (62)	1.000
Biological	31.5% (46)	40.0% (4)	30.9% (42)	0.725
Coronary artery bypass grafting	8.9% (13)	20.0% (2)	8.1% (11)	0.219
Operation time (minutes)	261 (105)	339 (149)	255 (99)	*0.072*
Cardiopulmonary bypass time (minutes)	152 (75)	233 (134)	146 (66)	0.017
Cross-clamp time (minutes)	113 (61)	155 (102)	110 (56)	0.189
In-hospital outcomes				
Operative mortality	6.8% (10)			
Composite morbidity	33.6% (49)	90.0% (9)	29.4% (40)	<0.001
Permanent stroke (%)	4.1% (6)	30.0% (3)	2.2% (3)	0.004
Renal failure (%)	6.2% (9)	20.0% (2)	5.1% (7)	0.117
Ventilation>24 hours (%)	28.8% (42)	80.0% (8)	25.0% (34)	0.001
Deep sternal wound infection (%)	1.4% (2)	10.0% (1)	0.7% (1)	0.133
Return to theatre (%)	14.4% 921)	40.0% (4)	12.5% (17)	0.038
Operation to discharge time (days)	15.4 (10.2)	9.5 (7.8)	15.9 (10.2)	0.042

AUCs for each risk score for detecting mortality and morbidity after surgery are listed in Table 7.3. Only STS Endocarditis Score with AUC 0.699 (p=0.036) and De Feo-Cotrufo Score with AUC 0.744 (p=0.010) reached statistical significance for detecting operative mortality. The optimal cut-points for detecting operative mortality are STS Score of 36 (sensitivity 70.0%, specificity 66.9%) and De Feo-Cotrufo Score of 25 (sensitivity 60.0%, specificity 86.0%).

EuroSCORE II, STS Endocarditis Score and De Feo-Cotrufo Score were good discriminators of composite morbidity with AUC 0.720 (p<0.001), 0.714 (p<0.001) and 0.774: (p<0.001). The best discriminator of permanent stroke was EuroSCORE II with AUC 0.837 (p=0.005). De Feo-Cotrufo Score had the highest AUC for ventilation>24 hours of 0.821 (p<0.001) and return to theatre of 0.712 (p=0.002). None of the scores were statistically significant at detecting renal failure or deep sternal wound infection.

Table 7-3 Receiver-operative characteristics analysis (area under curve and 95% confidence intervals)

Outcomes	EuroSCORE I additive	EuroSCORE I logistic	EuroSCORE II	STS Score	De Feo-Cotrufo Score
Operative mortality	0.653 (0.487-0.819)	0.645 (0.487-0.803)	0.656 (0.466-0.846)	0.699 (0.534-0.865)	0.744 (0.590-0.899)
Mortality during follow-up	0.588 (0.439-0.737)	0.579 (0.433-0.725)	0.686 (0.558-0.814)	0.735 (0.616-0.855)	0.751 (0.649-0.852)
Composite morbidity	0.632 (0.537-0.727)	0.625 (0.530-0.720)	0.720 (0.632-0.808)	0.714 (0.630-0.799)	0.774 (0.692-0.855)
Permanent stroke	0.649 (0.452-0.846)	0.645 (0.455-0.835)	0.837 (0.742-0.931)	0.681 (0.517-0.845)	0.770 (0.605-0.936)
Renal failure	0.448 (0.306-0.590)	0.431 (0.288-0.573)	0.520 (0.381-0.659)	0.429 (0.275-0.583)	0.622 (0.499-0.744)
Ventilation>24 hours	0.680 (0.586-0.775)	0.663 (0.568-0.759)	0.769 (0.683-0.855)	0.758 (0.675-0.841)	0.821 (0.740-0.901)
Deep sternal wound infection	0.311 (0.006-0.616)	0.344 (0.134-0.553)	0.455 (0.000-0.965)	0.681 (0.000-1.000)	0.592 (0.204-0.980)
Return to theatre	0.630 (0.513-0.746)	0.618 (0.499-0.736)	0.613 (0.478-0.748)	0.683 (0.572-0.794)	0.712 (0.595-0.823)

Longitudinal outcomes

Mean follow-up was 4.1 (2.4) years and all patients had at least 1-year follow-up. One, three and five year survivals of the entire cohort were 92.5%, 91.4% and 89.0% respectively. The scores were statistically significant at detecting mortality during follow-up were EuroSCORE II with AUC 0.686 (p=0.013), STS Endocarditis Score with AUC 0.735 (p=0.002) and De Feo-Cotrufo Score with AUC 0.751 (p=0.001), as shown in table 3.

Multivariate analyses

Predictors of mortality and morbidity in multivariate analyses are indicated in table 7.4. Independent predictors of both operative mortality and mortality during follow-up were inotrope or intra-aortic balloon pump treatment, previous coronary artery bypass grafting and dialysis. Predictors of composite morbidity included inotrope or intra-aortic balloon pump treatment and coronary artery bypass grafting performed during operation.

Table 7-4 Multivariable analyses

Predictors	Ratios	95% confidence interval	P-value
Operative mortality	**Odds ratio**		
Inotrope or intra-aortic balloon pump treatment	8.17	1.54-43.3	0.014
Diabetes mellitus	4.38	0.726-26.5	0.097
Current smoker	4.26	0.832-21.8	0.082
Previous coronary artery bypass grafting	5.08	2.13-12.4	0.002
Dialysis	7.25	1.23-42.9	0.029
Mortality during follow-up	**Hazards ratio**		
Inotrope or intra-aortic balloon pump treatment	5.17	1.64-16.3	0.005
Diabetes mellitus	6.33	1.68-23.9	0.006
Current smoker	3.91	1.31-11.7	0.015
Previous coronary artery bypass grafting	9.24	2.02-42.9	0.002
Dialysis	10.0	1.60-62.4	0.014
Composite Morbidity	**Odds ratio**		
Inotrope or intra-aortic balloon pump treatment	7.04	2.68-18.5	<0.001
Intracardiac abscess	2.32	0.929-5.80	0.072
Diabetes mellitus	2.85	0.84-9.73	0.094

Coronary artery bypass grafting performed	4.68	1.04-21.0	0.044
Return to theatre	**Odds ratio**		
Inotrope or intra-aortic balloon pump treatment	3.30	1.11-9.82	0.032
Cerebral event from endocarditis	6.29	1.18-33.3	0.032
Intracardiac abscess	2.62	0.921-7.43	0.071
Peripheral vascular disease	6.53	1.04-40.9	0.045
Permanent stroke	**Odds ratio**		
History of cerebrovascular accident	8.10	0.849-77.3	0.069
Peripheral vascular disease	5.33	1.47-19.3	0.009
Ventilation>24 hours	**Odds ratio**		
Male	3.39	0.842-13.6	0.086
Inotrope or intra-aortic balloon pump treatment	7.98	2.80-22.7	<0.001
Negative blood culture	8.20	0.985-68.3	0.052
Hypercholesterolaemia	3.79	1.02-14.1	0.047
Peripheral vascular disease	7.92	1.04-60.4	0.046

7.5   Discussion

This is the first study to validate both the STS Endocarditis Score and De Feo-Cotrufo Score as prognostic of operative mortality after surgery for active endocarditis, and better predictors than the EuroSCOREs in our cohort. Our second finding was that all five risk scores discriminated post-operative morbidity, however EuroSCORE II, STS Score and De Feo-Cotrufo Score were better predictors than both EuroSCORE I, particularly for permanent stroke and ventilation >24 hours. We also identified several independent predictors of mortality and morbidity after endocarditis surgery. The operative mortality of 6.8% was comparable to previously reported rates of 2.7-28.8% reported in various studies(45, 46, 111-119), reflecting the heterogeneity of the disease.

Additive and logistic EuroSCORE I are the only risk scores whose performance have been previously assessed in endocarditis surgery in three studies(46, 111, 112). These found good discrimination of additive EuroSCORE I with AUC 0.83 and 0.75 and logistic EuroSCORE I with AUC 0.84, 0.74 and 0.84 for operative mortality. Our results showed that EuroSCORE I both over-estimated and failed to discriminate operative mortality for endocarditis surgery. One reason may be that EuroSCORE I, based on cardiac surgery undertaken in 1995, is out-dated in the contemporary context of ever-improving surgical and peri-operative care, as observed in other studies of cardiac surgery(28, 65). Both additive

and logistic EuroSCORE 1 however did discriminate post-operative morbidities, particularly ventilation>24 hours, similar to that reported for other cardiac surgeries(88).

For detecting operative mortality, EuroSCORE II was no better than EuroSCORE I. EuroSCORE II however did discriminate mortality during follow-up which EuroSCORE I did not, and was able to detect post-operative morbidity, also shown in one other study(89), particularly permanent stroke. What could then be limiting the utility of EuroSCOREs in our setting is probably because these were derived predominantly from coronary and valve surgeries rather than patients with endocarditis.

STS Endocarditis score was constructed specifically from endocarditis operations(45) and was able to detect operative mortality, mortality during follow-up and post-operative morbidities in our cohort. Its constituents are quite similar to the EuroSCORE(26, 30). The use of pre-operative inotropes or intra-aortic balloon pump is a similar variable to the critical pre-operative state parameter of other scores, and shown to be an important predictor of mortality and morbidity in our study, as was renal failure. The STS score also distinguishes previous coronary artery bypass grafting from previous valve surgery unlike other scores, with exponential effect if both are present(45). The former is an important predictor of mortality in our study suggesting that underlying ischaemic heart disease and potentially heart failure are important risk factors. Unlike the EuroSCORE II, STS score includes all diabetes as parameters, not just those on insulin, which we and other studies(115, 118) have found to be associated with higher mortality, and this could explain why the STS Score appeared to have the highest AUC for deep sternal wound infection.

The De Feo-Cotrufo Score was derived from a smaller single-centred pilot study of native-valve endocarditis surgery patients as a preliminary to multicentre development(46). It was good discriminator for adverse outcomes after endocarditis surgery in our cohort, not inferior to the STS score, suggesting that it may also be applicable to prosthetic valve endocarditis. Its unique feature is incorporating endocarditis variables such as lack of pre-operative attainment of blood culture negativity and perivalvular involvement as parameters, which were not collected in the EuroSCORE(26, 30) or STS databases(31). Although these were not predictors of mortality in our study, they were associated with three of five post-operative complications that nearly reached statistical significance in multivariate analysis. Furthermore, De Feo-Cotrufo Score is relatively simple with only 6 parameters, putting a lot of weight on critical pre-operative state, which is a strong predictor of adverse outcomes. Further developments underway with the De Feo-Cotrufo Score will likely improve on its existing strengths.

Apart from the parameters of existing risk scores and characteristics we identified, there are several other variables associated with adverse outcomes after surgery for endocarditis reported in the literature. Prosthetic valve endocarditis was associated with higher mortality after surgery than native valve

endocarditis in some studies(116). The De Feo-Cotrufo Score in particular was derived entirely from native valve endocarditis, so its application can be widened if prosthetic valve and device infections are added into the model. Staphylococcus aureus grown from blood culture also predicted mortality in several cohorts but not ours, and the De Feo-Cotrufo Score instead had negative blood culture as a parameter(115, 118). Another study found pre-operative neurological impairment due to endocarditis to be associated with mortality(120).

We can infer from our results several aspects on how best to improve on existing risk scores for endocarditis surgery in the future. Firstly, the mechanisms of adverse outcomes are more complex in endocarditis, involving sepsis, inflammation and higher risk of embolic phenomenon in addition to ischaemia and heart failure, so the model may be best constructed from cardiac operations for endocarditis only. Secondly, many of the existing parameters of cardiac surgery risk scores, particularly pre-operative inotrope, intra-aortic balloon pump and/or ventilation, previous cardiac operations, renal function and diabetes are important risk factors and should be retained. Thirdly, variables unique to endocarditis such as valvular type and complications, blood culture results and embolic phenomenon should be tested in the model. Finally, as per all prognostic models, constant revision and large populations are required to strengthen the calibration of the score to match the ever-evolving clinical practice and assist in treatment selection, identification of adverse prognostic factors and patient counselling(121).

Study limitations

This was a single-centred retrospective observational study. We could not obtain sufficient information to investigate the efficacy of the logistic models of the STS and De Feo-Cotrufo Scores. The moderate sample size meant we only had a limited number of post-operative adverse events. Follow-up was limited given that we studied a contemporary cohort. We focused on patients having surgery for active endocarditis so our results do not necessarily apply to patients having surgery for treated endocarditis or having medical treatment only. Of note, 51% of the STS and 83% of the De Feo-Cotrufo Scores derivation cohort had active endocarditis contributing to their higher AUCs for adverse outcomes.

7.6   Conclusion

STS endocarditis Score and De Feo-Cotrufo Score detected mortality after operations for active endocarditis. Both of these and EuroSCORE II were also good discriminators for post-operative morbidities particularly permanent stroke and ventilation>24 hours. To optimise discriminative efficacy for post-operative outcomes after endocarditis surgery, operative risk scores should be derived and

applied specifically to endocarditis surgeries, incorporate endocarditis variables as parameters and be constantly revised to fit contemporary outcome.

8   Risk scores and surgery for infective endocarditis: a meta-analysis

As an extension from the work in chapter 7 and given the scarce literature in the field, a meta-analysis was performed to pool the performance results of risk scores for endocarditis surgery across related publications. The main focus is on EuroSCORE and EuroSCORE II as other endocarditis-specific risk models have not been externally validated in other studies except ours in chapter 7.

This manuscript was published in 2016 in the International Journal of Cardiology volume 222 pages 1001-1002. In contrast to other publications in this thesis which were full original articles, this was published as a brief communication letter. Additional unpublished tables and figures from this work are presented in the appendix at the end of the chapter. As of September 2018, it had one citation on google scholar. It was a poster presentation at Cardiac Society of Australia and New Zealand Annual Scientific Meeting and moderated poster presentation at European Society of Cardiology Congress in 2016.

Author list and affiliations:

Tom Kai Ming Wang[1,2], Michael Tzu Min Wang[1,2], James Pemberton[1]

1. Green Lane Cardiovascular Service, Auckland City Hospital, Auckland, New Zealand

2. Department of Medicine, University of Auckland, Auckland, New Zealand

Contribution of candidate: study planning and design, data collection, analysis and interpretation, writing manuscript, revision in response to co-authors and editors, conference presentations, journal submission

8.1 Introduction

Approximately half of patients with infective endocarditis, a heterogeneous and often lethal disease, require surgery for the management of valvular regurgitation, congestive heart failure, severe infection and/or embolic prevention(7). Risk models such as the EuroSCORE and the revised EuroSCORE II have played important roles in stratification and decision-making for cardiac surgery for over a decade, but remain under-utilised in the setting of endocarditis(30). There remains limited literature and uncertainty whether these general cardiac surgery risk models accurately predict mortality after the frequently complex cardiac surgery for infective endocarditis. This meta-analysis aims to pool and compare the prognostic utility of surgical risk scores in endocarditis surgery.

8.2 Methods

MEDLINE, Embase, Cochrane and Web of Science databases from 1 January 1980 to 31 December 2015 were searched for relevant studies. The search terms used were "endocarditis" AND ("surgery" OR "operation") AND ("score" OR "model" OR "index"). Endnote X7.0.1 (Thomson Reuters 1988-2013) was used to manage articles obtained from the literature search subsequent evaluation. Two authors (TKMW and MTMW) independently assessed studies for inclusion, and discrepancies were resolved by consensus. Studies needed to be on adult human subjects, and were included if they reported the c-statistic and its 95% confidence interval (95%CI) of one or more risk score at predicting operative mortality (in-hospital or within 30 days) after surgery for infective endocarditis.

Data pertaining to sample size, observed operative mortality and mean risk scores were then extracted. Discrimination was measured by c-statistics with 95%CI, pooled for each risk score if it was reported in two or more external validation studies. The Peto's odds ratios were also calculated and pooled for these scores as a measure of calibration. Review Manager 5.3 (Cochrane Collaboration, London, UK) program was used in this meta-analysis.

8.3 Results

Amongst 480 articles obtained from the initial search, the full-texts of 20 relevant articles were screened and 8 studies with 1,743 total patients included for analyses(6, 44, 46, 47, 111, 112, 122, 123). The twelve full-text articles excluded were because of reporting endocarditis patients not exclusively having surgery in four, not reporting risk scores in five and when risk scores are reported, not reporting c-statistics in three. Characteristics of the included studies are listed in table 8.1. Operative mortality was 19.4% overall. Only the EuroSCORE and EuroSCORE II were reported in two or more external validation studies for pooled analyses to be conducted.

Table 8-1 Characteristics of included studies and discrimination c-statistics (95% confidence interval) pooled analyses

Author	Year	Country	N	Operative mortality	EuroSCORE (mean)	C-statistic	EuroSCORE II (mean)	C-statistic
De Feo(46)	2012	Italy	252	9.4%	Not reported	0.84 (0.77-0.91)	Not reported	Not reported
Madeira(122)	2015	Portugal	128	16.4%	24.4%	0.75 (0.66-0.85)	11.9%	0.83 (0.75-0.91)
Martinez(44)	2014	Spain	437	24.3%	Not reported	0.73 (0.70-0.77)	Not reported	Not reported
Mestres(111)	2007	Spain	191	28.8%	27.1%	0.84 (0.77-0.91)	Not reported	Not reported
Olmos(47)	2015	Spain	247	28.1%	Not reported	0.74 (0.69-0.79)	Not reported	Not reported
Patrat-Delon(123)	2015	France	149	21.5%	Not reported	Not reported	15.8%	0.78 (0.70-0.84)
Rasmussen(112)	2011	Denmark	193	10.9%	16.0%	0.74 (0.64-0.84)	Not reported	Not reported
Wang(6)	2015	New Zealand	146	6.8%	13.2%	0.65 (0.49-0.80)	9.1%	0.66 (0.47-0.85)
Pooled			**1743**	**19.4%**		**0.76 (0.72-0.81)**		**0.79 (0.72-0.85)**

Pooled c-statistics (95%CI) for operative mortality were 0.76 (0.72-0.81) for EuroSCORE in seven studies and 0.79 (0.72-0.85) for EuroSCORE II in three studies. In terms of calibration, figure 8.1 illustrates the Peto's odds ratios calculated for included studies for both scores: pooled Peto's odds ratios (95% confidence interval) were 0.76 (0.57-1.01) for EuroSCORE and 1.25 (0.84-1.86) for EuroSCORE II. Additional figures of pooled c-statistics and Labbe Plots are illustrated in figures 8-2 and 8-3 and summary of measures in table 8-2 in the supplementary appendix.

Figure 8-1 Pooled Peto's odds ratios for EuroSCORE (top) and EuroSCORE II (bottom) at predicting mortality after endocarditis surgery

Study or Subgroup	Observed Events	Total	Expected Events	Total	O–E	Variance	Weight	Peto Odds Ratio Exp[(O–E) / V], Fixed, 95% CI
Madeira 2015	21	128	31	128	-5	10.4	21.9%	0.62 [0.34, 1.14]
Mestres 2007	55	191	52	191	1.5	19.31	40.6%	1.08 [0.69, 1.69]
Rasmussen 2011	21	193	31	193	-5	11.28	23.7%	0.64 [0.36, 1.15]
Wang 2015	10	146	19	146	-4.5	6.55	13.8%	0.50 [0.23, 1.08]
Total (95% CI)		658		658			100.0%	0.76 [0.57, 1.01]
Total events	107		133					

Heterogeneity: Chi² = 4.27, df = 3 (P = 0.23); I² = 30%
Test for overall effect: Z = 1.89 (P = 0.06)

Study or Subgroup	Observed Events	Total	Expected Events	Total	O–E	Variance	Weight	Peto Odds Ratio Exp[(O–E) / V], Fixed, 95% CI
Madeira 2015	21	128	15	128	3	7.76	31.7%	1.47 [0.73, 2.97]
Patrat-Delon 2015	32	149	24	149	4	11.41	46.6%	1.42 [0.79, 2.54]
Wang 2015	10	146	13	146	-1.5	5.32	21.7%	0.75 [0.32, 1.76]
Total (95% CI)		423		423			100.0%	1.25 [0.84, 1.86]
Total events	63		52					

Heterogeneity: Chi² = 1.75, df = 2 (P = 0.42); I² = 0%
Test for overall effect: Z = 1.11 (P = 0.27)

8.4 Discussion

Our findings suggest that the EuroSCORE and EuroSCORE II have moderate discrimination with c-statistics of 0.76-0.79 for operative mortality after endocarditis surgery. Notably endocarditis is a parameter in these scores(30). Calibration on the other hand was suboptimal, particularly for the EuroSCORE which had a trend to over-estimating operative mortality (P=0.06). This is to be expected as this score was originally developed in 1999 as an additive model and 2003 as a logistic model, and similar findings have been seen in other cardiac surgery studies, therefore the use of EuroSCORE II, published in 2012, is more preferable for risk stratification.

Despite this, there is clearly room for improvement for both discrimination and calibration of risk scores in endocarditis surgery. One way to logistically address this is to develop endocarditis-specific risk models, and a few such as the De Feo-Cotrufo(46) and Pulsuse Scores(44) have been published. Only one study have externally validated such scores and found them to have better discrimination than the EuroSCOREs(111), and this finding was also mentioned in recent guidelines(7). An issue faced however is that only additive rather than logistic forms of endocarditis-specific risk models have been published, which means calibration cannot be assessed and hinders its clinical utility. Further development of

endocarditis-specific risk models, publication of logistic scores and larger external validation studies will therefore be important. Finally, regular revision of risk models to fit improving contemporary surgical outcomes is also required to maintain adequate calibration. Being aware of the limitation of risk models when applied to endocarditis patients is critical, and management should be individualised and risk assessment incorporate other factors not captured by these scores.

This study has some limitations. The meta-analysis is based on a small number of small to moderate-sized studies which restricts its power. The studies were all retrospective and observational, and in particular retrospective risk score calculation may have inherent biases especially if not all parameters were prospectively collected or available. All but one study(6) reported whether risk scores predicted other outcomes such as long-term mortality and post-operative complications, and also performance of endocarditis-specific risk scores, so these could not be further evaluated.

8.5 Conclusion

In summary, the general EuroSCORE and EuroSCORE II had moderate discrimination for operative mortality after endocarditis surgery, however the EuroSCORE had a trend to over-estimating operative mortality and therefore less preferable for use than the EuroSCORE II. Development of endocarditis-specific risk scores may further improve the prognostic accuracy in this setting, and larger external validation studies are required.

8.6 Supplementary appendix

Figure 8-2 Pooled c-statistics for EuroSCORE (top) and EuroSCORE II (bottom) at predicting mortality after endocarditis surgery
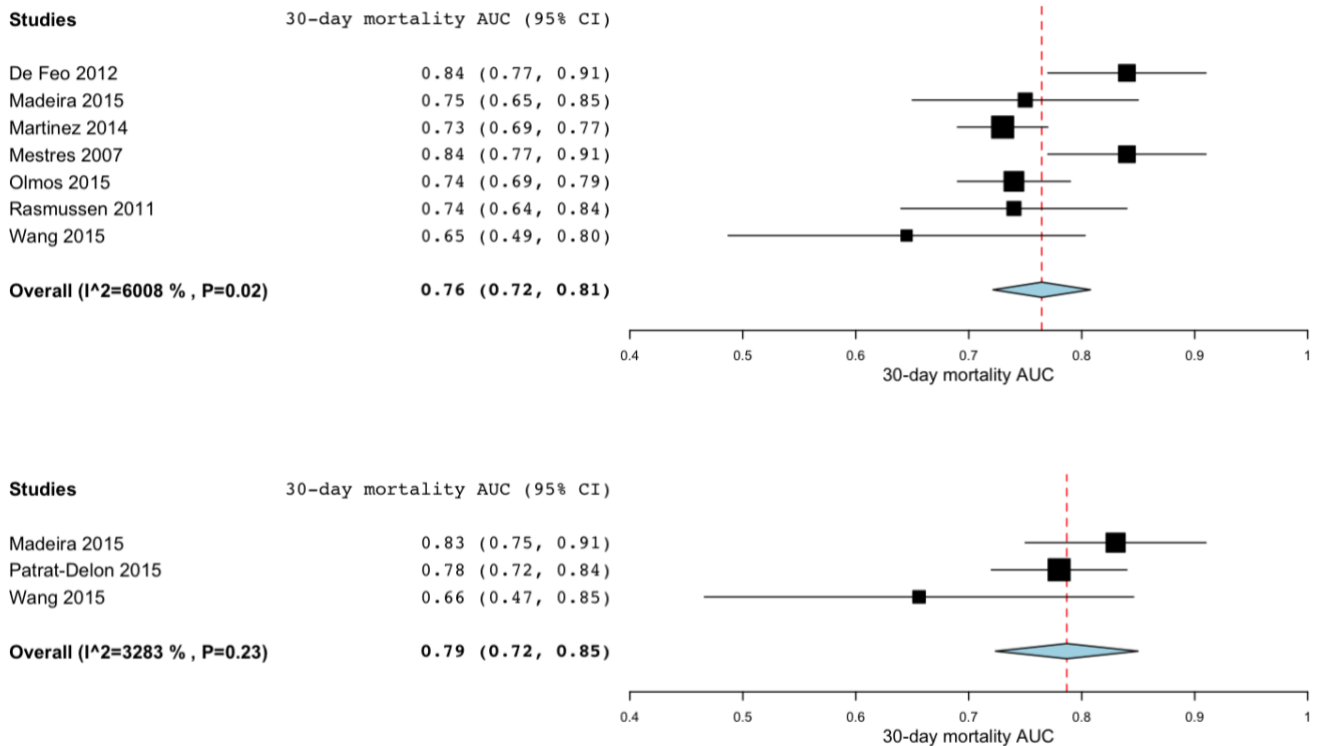
Table 8-2 Summary of risk scores performance measures

Measure	EuroSCORE	EuroSCORE II
Observed operative mortality	16.3%	14.9%
Estimated operative mortality (range)	20.2% (17.2%-23.5%)	12.3% (9.3%-15.8%)
Observed/expected ratio	0.80	1.21
Peto's Odds Ratios	0.76 (0.57-1.01)	1.25 (0.84-1.86)
C-statistics (30-day mortality)	0.76 (0.72-0.81)	0.79 (0.72-0.85)

Figure 8-3 Labbe Plots for a) EuroSCORE (top) and b) EuroSCORE II (bottom) at predicting operative mortality
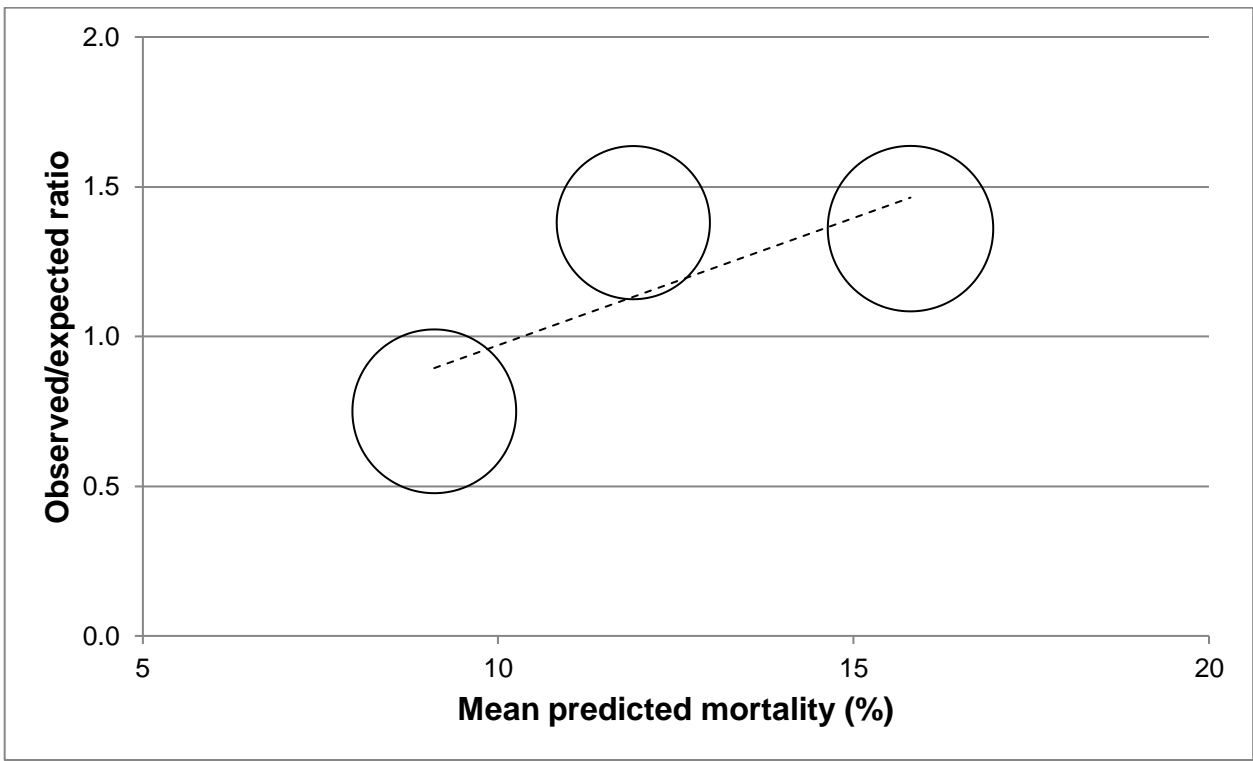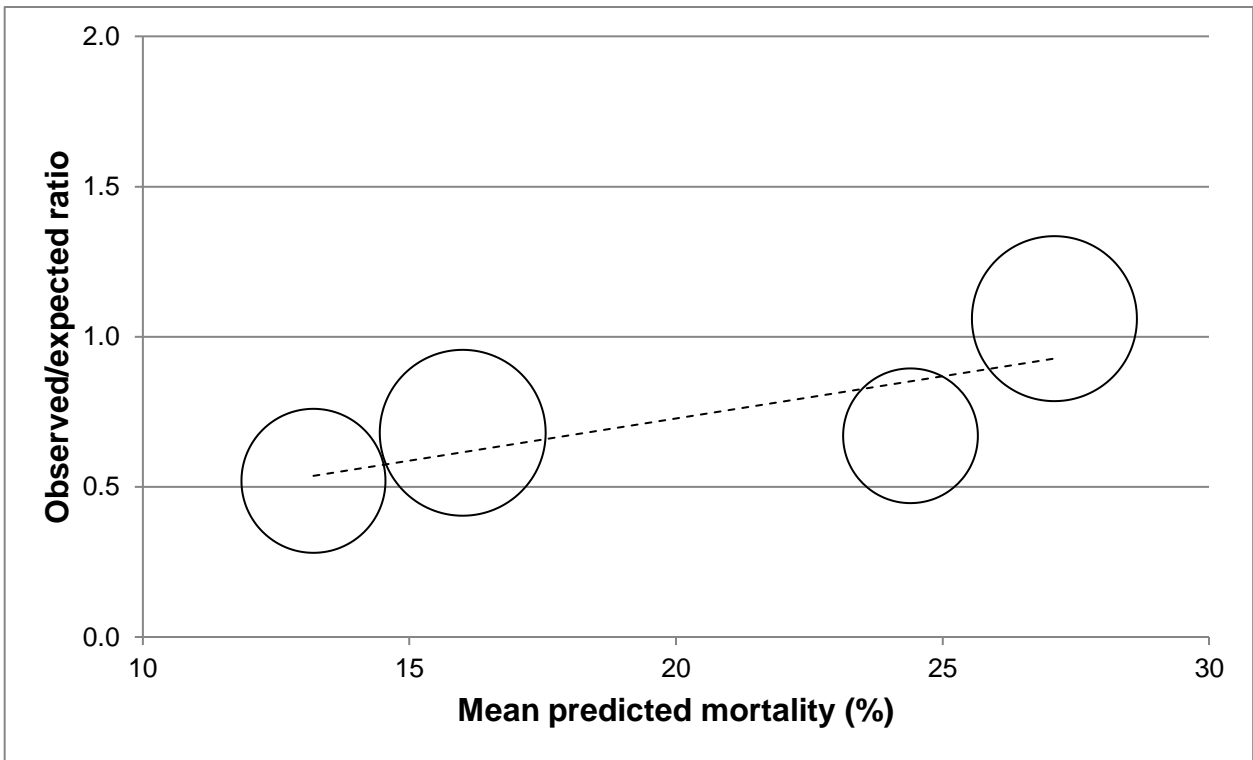
9   Performance of contemporary surgical risk scores for transcatheter aortic valve implantation: a meta-analysis.

Transcatheter aortic valve implantation (TAVI) is a percutaneous intervention to treat severe aortic valve disease(39). Its indications and roles have expanded significantly over the last decade as the alternative to aortic valve replacement to high and intermediate risk patients. In both clinical practice and research, surgical risk models are used for risk stratification to guide decision-making in these patients, but their accuracy here is unclear as they were derived from surgical not TAVI cohorts. This study aimed to pool the performance of surgical risk scores at predicting mortality after TAVI.

This manuscript was published in 2017 in the International Journal of Cardiology volume 236 pages 350-355. As of September 2018, it had four citations on Google scholar. It was a poster presentation at both the Cardiac Society of Australia and New Zealand Annual Scientific Meeting and the European Society of Cardiology Congress in 2016, receiving best poster award for aortic interventions in the latter conference.

Author list and affiliations:

Tom Kai Ming Wang[1,2], Michael Tzu Min Wang[1,2], Greg Gamble[2], Mark Webster[1], Peter Ruygrok[1,2]

1. Green Lane Cardiovascular Service, Auckland City Hospital, Auckland, New Zealand

2. Department of Medicine, University of Auckland, Auckland, New Zealand

Contribution of candidate: study planning and design, data collection, analysis and interpretation, writing manuscript, revision in response to co-authors and editors, conference presentations, journal submission

9.1 Abstract

Background: Transcatheter aortic valve implantation (TAVI) is considered for severe aortic valve disease at high and now intermediate risk for surgical aortic valve replacement. Risk stratification plays a critical role decision-making for intervention and modality. We compared the prognostic utility of surgical risk scores for TAVI in this meta-analysis.

Methods: MEDLINE, Embase, Cochrane and Web of Science databases from 1 January 1980 to 31 December 2015 were searched. Studies were systematically reviewed for inclusion, and data extracted for pooled analyses.

Results: Amongst 1688 articles searched, 47 full-text articles were screened and 24 studies (12,346 TAVI cases) included for analyses. Pooled c-statistics (95% confidence interval) for operative mortality were EuroSCORE 0.62 (0.57-0.67), EuroSCORE II 0.62 (0.59-0.66), STS Score 0.62 (0.59-0.65). Pooled observed/expected ratios (95%CI) were EuroSCORE 0.31 (0.25-0.38), EuroSCORE II 1.26 (1.06-1.51), STS 0.95 (0.72-1.27). C-statistics (95%CI) for 1-year mortality were EuroSCORE 0.62 (0.57-0.67), EuroSCORE II 0.66 (0.61-0.71) and STS Score 0.58 (0.53-0.64).

Conclusion: Surgical risk scores at most modestly discriminated operative and 1-year mortality. The EuroSCORE grossly over-estimated operative mortality while the EuroSCORE II and STS Scores fitted better to TAVI outcomes with their own limitations. There is a need for the development and validation of TAVI-specific risk models.

9.2 Introduction

Although surgical aortic valve replacement (AVR) is the gold standard for treating severe symptomatic aortic valve disease, transcatheter aortic valve implantation (TAVI) has recently established a role in the management of these patients(49, 78). TAVI was found to have comparable outcomes to surgery in high risk patients(51, 52) and more recently in intermediate risk patients(53, 124), and with several favourable characteristics use of this procedure will likely continue to grow in the near future.

Selection of the optimal treatment modality is complex, making accurate risk stratification critical. The most widely used surgical risk scores include The European System for Cardiac Operative Risk Evaluation (EuroSCORE)(26), its revised version EuroSCORE II(30), as well as the Society of Thoracic Surgeon's Score(32). They have been used for over a decade in a well-validated cohort of patients undergoing AVR, and more recently for those having TAVI in randomized trials(51-53), guidelines(49,

78) and clinical practice, however their accuracy in this context remains unclear. In this meta-analysis, we compared the performance of surgical risk scores at predicting mortality after TAVI.

9.3 Methods

Literature search

The PRIMSA guidelines were followed in the conduct of this meta-anaylsis. We searched MEDLINE, Embase, Cochrane and Web of Science databases for relevant studies and abstracts from 1 January 1980 to 31 December 2015, and also searched the reference lists of obtained articles. The search term used were "transcatheter aortic valve implantation" or "aortic valve replacement"; and "EuroSCORE", "STS", "Society of Thoracic Surgeon's Score", "risk", "score" or "model". Two authors (TKMW and MTMW) independently conducted the search and evaluated all studies for inclusion, and discrepancies were resolved by consensus. Endnote X7.0.1 (Thomson Reuters 1988-2013) was used to organize and evaluate the searched studies for inclusion.

Inclusion/exclusion criteria

All included studies must be original with adult (over 18 years of age) human subjects. Only studies reporting the results of c-statistics with 95% confidence interval (95%CI) of one or more logistic surgical risk scores: EuroSCORE, EuroSCORE II and STS Scores at predicting operative mortality after TAVI were included. Small studies with fewer than 100 patients, and reviews, were excluded. We also excluded the analyses for other scores like ACEF, Parsonnet and Ambler as they were each reported in less than 5 studies, and are much less used in clinical practice

Data extraction

All data extracted were carefully checked by two authors (TKMW and MTMW), for subsequent analyses. We recorded the study year, interval, location, number of subjects, age, sex, access site for each study when available. We also recorded the operative mortality rate (in-hospital and/or 30-day mortality), mean risk score and c-statistics with 95%CI for predicting operative mortality and 1-year mortality. Microsoft Excel was used for recording data

Statistical analysis

Use of random-effects model was pre-specified to pool data due to expected heterogeneity. Pooled analyses were conducted if the measured outcome of a specific risk model was reported in 2 or more studies. C-statistics and 95%CI were pooled as a measure of discrimination of scores, and pooled

observed/expected ratios for measuring calibration. L'Abbe plots were used to examine calibration at various mean predicted risks by study. Publication bias was assessed with Funnel Plots and Eggers and Beggs statistics. All tests were two tailed, with P<0.05 deemed statistically significant. Statistical analyses were conducted using Review Manager Version 5.3 (Cochrane Collaboration, Oxford, England).

9.4   Results

A total of 1,688 records were obtained from the searched, and after initial screening, 47 full-text articles were assisted. Subsequently 23 articles were excluded because of no c-statistic reported (16), no risk scores reported (2), no 95% confidence interval (1), TAVI-specific score (1), not TAVI (1), meta-analysis (1) and duplicate (1), leaving 24 studies included for analyses(55-58, 125-144). Characteristics of the 24 studies are listed in table 9.1. All but one study reported operative mortality rate, with 22 studies reporting 30-day mortality and 1 study reporting 30-day or in-hospital mortality as primary outcome.

Table 9-1 Characteristics of included studies

Author	Year	Study interval	N	Country	Centre	Age	Female	Transfemoral	Mortality 30 days	Mortality 1 year	ES	ESII	STS
Arai(125)	2015	Oct 2006-May 2013	703	France	1	83.3	49.1%	54.5%	96 (13.7%)	201 (28.6%)	21.4%	7.6%	7.6%
Balan(126)	2015	2011-2014	405	-	1	-	-	-	11 (2.7%)	-	-	-	-
Ben-Dor(127)	2011	Apr 2007-Jul 2010	111	United States	1	81.8	54.8%	-	13 (11.7%)	-	39.7%	-	11.5%
Beohar(128)	2014	-	2552	-	>1	85.6	47.6%	-	165 (6.5%)	-	26.6%	-	11.4%
Buchanan(129)	2012	Nov 2007-May 2011	417	Italy	1	79.5	-	-	20 (4.7%)	-	-	-	-
Capodanno(55)	2014	Dec 2010-Jun 2012	622	Italy	95	81.7	58.4%	-	37 (5.9%)	-	-	-	-
Conradi(130)	2013	-	300	Germany	1	-	-	-	32 (10.7%)	-	22.8%	7.3%	8.6%
D'Ascenzo(56)	2013	-	962	-	-	-	-	-	-	-	-	-	-
Debonnaire(58)	2015	-	511	-	-	82.0	62.0%	52%	29 (5.7%)	87 (17.0%)	18.3%	6.4%	16.6%
D'Onofrio(131)	2012	Jun 2007-Jan 2011	235	-	1	80.5	-	66.8%	10 (4.3%)	-	17.6%	5.4%	6.8%
Durand(132)	2013	May 2006-Oct 2011	250	France	1	83.0	54.0%	76.0%	19 (7.6%)	-	22.6%	7.7%	7.3%
Haensig(133)	2013	Feb 2006-May 2011	360	Germany	1	81.6	64.4%	0%	38 (10.6%)	-	30.0%	6.7%	11.7%
Iung(57)	2014	Jan 2010-Dec2011	1281	France/ Monaco	34	83.0	49.3%	73.2%	129 (10.1%)	-	-	-	-

Johansson(134)	2014	Jan 2008-Apr 2013	123	Sweden	1	-	-	69.1%	5 (4.1%)	-	25.0%	7.8%	7.3%
Piazza(135)	2010	Nov 2005-Jun 2009	168	Netherlands/ Switzerland	2	82.8	56.1%	-	19 (11.3%)	-	20.2%	-	6.7%
Sedaghat(136)	2013	2008-2012	206	Germany	3	80.5	47.1%	100%	14 (6.8%)	56 (27.2%)	29.2%	9.2%	9.5%
Silaschi(137)	2015	Jan 2008-Aug 2012	457	Germany	1	80.5	52.3%	-	44 (9.6%)	-	22.0%	7.0%	7.9%
Sinning(138)	2014	-	410	Germany	1	81.1	50.5%	-	29 (7.1%)	100 (24.4%)	26.8%	-	8.9%
Silva(139)	2015	Jan 2008-Jan 2013	418	Brazil	18	81.5	52.2%	96.2%	38 (9.1%)	-	20.2%	6.5%	14.7%
Sirotina(140)	2013	-	450	Germany	>1	-	-	63.8%	37 (8.2%)	-	21.0%	8.6%	7.5%
Stahli(141)	2013	-	350	Switzerland	1	82.2	51.1%	83%	32 (9.1%)	-	22.6%	8.0%	6.5%
Watanabe(142)	2014	Oct 2006-Nov 2011	453	France	1	83.1	49.7%	55.0%	57 (12.6%)	-	22.4%	8.1%	8.1%
Wendt(143)	2014	Jan 1999-Mar 2012	446	Germany	1	80.8	57.9%	65.2%	46 (10.3%)	-	20.7%	7.1%	7.6%
Zbronski(144)	2016	Mar 2010-Oct 2014	156	Poland	1	80.0	48.1%	77.6%	15 (9.6%)	-	30.6%	-	-

Pooled c-statistics (95%CI) for operative mortality were EuroSCORE 0.62 (0.57-0.67) in 21 studies, EuroSCORE II 0.62 (0.59-0.66) in 15 studies and STS Score 0.62 (0.59-0.65) in 21 studies. Figure 9.1 shows the pooled c-statistics analyses for a) EuroSCORE, b) EuroSCORE II and c) STS Score.

Figure 9-1 Pooled c-statistic (95%CI) for a) EuroSCORE, b) EuroSCORE II and c) Society of Thoracic Surgeon's Score at predicting 30-day operative mortality

Figure 9.2 presents the pooled Peto's odds ratio (95%CI) between and observed and predicted risk for the three main risk scores for operative mortality: EuroSCORE 0.31 (0.25-0.38), EuroSCORE II 1.26 (1.06-1.51) and STS 0.95 (0.72-1.27).

Figure 9-2 Pooled observed/expected ratio (95%CI) for a) EuroSCORE, b) EuroSCORE II and c) Society of Thoracic Surgeon's Score at predicting 30-day operative mortality

a) EuroSCORE

Study or Subgroup	Observed Events	Total	Expected Events	Total	Weight	Odds Ratio M-H, Random, 95% CI
Arai 2015	96	703	150	703	6.9%	0.58 [0.44, 0.77]
Ben-Dor 2011	13	111	44	111	4.2%	0.20 [0.10, 0.40]
Beohar 2014	165	2552	678	2552	7.4%	0.19 [0.16, 0.23]
Conradi 2013	32	300	68	300	5.7%	0.41 [0.26, 0.64]
D'Onofrio 2012	10	235	41	235	4.1%	0.21 [0.10, 0.43]
Delago 2015	29	511	94	511	5.8%	0.27 [0.17, 0.41]
Durand 2013	19	250	57	250	5.0%	0.28 [0.16, 0.48]
Haensig 2013	38	360	108	360	6.1%	0.28 [0.18, 0.41]
Johansson 2014	5	123	31	123	2.8%	0.13 [0.05, 0.34]
Piazza 2013	19	168	34	168	4.7%	0.50 [0.27, 0.92]
Sedaghat 2013	14	206	60	206	4.6%	0.18 [0.10, 0.33]
Silaschi 2015	44	457	101	457	6.2%	0.38 [0.26, 0.55]
Sinning 2014	29	410	110	410	5.8%	0.21 [0.13, 0.32]
Sinnot Silva 2015	38	418	84	418	6.0%	0.40 [0.26, 0.60]
Sirotina 2013	37	450	95	450	6.1%	0.33 [0.22, 0.50]
Stahli 2013	32	350	79	350	5.8%	0.35 [0.22, 0.54]
Watanabe 2014	57	453	101	453	6.4%	0.50 [0.35, 0.72]
Wendt 2014	46	446	108	446	6.3%	0.36 [0.25, 0.52]
Total (95% CI)		8503		8503	100.0%	0.31 [0.25, 0.38]
Total events	723		2043			

Heterogeneity: Tau² = 0.14; Chi² = 74.19, df = 17 (P < 0.00001); I² = 77%
Test for overall effect: Z = 11.12 (P < 0.00001)

b) EuroSCORE II

Study or Subgroup	Observed Events	Total	Expected Events	Total	Weight	Odds Ratio M-H, Random, 95% CI
Arai 2015	96	703	53	703	11.7%	1.94 [1.36, 2.76]
Conradi 2013	32	300	22	300	6.8%	1.51 [0.85, 2.66]
D'Onofrio 2012	10	235	13	235	3.7%	0.76 [0.33, 1.77]
Delago 2015	29	511	33	511	7.7%	0.87 [0.52, 1.46]
Durand 2013	19	250	19	250	5.4%	1.00 [0.52, 1.94]
Haensig 2013	38	360	24	360	7.4%	1.65 [0.97, 2.82]
Johansson 2014	5	123	10	123	2.3%	0.48 [0.16, 1.44]
Sedaghat 2013	14	206	19	206	4.8%	0.72 [0.35, 1.47]
Silaschi 2015	44	457	32	457	8.5%	1.41 [0.88, 2.28]
Sinnot Silva 2015	38	418	27	418	7.7%	1.45 [0.87, 2.42]
Sirotina 2013	37	450	39	450	8.6%	0.94 [0.59, 1.51]
Stahli 2013	32	350	28	350	7.4%	1.16 [0.68, 1.97]
Watanabe 2014	57	453	37	453	9.4%	1.62 [1.05, 2.50]
Wendt 2014	46	446	32	446	8.6%	1.49 [0.93, 2.38]
Total (95% CI)		5262		5262	100.0%	1.26 [1.06, 1.51]
Total events	497		388			

Heterogeneity: Tau² = 0.04; Chi² = 19.75, df = 13 (P = 0.10); I² = 34%
Test for overall effect: Z = 2.56 (P = 0.01)

c) STS Score

Study or Subgroup	Observed Events	Total	Expected Events	Total	Weight	Odds Ratio M-H, Random, 95% CI
Arai 2015	96	703	53	703	6.9%	1.94 [1.36, 2.76]
Ben-Dor 2011	13	111	13	111	4.7%	1.00 [0.44, 2.27]
Beohar 2014	165	2552	292	2552	7.4%	0.54 [0.44, 0.65]
Conradi 2013	32	300	26	300	6.0%	1.26 [0.73, 2.17]
D'Onofrio 2012	10	235	16	235	4.8%	0.61 [0.27, 1.37]
Delago 2015	29	511	85	511	6.5%	0.30 [0.19, 0.47]
Durand 2013	19	250	18	250	5.4%	1.06 [0.54, 2.07]
Haensig 2013	38	360	42	360	6.4%	0.89 [0.56, 1.42]
Johansson 2014	5	123	9	123	3.5%	0.54 [0.17, 1.65]
Piazza 2013	19	168	11	168	4.9%	1.82 [0.84, 3.95]
Sedaghat 2013	14	206	20	206	5.2%	0.68 [0.33, 1.38]
Silaschi 2015	44	457	36	457	6.4%	1.25 [0.79, 1.98]
Sinnot Silva 2015	38	418	61	418	6.6%	0.59 [0.38, 0.90]
Sirotina 2013	37	450	34	450	6.3%	1.10 [0.67, 1.78]
Stahli 2013	32	350	23	350	6.0%	1.43 [0.82, 2.48]
Watanabe 2014	57	453	37	453	6.5%	1.62 [1.05, 2.50]
Wendt 2014	46	446	34	446	6.4%	1.39 [0.88, 2.22]
Total (95% CI)		8093		8093	100.0%	0.95 [0.72, 1.27]
Total events	694		810			

Heterogeneity: Tau² = 0.28; Chi² = 93.57, df = 16 (P < 0.00001); I² = 83%
Test for overall effect: Z = 0.32 (P = 0.75)

Labbe plots of risk scores at predicting operative mortality are shown in figure 9.3. The EuroSCORE consistently over-estimated operative mortality, with a trend towards greater over-estimation when the mean score was higher. The EuroSCORE II marginally over-estimated operative mortality across all studies. The STS Score mildly under-estimated operative mortality at low mean scores and slightly over-estimated operative mortality at high mean scores.

Figure 9-3 Labbe Plots for a) EuroSCORE, b) EuroSCORE II and c) Society of Thoracic Surgeon's Score at predicting operative mortality

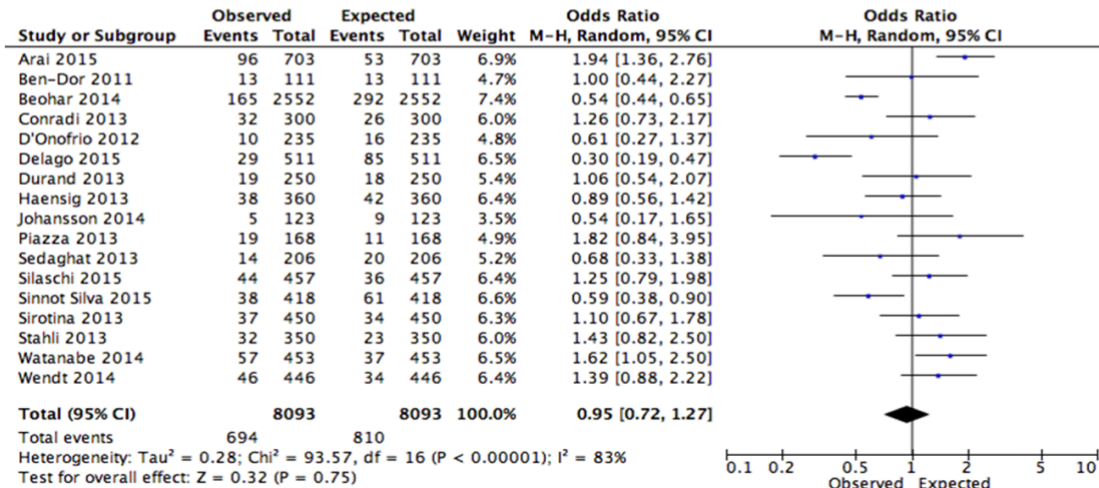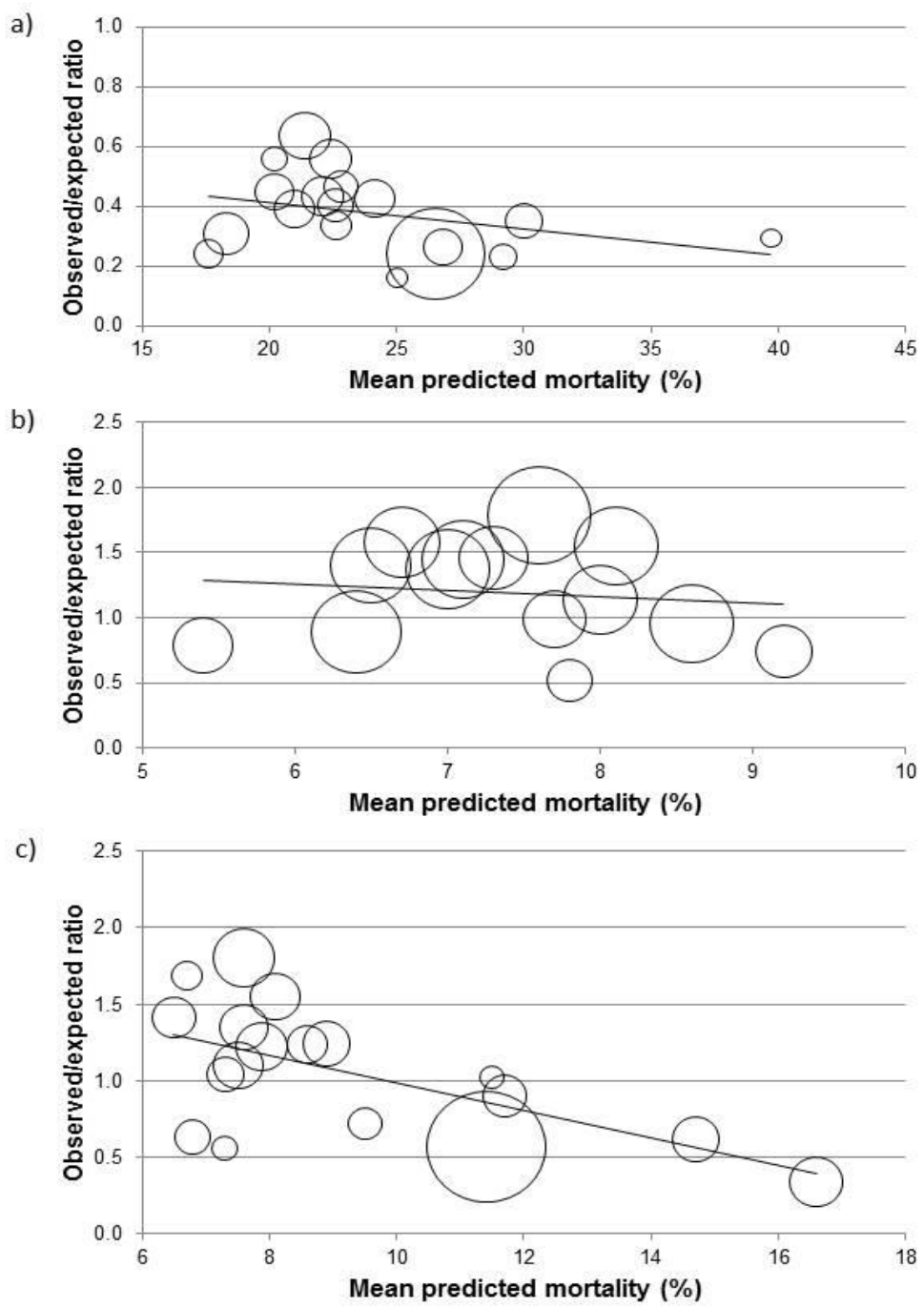A minority of studies reported c-statistics (95%CI) of risk scores for 1-year mortality. Pooled results were for EuroSCORE 0.62 (0.57-0.67) in 5 studies, EuroSCORE II 0.66 (0.61-0.71) in 4 studies and STS Score 0.58 (0.53-0.64) in 5 studies. None of the studies assessed risk scores at predicting long-term mortality beyond 1-year or other post-operative complications, except for Arai et al reporting c-statistics (95%CI) for predicting acute kidney injury for EuroSCORE 0.54 (0.47-0.60), EurosCORE II 0.57 (0.51-0.63) and STS Score0.54 (0.48-0.61)(125).

In the assessment for publication bias, Eggers and Beggs' statistics for all three scores were P>0.05. The Funnel plots for EuroSCORE and STS score are symmetrical, while that for EuroSCORE II suggests the possibility of publication bias for this score.

9.5 Discussion

This meta-analysis has produced several important findings. Firstly, the three main surgical risk scores all discriminated operative mortality after TAVI but did so only weakly, all with pooled c-statistics of 0.62. Secondly, the EuroSCORE considerably over-estimated mortality, the EuroSCORE II slightly under-estimated this, and the STS Score had the best calibration overall although its accuracy varied with the mean predicted risk score. There is therefore limitations in applying conventional surgical risk scores in TAVI.

Surgical risk scores overall appear to have better discrimination of operative mortality in cardiac surgery than TAVI. A meta-analysis of 22 cardiac surgery articles found EuroSCORE II to have a pooled c-statistic of 0.792(38), and another meta-analysis of 6 surgical aortic valve replacement studies reported 0.73-0.75 for EuroSCORE II and STS Score, respectively(145). These are higher than the 0.61-0.62 of 24 TAVI studies in our meta-analysis. This is to be expected, as these risk scores were derived from cardiac surgery cohorts, and in the case of the STS Score specific for different cardiac operations, rather than developed for TAVI(27, 30, 32). This suggests that not all parameters and their risk weightings are common to both procedures, and that there are other important factors not part of existing models.

On the other hand, similar findings are seen for calibration of surgical risk scores in TAVI compared with cardiac surgery. The original EuroSCORE developed over a decade ago significantly over-estimates operative mortality in patients undergoing cardiac surgery(28). We have found this also to be the case in those undergoing TAVI, despite similar discrimination to EuroSCORE II and STS Score, and therefore should no longer be used. In terms of EuroSCORE II and STS Scores, they slightly over-estimated mortality in one meta-analysis (Peto's odds ratio 0.74-0.86)(145), which may again reflect ongoing improvement in surgical outcomes even over the last 5 years. We found EuroSCORE II to have mild inaccuracies, and STS Score to have the best overall calibration so would be the preferred surgical score

in TAVI. Clinical trials have shown comparable outcomes of TAVI and AVR at least in high risk patients which may partially explain why surgical risk scores perform similarly in these settings, however the results also suggest the importance of regularly revising risk models to better fit contemporary outcomes(51-53).

A logical way to address this is developing risk models based on TAVI cohorts. The recently published OBSERVANT(55), SURTAVI(146), survival post TAVI (STT)(56), TAVI2-SCORe(58) and CoreValve study(59) risk models are examples of this. These studies generally found better performance of TAVI-specific scores, but other external validation studies mostly found these scores to not be superior to surgical risk models(144, 146). Clearly more studies of the development and validation of TAVI-specific scores are required. It is also important to publish logistic risk models which estimate mortality directly as a percentage, rather than additive risk models which tally up risk factors into a score, to enable evaluation of both discrimination and calibration and have superior clinical application. There should also be consideration of incorporating other prognostic parameters such as frailty which may further improve accuracy although it can add complexity as well. Furthermore, prediction and risk model development of long-term mortality may be just as important as operative mortality.

Given our findings, there needs to be caution and awareness of the limitations in applying these scores to TAVI as suggested by international guidelines(49, 78) and used in clinical trials of TAVI(51-53); the individualized approach full "Heart Team" assessment in patients considered for TAVI is critical(49, 78). The potential mismatch between surgical risk scores and heart team assessment is highlighted by the US CoreValve trial, which planned to enrol high-risk patients, with an estimated risk of death within 30 days of 15% or more, guided by STS Score(52). Approximately ¾ of those enrolled had an STS Score of 4-10%, indicating that the heart team considered other factors in assessing increased surgical risk.

This meta-analysis has several limitations. Included studies were retrospective and observational in nature, and in particular retrospective calculation of risk scores may introduce bias and not all parameters may have been available for this calculation. Secondly, not all studies reported the performance of all three main risk scores of interest which can affect interpretation when comparing the pooled results. Thirdly, the continuing improvements in outcomes with increased operator experience and improved valve technology are not accounted for by risk models. Furthermore, other sources of heterogeneity of studies such as different prosthesis and access site could affect the validity of the analyses despite the best attempts to adjust for these in the analyses with random effects model. We did not have patient-level data for those studies, so analyses such as calibration slope could not be performed. Also, publication bias particularly for EuroSCORE II may have had some influence on the pooled analyses. Finally, most individual studies were underpowered, which necessitated the need for a meta-analysis to be performed,

and more and larger studies required to validate both surgical and TAVI-specific risk models in the setting of TAVI.

9.6 Conclusion

In conclusion, the surgical EuroSCORE, EuroSCORE II and STS Score were able to modestly discriminate operative and 1-year mortality after TAVI. The EuroSCORE significantly over-estimated operative mortality so should not be used, while the EuroSCORE II and STS Scores performed better but still had some limitation. There is an imminent need to develop and validate TAVI-specific risk models for predicting mortality and morbidities with the hope of improved performance.

10 Diagnosis of myocardial infarction after coronary artery bypass grafting with high-sensitivity troponin T and new ECG or echocardiogram changes and relationship with mortality

Troponins are the preferred biomarkers for diagnosing myocardial infarction, but its use in the assessment and prognostication after cardiac surgery, including in diagnosing type 5 peri-operative myocardial infarction as per the universal definition, has not been well established and validated(72). This chapter evaluated whether rises in contemporary high-sensitivity cardiac troponins, with or without ECG and/or echocardiogram changes, were associated with mortality and morbidity after coronary artery bypass grafting.

This manuscript was published in 2013 in the European Heart Journal - Acute Cardiovascular Care volume 2 pages 323-333. As of September 2018, it had 23 citations on Google scholar, and by the Fourth Universal Definition of Myocardial Infarction guidelines published in 2018(11). It was an young investigator award finalist oral presentation at the Cardiac Society of Australia and New Zealand Annual Scientific Meeting 2013, poster at the European Society of Cardiology Congress in 2013, and also won the Royal Australasian College of Surgeons Young Investigator Prize in the cardiothoracic surgery section 2013.

Author list and affiliations:

Tom Kai Ming Wang[1,2], Ralph Stewart[1,2], Tharumenithiran Ramanathan[1], Nicholas Kang[1], Greg Gamble[2], Harvey White[,2]

1. Green Lane Cardiovascular Service, Auckland City Hospital, Auckland, New Zealand

2. Department of Medicine, University of Auckland, Auckland, New Zealand

Contribution of candidate: study planning and design, data collection, analysis and interpretation, writing manuscript, revision in response to co-authors and editors, conference presentations, journal submission

10.1 Abstract

Aims: Criteria for diagnosing myocardial infarction (MI) after coronary artery bypass grafting (CABG) are controversial. Uncertainties remain around the optimal threshold for biomarker elevation and the need for associated criteria. There are no studies of high-sensitivity troponin (hs-TnT) after CABG. We assessed whether using hs-TnT to define MI after CABG was associated with 30-day and medium-term mortality and evaluated the utility of adding to the troponin criteria new Q-waves or imaging evidence of new wall motion abnormality as suggested in the Universal Definition of MI.

Methods: Isolated CABG was performed in 818 patients from July 2010-June 2012 and hs-TnT was measured 12-24 hours after CABG. Patients with rising baseline or missing troponins (n=258) were excluded. Thresholds of 140ng/L (10 times 99th percentile upper reference limit (URL)) and 500ng/L (10 times coefficient of variation of 10% for 4th generation troponin T applied to hs-TnT) were pre-specified.

Results: Mean follow-up was 1.8 ± 0.6 years. On multivariate analyses isolated hs-TnT rise >140ng/L (n=360) or >500ng/L (n=162) were not associated with mortality. Additional ECG and/or echocardiographic criteria plus hs-TnT >140ng/L was associated with 30-day mortality hazards ratio (HR) 4.92 (95% CI 1.34-18.1), p=0.017 and medium-term mortality HR 3.44 (95% CI 1.13-10.5), p=0.030, whereas ECG and/or echocardiographic abnormalities with hs-TnT >500ng/L wasn't, (p=0.281 and p=0.123).

Conclusion: A definition for MI following CABG using hs-TnT with a cutpoint of 10 times the 99th percentile URL and ECG and/or echocardiographic criteria predicts 30-day and medium-term mortality. These findings validate the Universal Definition of type 5 MI.

10.2 Introduction

Cardiac troponins have been recommended by the Universal Definition of Myocardial Infarction (MI) as the preferred biomarkers for diagnosing MI after coronary artery bypass grafting (CABG)(72), due to superior sensitivity and specificity compared to traditional biomarkers like creatine kinase MB(147, 148). Several studies have found that isolated troponin rise after CABG, measured with contemporary assays, independently predicts adverse outcome(148-152).

High-sensitivity troponin assays have recently been developed(153), being more sensitive than contemporary assays at detecting lower troponin levels(72, 154). The properties and utility of hsTn may be different after CABG, and there are no data about hsTn use for the diagnosis of MI. The 2012 Third

Universal Definition defined MI (type 5) after CABG as requiring two criteria: 1. cardiac biomarkers (with troponins preferred) rise >10 times the 99% URL from a normal pre-operative level; and 2. a) new pathological Q-waves or new LBBB, and/or b) imaging or angiographic evidence of new occlusion of native vessels or grafts, new regional wall motion abnormality, or loss of viable myocardium(72).

We therefore assessed the ability of high-sensitivity troponin T (hs-TnT) to diagnose MI following CABG using several pre-specified criteria including the Universal Definition, and assessed its associations with mortality and morbidity.

10.3 Methods

Patient selection and data collection

Ethical approval of this study was obtained from our ethics review committee.

Patients undergoing isolated CABG without other concomitant cardiac surgery from the commencement of hs-TnT use in July 2010 to June 2012 were identified retrospectively from the cardiothoracic surgical unit database. Logistic EuroSCORE II which predicts operative risk was calculated(30). Buckberg cold blood cardioplegia was used for on-pump CABG.

Electrocardiograms (ECGs) were performed multiple times until discharge. Transthoracic echocardiograms were performed as indicated clinically. New Q-waves or left bundle branch block (LBBB) on the ECG or new regional wall motion abnormality on echocardiography were independently interpreted by two authors blinded to outcomes (TKMW and HDW), in accordance with the criteria of the Universal Definition(72).

Mortality data were checked against New Zealand's national registry up until 31 March 2013. Thirty-day and medium-term mortality were pre-specified as the primary outcomes. A composite of post-operative complications (renal failure, stroke, prolonged ventilation>24 hours, deep sternal wound infection and return to theatre) was a secondary outcome as determined according to the Society of Thoracic Surgeon's (SCS) definitions(32).

Troponin assays and classification

The hs-TnT assay (Roche Elecsys) which is guideline compliant(153) was introduced in July 2010. The 99th percentile upper reference limit (URL) of this assay is 14ng/L. The cutpoint commonly used clinically for the 4th generation troponin T with a 10% coefficient of variation (CV) (0.03ng/mL) when applied to the hs-TnT assay is 50ng/L(154). Patients routinely had hs-TnT measured 12 to 24 hours after

surgery. When there were more than one measurement during this time, the highest hs-TnT level was used.

Pre-operative troponin assays used at our hospital were the same hs-TnT. Two referral hospitals used the ABBOTT ARCHITECT troponin I assay with URL of 28ng/L or Siemens Dimension RxL troponin I assay with URL of 70ng/L(153). Baseline troponin levels measured >2 weeks before CABG were disregarded.

Patients were divided into three pre-specified categories based on pre-operative troponin levels: 1) `stable baseline troponins' included patients with pre-operative levels below the 99th percentile URL for the assay used, or with no troponin measurement and undergoing elective CABG; 2) `elevated stable baseline troponins' included patients with elevated pre-operative troponins above the 99th percentile URL, with at least two measurements <48 hours of each other, and the final pre-operative level either lower or less than 20% higher than the earlier measurement.; 3) excluded patients with elevated and rising, or no baseline troponin levels .

Two cutpoints for hs-TnT rise were pre-specified, the first was 10 times the 99th percentile URL(72). of the hs-TnT assay i.e. of 140ng/L and the second 10 times the CV of 10% for the previous cutpoint of the 4th generation troponin assay of 50ng/L(154) i.e. of 500ng/L.

For patients with `stable baseline troponins', a significant hs-TnT rise was defined as the post-operative hs-TnT being above the cutpoint. In patients with `elevated stable baseline troponins', a significant hs-TnT rise required the post-operative hs-TnT to be above the cutpoint, and also a 20% rise from the pre-operative level.

The prognostic significance of five pre-specified potential criteria for the diagnosis of MI were determined (1) hs-TnT rise >140ng/L alone, (2) hs-TnT rise >500ng/L alone, (3) new signs of MI on ECG and/or echocardiogram alone, (4) hs-TnT>140ng/L plus ECG and/or echocardiographic criteria, (5) hs-TnT rise >500ng/L plus ECG and/or echocardiographic criteria.

Statistical analyses

Continuous and categorical variables are presented as mean (standard deviation) or percentages respectively. Student t-test and Fisher's exact test for two groups or analysis of variance (ANOVA) and Chi-square for three or more groups were used for univariate analyses. Kaplan-Meier curves and log-rank (Mantel-Cox) test were performed for longitudinal survival analysis.

Receiver-operative characteristics (ROC) analysis was also used to assess how well combinations of post-operative hs-TnT levels and ECG and/or echocardiographic criteria correlated with mortality and morbidity. The area under ROC curve (AUC) and corresponding 95% confidence interval (95%CI) and p-values were calculated using the Wald Test for pairwise comparison with chance.

Baseline and operative variables with $p<0.10$ in univariate analyses were entered into multivariate analyses. Logistic regression was used to calculate odds ratios (OR) or Cox proportional hazards regression for hazards ratio (HR) and 95% CI. Independent predictors of five potential MI criteria as well as 30-day mortality, medium-term mortality and composite morbidity were determined. Each potential MI criteria was individually added to the models of post-operative outcomes to see whether they predicted these outcomes.

All tests were two tailed and p-values less than 0.05 were deemed statistically significant. SAS (Version 9.1, SAS Institute, Cary, NC, USA) and Prism (Version 5, GraphPad Software, San Diego, CA, USA) were used for analyses.

10.4 Results

Study Population

Figure 10.1 describes the study population. A total of 818 patients underwent isolated CABG during the 2-year study period. There were 258 patients that were excluded, 6 without post-operative hs-TnT measurements and 252 elevated and rising pre-operative troponins or lack of troponin measurements. There were 560 patients included, with 337 having 'stable baseline troponins' and 223 having `elevated stable baseline troponins'. Hs-TnT was measured 18.4 ± 3.5 hours after beginning surgery. Mean follow-up was 1.8 ± 0.6 years with all patients having at least 9 months follow-up. New Q-waves or new LBBB, were documented in 75 patients. Echocardiograms were performed in 51 patients, 11 of which showed new regional wall motion abnormalities. Only one post-operative coronary angiogram was performed, showing patent grafts and unchanged native vessels. No patient had magnetic resonance imaging.

Figure 10-1 Study population
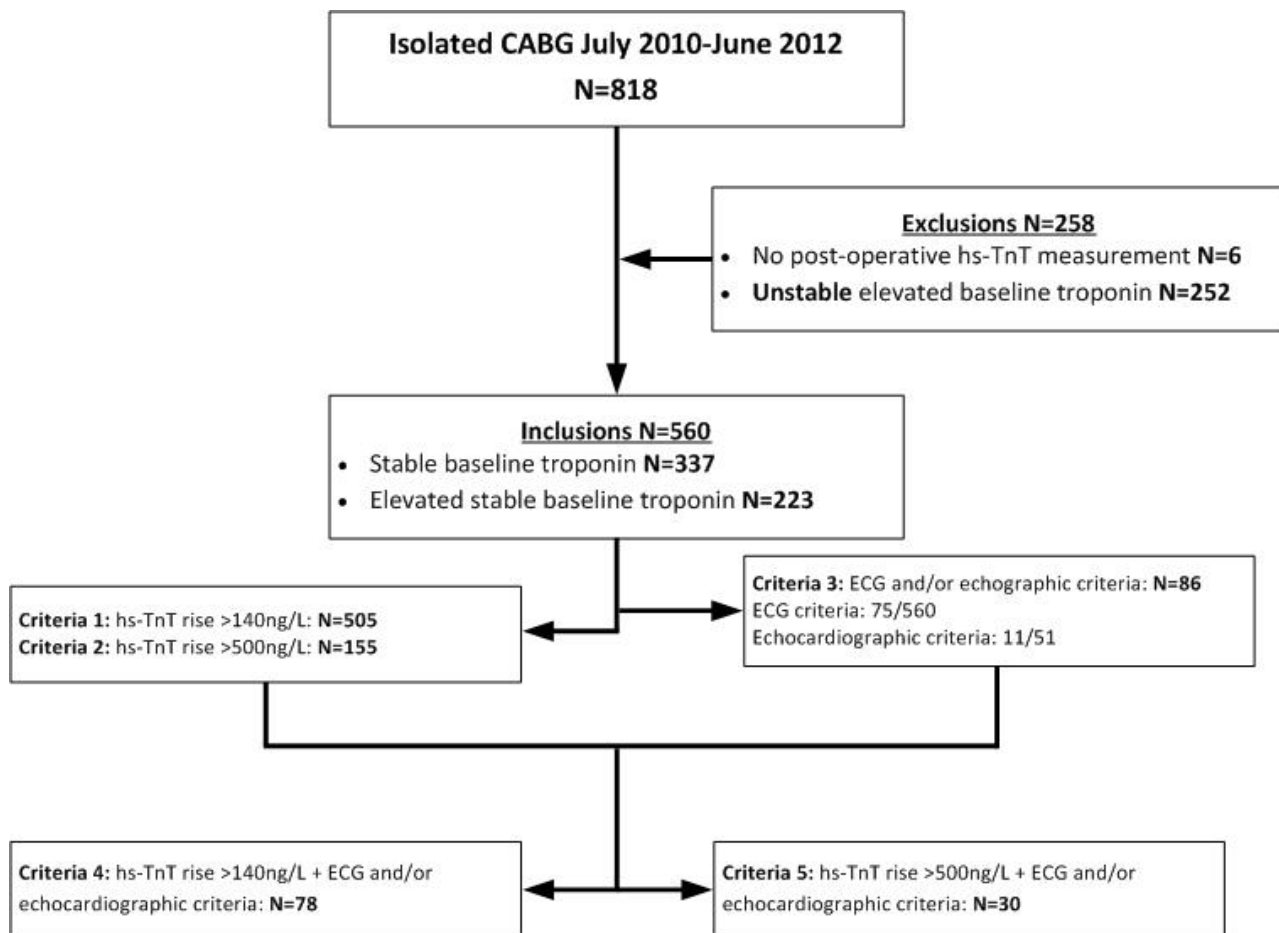
Clinical characteristics

Table 10.1 shows the baseline characteristics categorized by post-operative hs-TnT levels of ≤140ng/L, 141-500ng/L and >500ng/L. The medians (lower quartile, upper quartile) of hs-TnT in these groups were 116 (99, 129)ng/L, 282 (225, 375)ng/L and 730 (603, 1100)ng/L respectively. There was no association between the time from operation to post-operative hs-TnT measurement and hs-TnT level (Pearson coefficient r=0.01, p=0.37).

Table 10-1 Baseline characteristics

Post-operative high-sensitivity troponin T levels (ng/L)	≤140 (n=38)	141-500 (n=360)	>500 (n=162)	P-value
Demographics				
Age (years)	61.4 (9.9)	64.2 (9.9)	65.1 (9.6)	0.105
Male (%)	76.3% (29)	80.8% (291)	80.2% (130)	0.800
Ethnicity (%)				0.150
Caucasian	57.9% (22)	58.3% (210)	46.3% (75)	
Maori/Pacific Islander	23.7% (9)	20.8% (75)	25.3% (41)	
Other	23.7% (9)	20.8% (75)	28.4% (46)	
Body mass index (kg/m^2)	29.6 (5.8)	29.4 (5.4)	28.4 (4.9)	0.102
Presentation				
Canadian Cardiovascular Society class IV angina (%)	39.5% (15)	35.0% (126)	39.5% (64)	0.571
New York Heart Association class IV dyspnoea (%)	5.3% (2)	2.5% (9)	6.2% (10)	0.109
Recent myocardial infarction within six weeks (%)	36.8% (14)	43.1% (155)	50.6% (82)	0.163
Pre-operative intra-aortic balloon pump (%)	7.9% (3)	5.8% (21)	10.5% (17)	0.166
Urgent operation (%)	71.1% (27)	70.3% (253)	73.5% (119)	0.759
Past medical history				
Myocardial infarction (%)	55.3% (21)	63.3% (228)	66.7% (108)	0.405
Percutaneous coronary intervention (%)	13.2% (5)	9.2% (33)	14.2% (23)	0.209
Coronary artery bypass grafting (%)	2.6% (1)	0.6% (2)	3.7% (6)	0.026

Congestive heart failure (%)	5.3% (2)	4.1% (15)	4.9% (8)	0.897
Atrial fibrillation (%)	2.6% (1)	6.7% (24)	7.4% (12)	0.565
Diabetes (%)	36.8% (14)	35.0% (126)	43.2% (70)	0.200
Diabetes on insulin (%)	10.5% (4)	7.5% (27)	14.2% (23)	0.055
Hypercholesterolemia	92.1% (35)	91.7% (330)	90.7% (147)	0.930
Hypertension	63.2% (24)	71.1% (256)	59.9% (109)	0.464
Current smoker (%)	15.8% (6)	13.6% (49)	12.3% (20)	0.837
Stroke (%)	5.3% (2)	5.6% (20)	4.9% (8)	0.959
Peripheral vascular disease (%)	0.0% (0)	10.3% (37)	13.0% (21)	0.062
Chronic respiratory disease (%)	13.2% (5)	17.8% (64)	19.8% (32)	0.622
Dialysis (%)	0.0% (0)	0.6% (2)	8.0% (13)	<0.001
Investigations				
Left main stem stenosis >50% (%)	55.3% (21)	39.2% (141)	48.1% (78)	0.044
Three-vessel disease (%)	71.1% (27)	77.2% (278)	87.0% (141)	0.014
Ejection fraction (%)				0.432
Normal ($\geq$50%)	73.7% (28)	74.7% (269)	67.3% (109)	
Mild impairment (40-49%)	13.2% (5)	14.2% (51)	14.8% (24)	
Moderate impairment (30-39%)	7.9% (3)	6.4% (23)	12.3% (20)	
Severe impairment (<30%)	5.3% (2)	4.6% (17)	5.6% (9)	
Estimated glomerular filtration rate (mL/min)	95 (34)	82 (25)	74 (35)	<0.001

Pre-operative troponin groups				<0.001
Normal (%)	84.2% (32)	62.2% (224)	50.0% (81)	
Stable elevated (%)	15.8% (6)	37.8% (136)	50.0% (81)	
EuroSCORE II	1.8% (1.0%)	2.2% (3.1%)	2.7% (3.5%)	0.046

Post-operative outcomes

Table 10.2 shows operative and post-operative variables according to post-operative hs-TnT levels. Thirty day mortality, medium-term mortality and composite morbidity were 1.8%, 2.9% and 16.3% respectively.

Table 10-2 Operative and post-operative variables

| Post-operative high-sensitivity troponin T levels (ng/L)	≤140 (n=38)	141-500 (n=360)	>500 (n=162)	P-value
Operation details				
Off-pump (%)	13.2% (5)	1.7% (6)	2.5% (4)	<0.001
Number of distal anastomoses				0.002
1	5.3% (2)	1.1% (4)	0.6% (1)	
2	28.9% (11)	16.1% (58)	10.5% (17)	
3	44.7% (17)	47.2% (170)	55.6% (90)	
4	18.4% (7)	31.9% (115)	23.5% (38)	
5	2.6% (1)	3.3% (12)	9.3% (15)	
6	0.0% (0)	0.3% (1)	0.6% (1)	
Left internal mammary artery graft (%)	94.7% (36)	98.3% (354)	96.3% (156)	0.204
Right internal mammary artery graft (%)	2.6% (1)	3.9% (14)	10.5% (17)	0.008
Radial artery graft (%)	18.4% (7)	26.1% (94)	19.1% (31)	0.164
Saphenous vein grafts (%)	86.8% (33)	93.3% (336)	95.7% (155)	0.129
Cardiopulmonary bypass time (minutes)	85 (32)	89 (23)	97 (33)	0.004
Aortic cross-clamp time (minutes)	54 (25)	59 (19)	63 (23)	0.080
Post-operative Outcomes				
ECG (new Q wave or left bundle branch block %)	18.4% (7/38)	13.1% (47/358)	13.3% (21/158)	0.630
Echocardiogram (new regional wall motion abnormalities %)	0.0% (0/1)	8.7% (2/23)	33.3% (9/27)	0.094

ECG and/or echocardiographic criteria (%)	18.4% (7/38)	13.7% (49/358)	19.0% (30/158)	0.271
Composite morbidity (%)	15.8% (6)	10.6% (38)	29.0% (47)	<0.001
Stroke (%)	0.0% (0)	1.1% (4)	1.9% (3)	0.603
Renal failure (%)	5.3% (2)	1.1% (4)	1.9% (3)	0.147
Prolonged ventilation >24hours (%)	10.5% (4)	6.7% (24)	22.8% (37)	<0.001
Deep sternal wound infection (%)	0.0% (0)	0.3% (1)	0.0% (0)	0.757
Reoperation (%)	2.6% (1)	3.9% (14)	6.2% (10)	0.430
Operation to discharge (days)	6.7 (2.8)	7.8 (5.6)	8.9 (5.7)	0.028
Re-admit to hospital within 30 days (%)	26.3% (10)	19.7% (71)	20.4% (33)	0.631
30-day mortality (%)	2.6% (1)	1.1% (4)	3.1% (5)	0.266
Discharge medications				
Aspirin (%)	97.3% (36/37)	98.9% (353/357)	98.7% (155/157)	0.716
Statin (%)	94.6% (35/37)	89.9% (321/357)	86.0% (135/157)	0.227
Beta-blocker (%)	89.2% (33/37)	78.7% (281/357)	73.9% (116/157)	0.113
Angiotensin converting enzyme blockers or Angiotensin II receptor blocker (%)	32.4% (12/37)	30.0% (107/357)	29.3% (46/157)	0.932

Figure 10.2 shows the Kaplan-Meier survival curves for the five potential MI criteria. Isolated hs-TnT >140ng/L was not associated with mortality (Figure 10.2a), p=0.521. Isolated hs-TnT >500ng/L was associated with mortality HR 2.60, 95% CI (1.03-7.08), p=0.046 (Figure 10.2b). ECG and/or echocardiographic criteria alone was also associated with mortality; HR 3.32, 95%CI (1.21-9.15), p=0.020 (Figure 10.2c).

When the ECG and/or echocardiographic criteria for MI were added to hs-TnT rise, the associations with mortality were strengthened. One-year survival rates for those with and without hs-TnT >140ng/L and ECG and/or echocardiographic criteria were 92.3% and 97.9%; HR 3.76, 95% CI (1.37-10.4), p=0.010 (Figure 10.2d); and for hs-TnT >500ng/L and ECG and/or echocardiographic criteria were 86.7% and 97.7%; HR 6.08, 95% CI (1.96-18.9), p=0.002 (Figure 10.2e).

Figure 10-2 Kaplan-Meier survival curves for the five MI criteria

122
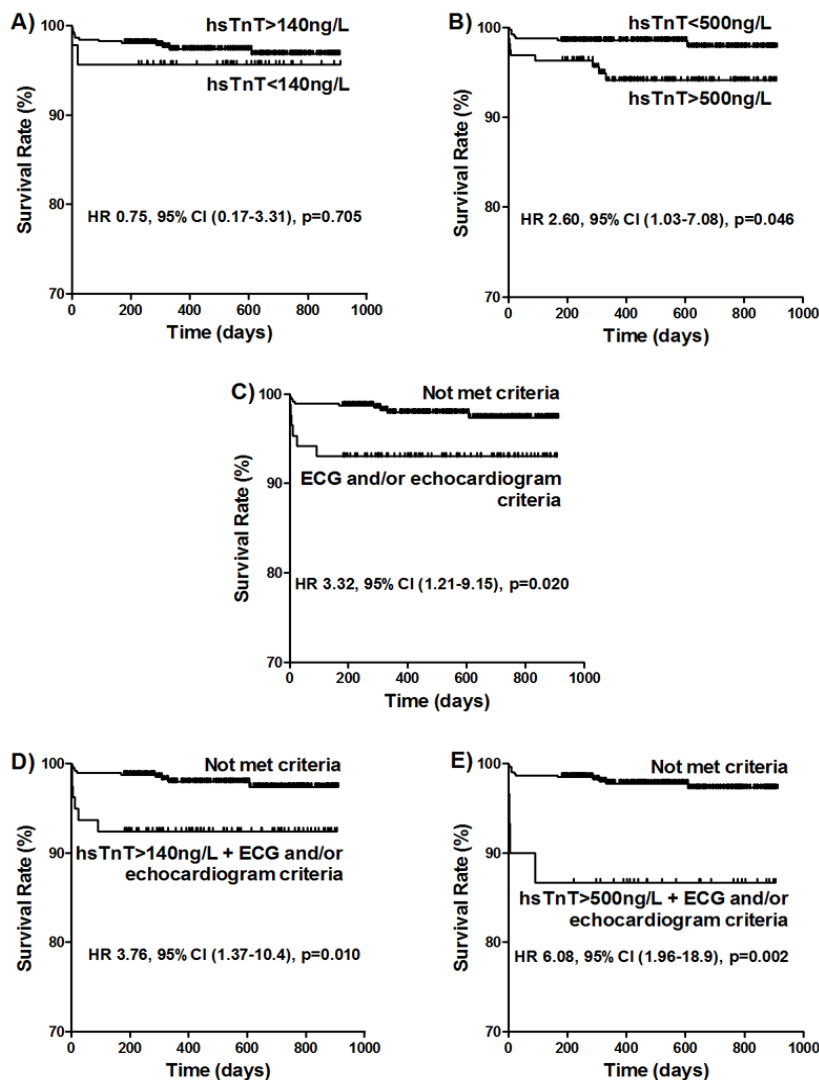
Receiver-operative characteristics analyses

Table 10.3 lists results from the ROC analyses. Dual criteria was best at detecting 30-day mortality AUC 0.693, 95%CI (0.521-0.865), p=0.036 and medium-term mortality AUC 0.682, 95%CI 0.521-0.844), p=0.027. Hs-TnT >140ng/L and ECG and/or echocardiographic data was the only criteria detecting 30-day mortality AUC 0.683, 95%CI (0.519-0.857), p=0.029 and medium-term mortality AUC 0.621 (0.507-0.743), p=0.045.

For detecting composite morbidity, hs-TnT alone had the highest AUC 0.658, (0.589-0.726), p<0.001, although dual criteria was nearly as good AUC 0.645, 95%CI (0.577-0.714), p=0.002. A cut-point of hs-TnT >500ng/L had the highest AUC 0.603, 95%CI (0.548-0.658), p<0.001.

The hs-TnT levels with maximum sensitivity x specificity for detecting 30-day and medium-term mortality were 580ng/L alone or 200ng/L as dual criteria. Corresponding figures for detecting composite morbidity were 410ng/L and 200ng/L.

Table 10-3 Receiver-operating characteristics analysis and areas under curve with corresponding 95% confidence interval

Criteria	Operative mortality	Medium-term mortality	Composite morbidity
hs-TnT (continues)	0.606 (0.379-0.832)	0.634 (0.465-0.804)	0.658 (0.589-0.726)
ECG and/or echo criteria	0.676 (0.512-0.840)	0.613 (0.490-0.737)	0.540 (0.494-0.586)
hs-TnT (continues) + ECG and/or echocardiographic criteria	0.693 (0.521-0.865)	0.682 (0.521-0.844)	0.645 (0.577-0.714)
hs-TnT >140ng/L	0.551 (0.420-0.683)	0.513 (0.429-0.598)	0.520 (0.483-0.558)
hs-TnT >500ng/L	0.565 (0.404-0.726)	0.617 (0.489-0.745)	0.603 (0.548-0.658)
hs-TnT >140ng/L + ECG and/or echocardiographic criteria	0.683 (0.519-0.857)	0.641 (0.527-0.763)	0.549 (0.503-0.594)
hs-TnT>500ng/L + ECG and/or echocardiographic criteria	0.625 (0.475-0.775)	0.601 (0.491-0.711)	0.561 (0.522-0.599)

All figures are areas under receiver-operative characteristics curve and 95% confidence intervals, hs-TnT = high-sensitivity troponin T

Multivariate analyses

Table 10.4 shows on multivariate analyses, the features associated with patients having hs-TnT >140ng/L and ECG and/or echocardiographic criteria for MI (n=78/560, 13.9%) were CCS class 4 angina (p=0.017), and longer cardiopulmonary bypass time (p=0.048). The features associated with patients having hs-TnT >500ng/L and ECG and/or echocardiographic criteria for MI (n=30/560, 5.4%) were dialysis (p=0.032) and longer cardiopulmonary bypass time (p=0.006).

Table 10-4 Multivariate predictors of troponin rise and criteria for myocardial infarction

Predictors	OR	95%CI	P-value
hs-TnT >140ng/L (n=505)			
Peripheral vascular disease	11.5	0.74-179	0.082
Estimated glomerular filtration rate (per 10ml/min decrease)	1.15	1.04-1.27	0.024
Off-pump	0.10	0.04-0.32	<0.001
hs-TnT >500ng/L (n=155)			
Body mass index (per 1 kg/m^2)	0.96	0.92-1.00	0.075
Dialysis	8.78	1.79-43.1	0.008
Previous coronary artery bypass grafting	4.37	1.02-18.6	0.047
Three-vessel disease	1.75	0.97-3.15	0.063
Cardiopulmonary bypass time (per 10 minutes)	1.08	1.00-1.17	0.038
ECG and/or echocardiographic criteria (n=86) None			
hs-TnT >140ng/L + ECG and/or echocardiographic criteria (n=78)			
Canadian Cardiovascular Society Class IV angina	1.94	1.13-3.33	0.017
Cardiopulmonary bypass time (per 10 minutes)	1.09	1.00-1.18	0.048
hs-TnT >500ng/L + ECG and/or echocardiographic criteria (n=30)			
Diabetes on insulin	2.46	0.91-6.71	0.078
Dialysis	4.58	1.14-18.4	0.032
EuroSCORE II [30]	1.11	1.00-1.22	0.061
Cardiopulmonary bypass time (per 10 minutes)	1.17	1.05-1.31	0.006

OR = odds ratio, 95% CI = 95% confidence interval, hs-TnT high sensitivity troponin T

Table 10.5 shows the variables predictive for 30-day and medium-term mortality. Both hs-TnT >140ng/L with ECG and/or echocardiographic criteria and ECG and/or echocardiographic criteria alone, predicted 30-day mortality; HR 6.12, 95% CI 1.50 – 25.0, p=0.012 and HR 5.93, 95% CI 1.46 – 24.2, p=0.013, respectively. Other predictors of 30-day mortality were NYHA class 4, dyspnea, (p=0.016); not using left internal mammary grafts, (p=0.008); and return to theatre (p=0.049).

Factors predicting medium-term mortality were estimated glomerular filtration rate (p=0.021), not using left internal mammary grafts (p=0.006) and level of post-operative hs-TnT (p=0.045). Levels of hs-TnT >140ng/L with ECG and/or echocardiographic criteria also predicted medium-term survival; HR 3.44, 95% CI 1.13 – 10.5, p=0.031 but hs-TnT >500ng/L with ECG and/or echocardiographic criteria did not, p=0.125.

Table 10-5 Multivariate predictors of mortality

Predictors	OR or HR	95%CI	P-value
30-day mortality	OR		
Female	3.74	1.03-13.6	0.045
Maori/Pacific Islander	3.25	0.82-13.0	0.095
New York Heart Association Class IV dyspnoea	11.4	2.58-50.8	0.001
Left internal mammary artery graft	0.06	0.01-0.63	0.02
Post-operative hs-TnT (per 100ng/L)	1.07	0.93-1.23	0.36
hs-TnT >140ng/L	0.37	0.06-2.23	0.277
hs-TnT >500ng/L	0.84	0.21-3.34	0.803
ECG and/or echocardiographic criteria	4.68	1.28-17.2	0.020
hs-TnT >140ng/L + ECG and/or echocardiographic criteria	4.92	1.34-18.1	0.017
hs-TnT>500ng/L + ECG and/or echocardiographic criteria	2.59	0.46-14.6	0.281
Medium-term mortality	HR		
Estimated glomerular filtration rate (per 10ml/min decrease)	1.34	1.03-1.73	0.021
Left internal mammary artery graft	0.11	0.02-0.53	0.006
Post-operative hs-TnT (per 100ng/L)	1.04	1.00-1.08	0.045
hs-TnT >140ng/L	1.03	0.20-5.41	0.970
hs-TnT >500ng/L	1.51	0.51-4.42	0.456
ECG and/or echocardiographic criteria	3.28	1.08-9.98	0.036

hs-TnT >140ng/L + ECG and/or echocardiographic criteria	3.44	1.13-10.5	0.030
hs-TnT >500ng/L + ECG and/or echocardiographic criteria	2.73	0.76-9.80	0.123

OR = odds ratio, HR=hazards ratio, 95%CI = 95% confidence interval, hs-TnT = high-sensitivity troponin T

Table 10.6 shows the predictors of post-operative composite morbidity. These included Maori or Pacific Island ethnicity, p=0.021, use of an intra-aortic balloon pump, p<0.001, previous CABG, p=0.037 and cardiopulmonary bypass time, p=0.002.

The level of post-operative troponins was also predictive of composite morbidity, p=0.014. Isolated levels of hs-TnT >500ng/L were predictive, p=0.005 as well as hs-TnT levels >500ng/L and ECG and/or echocardiographic criteria for MI (p=0.022).

Table 10-6 Multivariate predictors of composite morbidity

Predictor	OR	95%CI	P-value
Maori/Pacific Islander	1.86	1.02-3.40	0.043
Pre-operative intra-aortic balloon pump	5.39	2.29-12.7	<0.001
Previous coronary artery bypass grafting	7.12	1.13-44.7	0.037
Ejection fraction (<40%)	1.95	0.95-4.01	0.072
EuroScore II	1.09	1.01-1.16	0.035
Cardiopulmonary bypass time (per 10 minutes)	1.16	1.06-1.27	0.002
Post-operative hs-TnT (per 100ng/L)	1.06	1.01-1.10	0.014
hs-TnT >140ng/L	0.91	0.36-2.32	0.846
hs-TnT >500ng/L	2.20	1.27-3.82	0.005
ECG and/or echocardiographic criteria	1.64	0.84-3.21	0.151
hs-TnT >140ng/L + ECG and/or echocardiographic criteria	1.75	0.89-3.46	0.098
hs-TnT >500ng/L + ECG and/or echocardiographic criteria	3.02	1.18-7.76	0.022

OR = odds ratio, 95%CI = 95% confidence interval, hs-TnT = high-sensitivity troponin T

10.5 Discussion

There are several novel findings of this study. Firstly, using a guideline compliant high sensitivity assay(153) dual criteria of hs-TnT levels >140ng/L associated with ECG and/or echocardiographic criteria, were predictive of 30-day and medium-term mortality after CABG. A higher cutpoint of >500ng/L associated with ECG and/or echocardiographic changes was not associated with mortality.

Secondly, the finding of elevated hs-TnT >10 times the URL and associated ECG and/or echocardiographic evidence of MI validates the recommendation of Third Universal Definition of MI with hs-TnT(72). Thirdly, higher hs-TnT levels predicted post-operative morbidity.

Prognostic value of isolated high-sensitivity troponin T rise

Post-operative hs-TnT alone as a continuous parameter independently predicted medium-term mortality and composite morbidity. Previous studies with various contemporary troponin T and troponin I assays have reported similar findings(148-151, 155). The level of post-operative hs-TnT >140ng/L (10 times the URL) was found in 90% of patients, whereas hs-TnT >500ng/L (36 times the URL) were found in 29% of patients.

Isolated hs-TnT >140ng/L was not associated with mortality or composite morbidity. Isolated hs-TnT >500ng/L was associated with medium-term mortality on univariate analysis and predicted composite morbidity. These findings of a requirement of a high cut-point are similar to the findings of other studies with various assays, including point of care, none of which are guideline compliant(156), using cut-points 7.8-170 times URL(149, 152, 157, 158). However we found on multivariate analysis that isolated hs-TnT >500ng/L was not related to mortality. Hs-TnT >500ng/L was an independent predictor of composite morbidity, indicating elevation in hs-TnT is a marker of myocardial injury due to multiple causes and not restricted to ischemia e.g.: heart failure, sepsis and renal failure(60).

ECG and/or echocardiographic features of myocardial infarction

This study shows that new signs of infarction on the ECG and/or evidence of new wall motion abnormality on echocardiography are important predictors of post-CABG mortality. The prevalence of ECG signs of infarction was relatively uniform across the three thresholds of hs-TnT rise, including below 140ng/L. This suggests that if an isolated hs-TnT threshold is set too high, some patients with poor prognosis will be missed. ECG and/or echocardiographic criteria however did not predict composite post-operative morbidity, showing their specificity for MI.

There were 11 of 51 patients that had new regional wall motion abnormalities on post-operative echocardiograms, all of whom had no new ECG changes but met the Universal Definition criteria for MI. The diagnosis of MI would have been missed if an echocardiogram was not performed in these patients.

Dual elevation of high sensitivity troponin plus ECG and/or echocardiographic criteria

Our results showed that dual criteria are more prognostic of mortality than single criteria. Fourteen percent (78/560) of patients had an MI as defined by the dual criteria corresponding to the Universal

Definition in our cohort, where hs-TnT threshold was 140ng/L (10 times 99th percentile URL). This criteria was the strongest predictor of both 30-day and medium-term mortality in both the ROC and multivariate analyses amongst the five criteria assessed. It utilises the strengths of both hs-TnT's sensitivity for myocardial injury with the specificity of new Q-waves and new LBBB on ECG for myocardial infarction.

Interestingly, a higher threshold for hs-TnT >500ng/L with ECG and/or echocardiographic criteria did not independently predicted mortality at 30-days on medium-term. This may be because the dual criteria, although very specific for adverse outcomes, compromise its sensitivity and therefore fails to capture the majority of post-operative mortality throughout the follow-up period. As the ECG and/or echocardiogram is very specific for MI after CABG, a lower threshold for hs-TnT rise for dual criteria than single criteria has better prognostic value, as shown by our results for hs-TnT>140ng/L with ECG and/or echocardiogram criteria.

The highest AUCs for mortality we found even with dual criteria were 0.64-0.70, while previous studies have reported area under ROC curves of 0.73-0.82 using single criteria with contemporary I or T troponin assays(148, 152). A potential reason for our finding of a lower area under the ROC curve may be that the superior sensitivity of the hs-TnT assay may have compromised its specificity and positive predictive value

The biomarker cutpoint for the Universal Definition for CABG MI in 2007 was 5 times the URL(159). The increase to 10 times the URL threshold for biomarker evaluation in the 2012 Universal Definition of MI URL was arbitrary(72). Troponins are preferred because of their sensitivity and specificity(72, 154). Creatine kinase MB, although a proven predictor of mortality after CABG, has been shown to be inferior to troponins in predicting adverse cardiovascular outcomes(147, 148).

The importance of stable baseline troponin levels

To diagnose MI appropriately, a stable baseline of troponin is required. We thus included patients with 'normal stable baseline', and `elevated stable baseline troponins'. Patients with elevated but rising troponin levels were excluded because post-operative troponin rises cannot be accurately distinguished from an index MI and be attributed to an MI after CABG(160).

The Universal Definition for MI after CABG does not define the amount of rise required in patients with 'elevated stable baseline troponins'. We used the 20% elevation from baseline criteria, defined by the Universal Definition for MI with PCI(72).

Independent predictors of myocardial infarction and mortality

Unstable angina and cardiopulmonary bypass time were identified as independent predictors of MI (defined as hs-TnT >140ng/L with ECG and/or echocardiographic criteria), which have been previously reported with other definitions of MI(161).

In addition to MI, predictors of 30-day or medium term mortality were female sex, dyspnea NYHA class IV, and reduced renal function, all of which are incorporated in the EuroSCORE(26) and Society of Thoracic Surgeon's (STS) score(32). We also found, as reported by others(162), that revascularization with the left internal mammary artery was significantly associated with lower mortality.

Implications

The presence of predictors of MI should alert clinicians of the importance of their management. As the occurrence of MI is associated with mortality and other adverse outcomes, such as inotropic requirement(150) and longer ventilation time(161), treatments proven to improve survival after CABG, including aspirin(163) statins(164) and beta-blockers(165), should be administered. A lower threshold for angiography and return to theatre, in patients requiring inotropes or with prolonged ventilation should be considered(149).

The 30-day mortality rate is similar to that reported for 2009 by the STS database; 1.8% vs 1.9% STS(166). However, the rate of MI we found (14%) is higher than previously reported using conservative definitions. This means that more patients will be labelled with the diagnosis. However there is an opportunity to improve outcomes in these patients. More research is required to assess which therapies may reduce the occurrence of MI, and adverse outcomes after MI.

Study limitations

This is a retrospective observational study from one center. About 20% of patients were excluded because of rising or missing baseline troponin levels. The timing of troponin measurements post-operatively was at 12-24 hours and not at a fixed time. The moderate cohort size means some of the outcomes might have been underpowered for analyses. A small number of patients had post-operative cardiac imaging such as echocardiography, based on clinical indications.

10.6 Conclusion

This study shows the utility of hs-TnT in detecting MI using a cutpoint of 10 times the 99th percentile URL with ECG and/or echocardiographic findings. Diagnosis of MI with dual criteria was related to 30-day and medium-term mortality. These results validate the Third Universal Definition of MI using hs-TnT as the biomarker.

11 Conclusions

11.1 Summary of findings

My thesis has contributed novel research and insights into risk modelling for cardiac surgery and interventions. After reviewing the literature on the statistics of risk modelling and cardiac surgical risk models, the applicability of contemporary surgical risk models in New Zealand cardiac surgery cohorts was investigated, followed by assessment of the performance of risk models in the important clinical settings of infective endocarditis surgery and TAVI, and finally evaluation of the prognostic utility of the cardiac biomarker high-sensitivity troponin T for mortality and diagnosing peri-operative MI after cardiac surgery.

In Chapters 3-6, performance of EuroSCORE, EuroSCORE II, STS Score and ANZCTS Scores (where applicable) for isolated CABG, isolated AVR, isolated mitral valve repair or replacement, and AVR+CABG were assessed and compared. For isolated CABG, all scores had similar discrimination (c-statistic 0.64-0.68), however the EuroSCORE significantly over-estimated operative mortality while the other scores provided adequate calibration to fit contemporary outcomes. In the isolated AVR cohort, all scores again had similar moderate discrimination (c-statistic 0.68-0.75), but the STS Score performed best for high risk patients and complications. The EuroSCORE II and STS Score had improved discrimination and calibration over the EuroSCORE in AVR+CABG patients. Lastly, all scores had high discrimination (c-statistic 0.82-0.85) for mortality after isolated mitral valve repair or replacement, although again the STS Score was the best predictor for morbidities.

Comparing the types of surgeries, the scores were most accurate for mitral surgery, followed by AVR, AVR+CABG and lastly isolated CABG. This is somewhat surprising given that isolated CABG is the most common cardiac operation performed and constitutes the largest proportion of the derivation cohort for risk models. In part this is reflected by the unique characteristics of our cohort, in particular, Maori and Pacific ethnicity was an independent predictor of adverse outcomes after CABG. Possible reasons that these patients have higher risk include on average greater severity of cardiac disease at presentation, prevalence of comorbidities, socioeconomic deprivation, and reduced access to healthcare and understanding of their health condition. Effective interventions to reduce ethnic disparities in outcomes are therefore urgently needed. The STS Score had the best discrimination for post-operative complications especially for AVR and mitral valve surgery, an advantage of its overall and separate morbidities risk models. It was also the best score for calibration of AVR patients in the highest surgical risk quintile, suggesting it is the best one to use in higher risk patients to guide treatment modality selection and consideration for TAVI. The fact that risk models performed best for MVR may be due to

this cohort being the youngest, and therefore less affected by factors such as frailty not included in current surgical risk models, as well as having the highest proportion of stable patient having outpatient non-urgent surgery compared to other types of cardiac surgery.

Risk scores for infective endocarditis surgery, a high risk and heterogeneous condition, was explored in Chapters 7 and 8 with some novel findings. In the local endocarditis surgery cohort, the general cardiac surgery scores only performed modestly. The De Feo-Cotrufo Score which is endocarditis-specific had superior discrimination to the EuroSCOREs at predicting mortality and post-operative complications after endocarditis surgery, and is therefore preferred, however calibration of endocarditis-specific scores couldn't be assessed as the scores were additive. The meta-analysis did find the EuroSCOREs to have moderate discrimination for operative mortality but could not assess endocarditis-specific scores which had not been externally validated in the past, except in our study.

The widely expanding TAVI procedure and its risk modelling were assessed as a meta-analysis in Chapter 9. The EuroSCORE, EuroSCORE II and STS score weakly predicted TAVI outcomes compared to other cardiac surgery. Furthermore, the EuroSCORE over-estimated risk whereas the other two scores had better calibration. The modest performance of surgical risk scores is within expectations, given that they were derived from surgical cohorts rather than TAVI, which present unique benefits and challenges compared to cardiac surgery. There is clearly a need to developing more accurate risk models for TAVI in the future.

In Chapter 10 peri-operative MI diagnosed using high-sensitivity troponin rises with or without concurrent ECG and/or echocardiographic abnormalities after isolated CABG was assessed. The universal definition's dual criteria including cutpoint at 10 times 99th percentile upper reference limit of troponin was shown to be most prognostic for short and medium-term mortality and therefore validated for the first time. For composite morbidity, isolated high-sensitivity troponin elevation was the best predictor compared to other criteria. Important independent predictors of peri-operative type 5 MI and other adverse outcomes were also identified. Having a validated universal definition for diagnosis enables identification and targeted management in clinical practice and consistency in endpoint reporting in clinical research.

11.2 Clinical implications

Risk scores are used by clinicians to guide clinical decision-making. For cardiac procedures, this include whether to intervene, and which modality. Such objectively measures allow evaluation of benefit and risk, providing context for discussion amongst the multidisciplinary heart team, and discussion with patients as part of informed consent. Publication of full logistic models such as the EuroSCOREs and

STS Scores rather than just additive models give a quantitative measure of absolute risk and is strongly encouraged. The clinical decision is usually easier when the risk is low to go ahead, or prohibitively high to choose for conservative non-operative management. Their greatest value is in the moderate to high risk patients, to not only guide decision-making, but also to alert which patients need closer attention and monitoring as well as optimisation of their clinical state peri-operatively. This is true even in diseases where surgery is mandatory regardless of risk assessment as the prognosis is otherwise exceedingly poor without it.

How do the findings of this thesis apply to clinical practice? Amongst existing surgical risk models, the original EuroSCORE should no longer be used, and be replaced with the more contemporary EuroSCORE II or STS Score for cardiac surgery workup. Although EuroSCORE II has the advantage for being easier to calculate, the STS Score has benefits of not only good performance, but also accuracy in high risk patients and for discriminating and calibrating post-operative complication, so is generally preferred. The ANZSCTS Scores have not demonstrated improved accuracy in New Zealand cohorts over the other scores and also don't have online calculators, so is less user friendly. Limitations of the scores particularly for surgeries which it performs less well in, or certain subgroups such as high risk patients, need to be appreciated, and shouldn't take away the holistic and multidisciplinary assessment and clinical judgement for each individual patient.

The same concepts apply to using risk models in contexts of endocarditis surgery, TAVI and other cardiac interventions. Endocarditis-specific risk scores appear to have improved discrimination, however they are not published as logistic scores which limits their use in terms of calibration and providing an estimate of mortality risk. Nevertheless, the presence of proven mortality predictors and higher score warrants careful consideration whether the surgery should proceed, and if so, pre-operative and peri-operative optimisation of medical care. The suboptimal performances of surgical risk models for TAVI compared to cardiac surgery are clear. Clinicians should understand this and that their use in this context should be mainly for estimating and stratifying the risk of the alternative modality of surgical AVR for the patient which remains important, rather than TAVI directly.

Findings from Chapter 10 also have important implications. Assessment of type 5 peri-operative MI requires routine measurements of high-sensitivity troponins pre and post-operatively, to firstly establish whether baseline troponins are normal, stable or not as a prerequisite, and then the degree of rise 12-24 hours after cardiac surgery to establish if it meets the guideline threshold or not. Similarly, ECGs should be performed in everyone before and daily after surgery to detect new Q-waves or LBBB, and where there is clinical suspicion, post-operative echocardiogram should be performed to identify any new regional wall motion abnormalities. The dual criteria is strongly prognostic of mortality both short and

long-term, while an isolated troponin rise is associated with morbidity, and therefore additional care must be taken when these or their predictors are present to attempt to enhance outcomes, with a lower threshold for subsequent treatment and interventions.

11.3 Limitations

The studies in this thesis had several limitations to acknowledge. The six cohort studies were all single-centre, retrospective and observational, although this is the commonest way for evaluating and externally validating risk models, while reflecting real world experience. The size of each cohort and clinical events meant that power was moderate at best. This meant that not all statistically significant differences in risk model performance and predictors of outcomes could be identified, and also insufficient for constructing a new model. For example, the confidence intervals for c-statistics were often wide and overlapping, making it hard to determine if one score truly performs better than another. As such, c-statistics were not compared in a statistical manner with either the DeLong or the Hanley and McNeil tests. Where risk models had only been published as an additive score, a complete assessment including calibration was not possible. Some important pre and post-operative characteristics were not collected, such as frailty and long-term outcomes such as symptoms, quality of life, morbidities and investigations at follow-up. Duration of follow-up was restricted given that contemporary surgical cohorts were studied, but long term outcomes beyond 5-10 years are important considerations for all patients. Yet each study was able to, within these limitations, assess performance of risk models specifically applied to a local New Zealand cohort and reported their unique characteristics, outcomes and predictors.

Existing surgical risk models have their own problems too that can restrict utility. They are all based on historical populations, with parameters limited to what was routinely collected in their respective databases, which limit the accuracy to contemporary outcomes. The developmental cohort's demographics, clinical status, past history, investigations, management and outcomes may all differ from the cohort of interest. Retrospective calculation of risk scores could also introduce unintended bias, and in particular may affect those having inpatient urgent operations more whereby their clinical status may change rapidly and day-to-day which can alter the risk calculation. These can explain the performance of risk models especially when it is modest, and again the need to evaluate their performance locally to determine their use with or without recalibration, or should not be used. Regular revisions of risk models which have been done for EuroSCOREs, STS Scores and ANZCTS Scores are also important to reflect changes in disease patterns, clinical practice and outcomes.

The meta-analyses performed also had some limitations in power, which makes identify all statistically significant differences and associations difficult, although they were still much higher than individual

studies. Another weakness was that patient-level data were unavailable despite attempts to obtaining this for the majority of the individual studies. This enables closer evaluation of the heterogeneity of study design, patient characteristics, management and outcomes between the cohorts pooled. It also allows meta-regression and other techniques to be performed to look for predictors of adverse outcomes and subgroup analyses. Publication bias is another potentially limitation for meta-analysis where studies with less impressive or non-significant findings are less likely to be published, although this was evaluated in both meta-analyses performed and not felt to have significant impact on the findings.

11.4 Future directions

There is great scope for further research pertinent to all the topics of this thesis. Firstly, there will be a move towards procedure, disease and at times subgroup specific models, with a wider range of outcomes. The need can be seen from the modest performance of general cardiac surgery scores in endocarditis surgery and TAVI in previous chapters. TAVI-specific risk models and their validation are urgently needed due to the blossoming of this procedure. In some recent randomised trials TAVI has superior outcomes to AVR so calibration of surgical scores would be suboptimal, but TAVI has unique adverse predictors to cardiac surgery that aren't part of surgical scores such as transfemoral versus other access, annular and outflow tract calcification and coronary heights from the annular plane. Having large populations and numbers of events are mandatory to risk model construction, although this is not always available especially for rare performed procedures and diseases. Multi-centre registries and cooperation and longer recruitment times may be necessary. The range of models developed would then need to be assessed to determine which performs best. Taking the subgroup gender as an example, comparison can be made between a general risk model with gender as a parameter, and gender-specific risk models to see which one performs better. If the general risk model performs similarly then it can be utilised, but if significantly inferior, then this implies that other parameters have different level of importance and interactions in males and females, and gender-specific risk models should be used. Such comparison can help clinicians understand what factors are important for each subgroup to better stratify risk. Separate risk models should also be made for other outcomes beyond operative mortality, such as long-term mortality and morbidities like the STS scores, as they all have important clinical, quality of life and cost implications. It remains important to have user friendly interfaces and calculators for this range of models that will likely be developed in the future.

Improving the discrimination of risk models will need more than just large cohorts and separate risk models. There are many known independent predictors of adverse outcomes not part of existing risk models, including clinical parameters such as frailty, imaging parameters such as right ventricular function, global longitudinal strain and late gadolinium enhancement, and biomarkers including

troponins and B-type natriuretic peptide. Ongoing research of all these classes of variables as well as novel biologically plausible markers including genomics and metabolomics will be important. These will need to be collected in large cohorts to identify their significant and interactions with other existing parameters to then be incorporated into newer risk models, otherwise they will be difficult to use as separate entities during risk assessment. Another problem is that only a minority of data in routinely collected in clinical records are accessible and entered into research databases for statistical modelling. Data access needs to be expanded to obtain a wider variety of parameters for research, and this includes seeking patient-level data for meta-analysis. With the big data available, incremental improvements in risk modelling can then be made, and undoubtedly newer statistical techniques including machine learning with artificial intelligence will play critical roles. These will help pave the way towards the future of personalised medicine. Finally, clinical trials are also warranted to assess whether applying risk model assessment of patients, and the resultant changes in management based on risk thresholds, improve clinical outcomes.

From the New Zealand standpoint, it is challenging but certainly achievable to try and develop local cardiac surgery risk models. This may be particularly important for our unique demographics such as Maori and Pacific ethnicities, and spread of aetiologies for cardiac diseases including rheumatic heart disease and endocarditis. Accurate registries with sufficient patients and clinical events would be required. Given our small country, national multi-centre participation and support and funding from the Ministry of Health and other related organisations are very helpful. This has been achieved locally in a different cardiology setting, with the substantial work over the last two decades in primary cardiovascular risk prediction tools and guidelines, and more recently, the undertaking of the All New Zealand acute coronary syndrome quality improvement registry(167, 168). Widespread uptake of these programmes has and will continue to provide quality data and risk prediction to inform local practice, and a national registry is urgently needed for cardiac surgery and procedures beyond cardiac catheterisation. Although some of the New Zealand cardiac surgery centres contribute to the ANZCTS database, our studies have shown them to provide no advantage over international EuroSCORE II and STS Scores, highlighting differences even in the populations on the two sides of the Tasman Sea. Enthusiasm amongst the cardiothoracic surgeons and the multidisciplinary cardiac team are needed to drive this initiative forward.

Finally, defining MI after cardiac surgery remains challenging. The current definition of and biomarker cutpoint for peri-operative type 5 MI was set arbitrarily(72), and although we have validated this larger studies should try to evaluate this criteria and potentially replicate our findings which is still lacking. The clinical utilities of other techniques in this context such as cardiac magnetic resonance imaging remain unknown. The existing criteria may or may not be applicable to other types of cardiac surgery and interventions beyond isolated CABG which needs evaluation. In addition, the optimal management

strategy to both prevent and treat this diagnosis is largely unknown where clinical trials are warranted. All of these points require further research to try and improve patient outcomes after cardiac surgery.

I am fascinated by and look forward to clinical advances in the field of my thesis and beyond in the literature, and am excited by my potential contributions and collaborations in this important area of cardiology and cardiac surgery in the future.

References

1.   Wang TK, Li AY, Ramanathan T, Stewart RA, Gamble G, White HD. Comparison of four risk scores for contemporary isolated coronary artery bypass grafting. Heart, lung & circulation. 2014;23(5):469-74.

2.   Wang TK, Choi DH, Stewart R, Gamble G, Haydock D, Ruygrok P. Comparison of four contemporary risk models at predicting mortality after aortic valve replacement. The Journal of thoracic and cardiovascular surgery. 2015;149(2):443-8.

3.   Wang TK, Choi DH, Haydock D, Gamble G, Stewart R, Ruygrok P. Comparison of Risk Scores for Prediction of Complications following Aortic Valve Replacement. Heart, lung & circulation. 2015;24(6):595-601.

4.   Wang TK, Choi DH, Ramanathan T, Ruygrok PN. Comparing Performance of Risk Scores for Combined Aortic Valve Replacement and Coronary Bypass Grafting Surgery. Heart, lung & circulation. 2016;25(11):1118-23.

5.   Wang TK, Harmos S, Gamble GD, Ramanathan T, Ruygrok PN. Performance of contemporary surgical risk scores for mitral valve surgery. Journal of cardiac surgery. 2017;32(3):172-6.

6.   Wang TK, Oh T, Voss J, Gamble G, Kang N, Pemberton J. Comparison of contemporary risk scores for predicting outcomes after surgery for active infective endocarditis. Heart and vessels. 2015;30(2):227-34.

7.   Habib G, Lancellotti P, Antunes MJ, Bongiorni MG, Casalta JP, Del Zotti F, et al. 2015 ESC Guidelines for the management of infective endocarditis: The Task Force for the Management of Infective Endocarditis of the European Society of Cardiology (ESC). Endorsed by: European Association for Cardio-Thoracic Surgery (EACTS), the European Association of Nuclear Medicine (EANM). European heart journal. 2015;36(44):3075-128.

8.   Pettersson GB, Coselli JS, Pettersson GB, Coselli JS, Hussain ST, Griffin B, et al. 2016 The American Association for Thoracic Surgery (AATS) consensus guidelines: Surgical treatment of infective endocarditis: Executive summary. The Journal of thoracic and cardiovascular surgery. 2017;153(6):1241-58.e29.

9.   Wang TKM, Wang MTM, Gamble GD, Webster M, Ruygrok PN. Performance of contemporary surgical risk scores for transcatheter aortic valve implantation: A meta-analysis. International journal of cardiology. 2017;236:350-5.

10.   Wang TK, Stewart RA, Ramanathan T, Kang N, Gamble G, White HD. Diagnosis of MI after CABG with high-sensitivity troponin T and new ECG or echocardiogram changes: relationship with mortality and validation of the universal definition of MI. European heart journal Acute cardiovascular care. 2013;2(4):323-33.

11.   Thygesen K, Alpert JS, Jaffe AS, Chaitman BR, Bax JJ, Morrow DA, et al. Fourth Universal Definition of Myocardial Infarction (2018). Journal of the American College of Cardiology. 2018.

12.   Wasson JH, Sox HC, Neff RK, Goldman L. Clinical prediction rules. Applications and methodological standards. The New England journal of medicine. 1985;313(13):793-9.

13.   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 (Cambridge, Mass). 2010;21(1):128-38.

14.   Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. The TRIPOD Group. Circulation. 2015;131(2):211-9.

15.   Royston P, Moons KG, Altman DG, Vergouwe Y. Prognosis and prognostic research: Developing a prognostic model. BMJ (Clinical research ed). 2009;338:b604.

16.   Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Statistics in medicine. 1996;15(4):361-87.

17.   Lip GY, Nieuwlaat R, Pisters R, Lane DA, Crijns HJ. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation. Chest. 2010;137(2):263-72.

18.   Wells PS, Anderson DR, Rodger M, Stiell I, Dreyer JF, Barnes D, et al. Excluding pulmonary embolism at the bedside without diagnostic imaging: management of patients with suspected pulmonary embolism presenting to the emergency department by using a simple clinical model and d-dimer. Annals of internal medicine. 2001;135(2):98-107.

19.   Steyerberg EW, Moons KG, van der Windt DA, Hayden JA, Perel P, Schroter S, et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. PLoS medicine. 2013;10(2):e1001381.

20.   Heagerty PJ, Zheng Y. Survival model predictive accuracy and ROC curves. Biometrics. 2005;61(1):92-105.

21.   Miller ME, Langefeld CD, Tierney WM, Hui SL, McDonald CJ. Validation of probabilistic predictions. Medical decision making : an international journal of the Society for Medical Decision Making. 1993;13(1):49-58.

22.   Hu B, Palta M, Shao J. Properties of R(2) statistics for logistic regression. Statistics in medicine. 2006;25(8):1383-95.

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

24.   Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Statistics in medicine. 2008;27(2):157-72; discussion 207-12.

25.   Roques F, Nashef SA, Michel P, Gauducheau E, de Vincentiis C, Baudet E, et al. Risk factors and outcome in European cardiac surgery: analysis of the EuroSCORE multinational database of 19030 patients. European journal of cardio-thoracic surgery : official journal of the European Association for Cardio-thoracic Surgery. 1999;15(6):816-22; discussion 22-3.

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

27.	Roques F, Michel P, Goldstone AR, Nashef SA. The logistic EuroSCORE. European heart journal. 2003;24(9):881-2.

28.	Parolari A, Pesce LL, Trezzi M, Cavallotti L, Kassem S, Loardi C, et al. EuroSCORE performance in valve surgery: a meta-analysis. The Annals of thoracic surgery. 2010;89(3):787-93, 93.e1-2.

29.	Siregar S, Groenwold RH, de Heer F, Bots ML, van der Graaf Y, van Herwerden LA. Performance of the original EuroSCORE. European journal of cardio-thoracic surgery : official journal of the European Association for Cardio-thoracic Surgery. 2012;41(4):746-54.

30.	Nashef SA, Roques F, Sharples LD, Nilsson J, Smith C, Goldstone AR, et al. EuroSCORE II. European journal of cardio-thoracic surgery : official journal of the European Association for Cardio-thoracic Surgery. 2012;41(4):734-44; discussion 44-5.

31.	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 Suppl):S2-22.

32.	O'Brien SM, Shahian DM, Filardo G, Ferraris VA, Haan CK, Rich JB, et al. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 2--isolated valve surgery. The Annals of thoracic surgery. 2009;88(1 Suppl):S23-42.

33.	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 3--valve plus coronary artery bypass grafting surgery. The Annals of thoracic surgery. 2009;88(1 Suppl):S43-62.

34.	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.

35.	Ariyaratne TV, Billah B, Yap CH, 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 : official journal of the European Association for Cardio-thoracic Surgery. 2011;39(6):815-21.

36.	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.

37.	Sullivan PG, Wallach JD, Ioannidis JP. Meta-Analysis Comparing Established Risk Prediction Models (EuroSCORE II, STS Score, and ACEF Score) for Perioperative Mortality During Cardiac Surgery. The American journal of cardiology. 2016;118(10):1574-82.

38.	Guida P, Mastro F, Scrascia G, Whitlock R, Paparella D. Performance of the European System for Cardiac Operative Risk Evaluation II: a meta-analysis of 22 studies involving 145,592 cardiac surgery procedures. The Journal of thoracic and cardiovascular surgery. 2014;148(6):3049-57.e1.

39.	Baumgartner H, Falk V, Bax JJ, De Bonis M, Hamm C, Holm PJ, et al. 2017 ESC/EACTS Guidelines for the management of valvular heart disease. European heart journal. 2017;38(36):2739-91.

40.	Edwards FH. Evolution of the Society of Thoracic Surgeons National Cardiac Surgery Database. The Journal of invasive cardiology. 1998;10(8):485-8.

41.   Yap CH, Reid C, Yii M, Rowland MA, Mohajeri M, Skillington PD, et al. Validation of the EuroSCORE model in Australia. European journal of cardio-thoracic surgery : official journal of the European Association for Cardio-thoracic Surgery. 2006;29(4):441-6; discussion 6.

42.   Wang TKM, Wang MTM, Pemberton J. Risk scores and surgery for infective endocarditis: A meta-analysis. International journal of cardiology. 2016;222:1001-2.

43.   Costa MA, Wollmann DR, Jr., Campos AC, Cunha CL, Carvalho RG, Andrade DF, et al. Risk index for death by infective endocarditis: a multivariate logistic model. Revista brasileira de cirurgia cardiovascular : orgao oficial da Sociedade Brasileira de Cirurgia Cardiovascular. 2007;22(2):192-200.

44.   Martinez-Selles M, Munoz P, Arnaiz A, Moreno M, Galvez J, Rodriguez-Roda J, et al. Valve surgery in active infective endocarditis: a simple score to predict in-hospital prognosis. International journal of cardiology. 2014;175(1):133-7.

45.   Gaca JG, Sheng S, Daneshmand MA, O'Brien S, Rankin JS, Brennan JM, et al. Outcomes for endocarditis surgery in North America: a simplified risk scoring system. The Journal of thoracic and cardiovascular surgery. 2011;141(1):98-106.e1-2.

46.   De Feo M, Cotrufo M, Carozza A, De Santo LS, Amendolara F, Giordano S, et al. The need for a specific risk prediction system in native valve infective endocarditis surgery. TheScientificWorldJournal. 2012;2012:307571.

47.   Olmos C, Vilacosta I, Habib G, Maroto L, Fernandez C, Lopez J, et al. Risk score for cardiac surgery in active left-sided infective endocarditis. Heart (British Cardiac Society). 2017;103(18):1435-42.

48.   Gatti G, Benussi B, Gripshi F, Della Mattia A, Proclemer A, Cannata A, et al. A risk factor analysis for in-hospital mortality after surgery for infective endocarditis and a proposal of a new predictive scoring system. Infection. 2017;45(4):413-23.

49.   Nishimura RA, Otto CM, Bonow RO, Carabello BA, Erwin JP, 3rd, Fleisher LA, et al. 2017 AHA/ACC Focused Update of the 2014 AHA/ACC Guideline for the Management of Patients With Valvular Heart Disease: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Journal of the American College of Cardiology. 2017;70(2):252-89.

50.   Leon MB, Smith CR, Mack M, Miller DC, Moses JW, Svensson LG, et al. Transcatheter aortic-valve implantation for aortic stenosis in patients who cannot undergo surgery. The New England journal of medicine. 2010;363(17):1597-607.

51.   Smith CR, Leon MB, Mack MJ, Miller DC, Moses JW, Svensson LG, et al. Transcatheter versus surgical aortic-valve replacement in high-risk patients. The New England journal of medicine. 2011;364(23):2187-98.

52.   Adams DH, Popma JJ, Reardon MJ. Transcatheter aortic-valve replacement with a self-expanding prosthesis. The New England journal of medicine. 2014;371(10):967-8.

53.   Leon MB, Smith CR, Mack MJ, Makkar RR, Svensson LG, Kodali SK, et al. Transcatheter or Surgical Aortic-Valve Replacement in Intermediate-Risk Patients. The New England journal of medicine. 2016;374(17):1609-20.

54.   Reardon MJ, Van Mieghem NM, Popma JJ, Kleiman NS, Sondergaard L, Mumtaz M, et al. Surgical or Transcatheter Aortic-Valve Replacement in Intermediate-Risk Patients. The New England journal of medicine. 2017;376(14):1321-31.

55.   Capodanno D, Barbanti M, Tamburino C, D'Errigo P, Ranucci M, Santoro G, et al. A simple risk tool (the OBSERVANT score) for prediction of 30-day mortality after transcatheter aortic valve replacement. The American journal of cardiology. 2014;113(11):1851-8.

56.   D'Ascenzo F, Capodanno D, Tarantini G, Nijhoff F, Ciuca C, Rossi ML, et al. Usefulness and validation of the survival posT TAVI score for survival after transcatheter aortic valve implantation for aortic stenosis. The American journal of cardiology. 2014;114(12):1867-74.

57.   Iung B, Laouenan C, Himbert D, Eltchaninoff H, Chevreul K, Donzeau-Gouge P, et al. Predictive factors of early mortality after transcatheter aortic valve implantation: individual risk assessment using a simple score. Heart (British Cardiac Society). 2014;100(13):1016-23.

58.   Debonnaire P, Fusini L, Wolterbeek R, Kamperidis V, van Rosendael P, van der Kley F, et al. Value of the "TAVI2-SCORe" versus surgical risk scores for prediction of one year mortality in 511 patients who underwent transcatheter aortic valve implantation. The American journal of cardiology. 2015;115(2):234-42.

59.   Hermiller JB, Jr., Yakubov SJ, Reardon MJ, Deeb GM, Adams DH, Afilalo J, et al. Predicting Early and Late Mortality After Transcatheter Aortic Valve Replacement. Journal of the American College of Cardiology. 2016;68(4):343-52.

60.   Edwards FH, Cohen DJ, O'Brien SM, Peterson ED, Mack MJ, Shahian DM, et al. Development and Validation of a Risk Prediction Model for In-Hospital Mortality After Transcatheter Aortic Valve Replacement. JAMA cardiology. 2016;1(1):46-52.

61.   Windecker S, Kolh P, Alfonso F, Collet JP, Cremer J, Falk V, et al. 2014 ESC/EACTS Guidelines on myocardial revascularization: The Task Force on Myocardial Revascularization of the European Society of Cardiology (ESC) and the European Association for Cardio-Thoracic Surgery (EACTS)Developed with the special contribution of the European Association of Percutaneous Cardiovascular Interventions (EAPCI). European heart journal. 2014;35(37):2541-619.

62.   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 Suppl I:I-3-I-12.

63.   Gogbashian A, Sedrakyan A, Treasure T. EuroSCORE: a systematic review of international performance. European journal of cardio-thoracic surgery : official journal of the European Association for Cardio-thoracic Surgery. 2004;25(5):695-700.

64.   Nashef SA, Roques F, Hammill BG, Peterson ED, Michel P, Grover FL, et al. Validation of European System for Cardiac Operative Risk Evaluation (EuroSCORE) in North American cardiac surgery. European journal of cardio-thoracic surgery : official journal of the European Association for Cardio-thoracic Surgery. 2002;22(1):101-5.

65.   Bhatti F, Grayson AD, Grotte G, Fabri BM, Au J, Jones M, et al. The logistic EuroSCORE in cardiac surgery: how well does it predict operative risk? Heart (British Cardiac Society). 2006;92(12):1817-20.

66.   Biancari F, Vasques F, Mikkola R, Martin M, Lahtinen J, Heikkinen J. Validation of EuroSCORE II in patients undergoing coronary artery bypass surgery. The Annals of thoracic surgery. 2012;93(6):1930-5.

67.   Chalmers J, Pullan M, Fabri B, McShane J, Shaw M, Mediratta N, et al. Validation of EuroSCORE II in a modern cohort of patients undergoing cardiac surgery. European journal of cardio-thoracic surgery : official journal of the European Association for Cardio-thoracic Surgery. 2013;43(4):688-94.

68.   Di Dedda U, Pelissero G, Agnelli B, De Vincentiis C, Castelvecchio S, Ranucci M. Accuracy, calibration and clinical performance of the new EuroSCORE II risk stratification system. European journal of cardio-thoracic surgery : official journal of the European Association for Cardio-thoracic Surgery. 2013;43(1):27-32.

69.   Grant SW, Hickey GL, Dimarakis I, Trivedi U, Bryan A, Treasure T, et al. How does EuroSCORE II perform in UK cardiac surgery; an analysis of 23 740 patients from the Society for Cardiothoracic Surgery in Great Britain and Ireland National Database. Heart (British Cardiac Society). 2012;98(21):1568-72.

70.   Carnero-Alcazar M, Silva Guisasola JA, Reguillo Lacruz FJ, Maroto Castellanos LC, Cobiella Carnicer J, Villagran Medinilla E, et al. Validation of EuroSCORE II on a single-centre 3800 patient cohort. Interactive cardiovascular and thoracic surgery. 2013;16(3):293-300.

71.   Barili F, Pacini D, Capo A, Rasovic O, Grossi C, Alamanni F, et al. Does EuroScore II perform better than its original versions? A multicentre validation study. European heart journal. 2013;34:22-9.

72.   Thygesen K, Alpert JS, Jaffe AS, Simoons ML, Chaitman BR, White HD, et al. Third Universal Definition of Myocardial Infarction. Circulation. 2012;126(16):2020-35.

73.   Castellanos LR, Li Z, Yeo KK, Young JN, Ayanian JZ, Amsterdam EA. Relation of race, ethnicity and cardiac surgeons to operative mortality rates in primary coronary artery bypass grafting in California. The American journal of cardiology. 2011;107(1):1-5.

74.   Biancari F, Kangasniemi OP, Luukkonen J, Vuorisalo S, Satta J, Pokela R, et al. EuroSCORE predicts immediate and late outcome after coronary artery bypass surgery. The Annals of thoracic surgery. 2006;82(1):57-61.

75.   De Maria R, Mazzoni M, Parolini M, Gregori D, Bortone F, Arena V, et al. Predictive value of EuroSCORE on long term outcome in cardiac surgery patients: a single institution study. Heart (British Cardiac Society). 2005;91(6):779-84.

76.   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. European journal of cardio-thoracic surgery : official journal of the European Association for Cardio-thoracic Surgery. 2012;42(1):101-6; discussion 6-7.

77.   Toumpoulis IK, Anagnostopoulos CE, DeRose JJ, Swistel DG. Does EuroSCORE predict length of stay and specific postoperative complications after coronary artery bypass grafting? International journal of cardiology. 2005;105(1):19-25.

78.   Vahanian A, Alfieri O, Andreotti F, Antunes MJ, Baron-Esquivias G, Baumgartner H, et al. Guidelines on the management of valvular heart disease (version 2012). European heart journal. 2012;33(19):2451-96.

79.   Varadarajan P, Kapoor N, Bansal RC, Pai RG. Survival in elderly patients with severe aortic stenosis is dramatically improved by aortic valve replacement: Results from a cohort of 277 patients aged > or =80 years. European journal of cardio-thoracic surgery : official journal of the European Association for Cardio-thoracic Surgery. 2006;30(5):722-7.

80. Nkomo VT, Gardin JM, Skelton TN, Gottdiener JS, Scott CG, Enriquez-Sarano M. Burden of valvular heart diseases: a population-based study. Lancet (London, England). 2006;368(9540):1005-11.

81. Dewey TM, Brown D, Ryan WH, Herbert MA, Prince SL, Mack MJ. Reliability of risk algorithms in predicting early and late operative outcomes in high-risk patients undergoing aortic valve replacement. The Journal of thoracic and cardiovascular surgery. 2008;135(1):180-7.

82. Wendt D, Osswald BR, Kayser K, Thielmann M, Tossios P, Massoudy P, et al. Society of Thoracic Surgeons score is superior to the EuroSCORE determining mortality in high risk patients undergoing isolated aortic valve replacement. The Annals of thoracic surgery. 2009;88(2):468-74; discussion 74-5.

83. Basraon J, Chandrashekhar YS, John R, Agnihotri A, Kelly R, Ward H, et al. Comparison of risk scores to estimate perioperative mortality in aortic valve replacement surgery. The Annals of thoracic surgery. 2011;92(2):535-40.

84. Barili F, Pacini D, Capo A, Ardemagni E, Pellicciari G, Zanobini M, et al. Reliability of new scores in predicting perioperative mortality after isolated aortic valve surgery: a comparison with the society of thoracic surgeons score and logistic EuroSCORE. The Annals of thoracic surgery. 2013;95(5):1539-44.

85. Kappetein AP, Head SJ, Genereux P, Piazza N, van Mieghem NM, Blackstone EH, et al. Updated standardized endpoint definitions for transcatheter aortic valve implantation: the Valve Academic Research Consortium-2 consensus document. European heart journal. 2012;33(19):2403-18.

86. Puskas JD, Kilgo PD, Thourani VH, Lattouf OM, Chen E, Vega JD, et al. The society of thoracic surgeons 30-day predicted risk of mortality score also predicts long-term survival. The Annals of thoracic surgery. 2012;93(1):26-33; discussion -5.

87. Lagercrantz E, Lindblom D, Sartipy U. Survival and quality of life in cardiac surgery patients with prolonged intensive care. The Annals of thoracic surgery. 2010;89(2):490-5.

88. Ettema RG, Peelen LM, Schuurmans MJ, Nierich AP, Kalkman CJ, Moons KG. Prediction models for prolonged intensive care unit stay after cardiac surgery: systematic review and validation study. Circulation. 2010;122(7):682-9, 7 p following p 9.

89. Widyastuti Y, Stenseth R, Wahba A, Pleym H, Videm V. Length of intensive care unit stay following cardiac surgery: is it impossible to find a universal prediction model? Interactive cardiovascular and thoracic surgery. 2012;15(5):825-32.

90. Widyastuti Y, Stenseth R, Pleym H, Wahba A, Videm V. Pre-operative and intraoperative determinants for prolonged ventilation following adult cardiac surgery. Acta anaesthesiologica Scandinavica. 2012;56(2):190-9.

91. Hirose H, Inaba H, Noguchi C, Tambara K, Yamamoto T, Yamasaki M, et al. EuroSCORE predicts postoperative mortality, certain morbidities, and recovery time. Interactive cardiovascular and thoracic surgery. 2009;9(4):613-7.

92. Bucerius J, Gummert JF, Borger MA, Walther T, Doll N, Onnasch JF, et al. Stroke after cardiac surgery: a risk factor analysis of 16,184 consecutive adult patients. The Annals of thoracic surgery. 2003;75(2):472-8.

93. Antunes PE, de Oliveira JF, Antunes MJ. Predictors of cerebrovascular events in patients subjected to isolated coronary surgery. The importance of aortic cross-clamping. European journal of

cardio-thoracic surgery : official journal of the European Association for Cardio-thoracic Surgery. 2003;23(3):328-33.

94.   Wang L, Han QQ, Qiao F, Wang C, Zhang XW, Han L, et al. Performance of EuroSCORE II in patients who have undergone heart valve surgery: a multicentre study in a Chinese population. European journal of cardio-thoracic surgery : official journal of the European Association for Cardio-thoracic Surgery. 2014;45(2):359-64.

95.   Doddakula K, Al-Sarraf N, Gately K, Hughes A, Tolan M, Young V, et al. Predictors of acute renal failure requiring renal replacement therapy post cardiac surgery in patients with preoperatively normal renal function. Interactive cardiovascular and thoracic surgery. 2007;6(3):314-8.

96.   Gummert JF, Bucerius J, Walther T, Doll N, Falk V, Schmitt DV, et al. Requirement for renal replacement therapy in patients undergoing cardiac surgery. The Thoracic and cardiovascular surgeon. 2004;52(2):70-6.

97.   Fowler VG, Jr., O'Brien SM, Muhlbaier LH, Corey GR, Ferguson TB, Peterson ED. Clinical predictors of major infections after cardiac surgery. Circulation. 2005;112(9 Suppl):I358-65.

98.   Paul M, Raz A, Leibovici L, Madar H, Holinger R, Rubinovitch B. Sternal wound infection after coronary artery bypass graft surgery: validation of existing risk scores. The Journal of thoracic and cardiovascular surgery. 2007;133(2):397-403.

99.   Tiveron MG, Fiorelli AI, Mota EM, Mejia OA, Brandao CM, Dallan LA, et al. Preoperative risk factors for mediastinitis after cardiac surgery: analysis of 2768 patients. Revista brasileira de cirurgia cardiovascular : orgao oficial da Sociedade Brasileira de Cirurgia Cardiovascular. 2012;27(2):203-10.

100.   Alsoufi B, Karamlou T, Slater M, Shen I, Ungerleider R, Ravichandran P. Results of concomitant aortic valve replacement and coronary artery bypass grafting in the VA population. The Journal of heart valve disease. 2006;15(1):12-8; discussion 8-9.

101.   Roberts WC, Roberts CC, Vowels TJ, Ko JM, Filardo G, Hamman BL, et al. Effect of coronary bypass and valve structure on outcome in isolated valve replacement for aortic stenosis. The American journal of cardiology. 2012;109(9):1334-40.

102.   Beach JM, Mihaljevic T, Svensson LG, Rajeswaran J, Marwick T, Griffin B, et al. Coronary artery disease and outcomes of aortic valve replacement for severe aortic stenosis. Journal of the American College of Cardiology. 2013;61(8):837-48.

103.   Koene BM, van Straten AH, van Geldorp MW, Ter Woorst JF, Elenbaas TW, Soliman Hamad MA. Can the EuroSCORE predict midterm survival after aortic valve replacement? Journal of cardiothoracic and vascular anesthesia. 2012;26(4):617-23.

104.   Feldman T, Foster E, Glower DD, Kar S, Rinaldi MJ, Fail PS, et al. Percutaneous repair or surgery for mitral regurgitation. The New England journal of medicine. 2011;364(15):1395-406.

105.   Nishimura RA, Vahanian A, Eleid MF, Mack MJ. Mitral valve disease--current management and future challenges. Lancet (London, England). 2016;387(10025):1324-34.

106.   Chan V, Ahrari A, Ruel M, Elmistekawy E, Hynes M, Mesana TG. Perioperative deaths after mitral valve operations may be overestimated by contemporary risk models. The Annals of thoracic surgery. 2014;98(2):605-10; discussion 10.

107.   Adamo M, Capodanno D, Cannata S, Giannini C, Laudisa ML, Barbanti M, et al. Comparison of three contemporary surgical scores for predicting all-cause mortality of patients undergoing percutaneous mitral valve repair with the MitraClip system (from the multicenter GRASP-IT registry). The American journal of cardiology. 2015;115(1):107-12.

108.   Moreillon P, Que YA. Infective endocarditis. Lancet (London, England). 2004;363(9403):139-49.

109.   Bonow RO, Carabello BA, Chatterjee K, de Leon AC, Jr., Faxon DP, Freed MD, et al. ACC/AHA 2006 guidelines for the management of patients with valvular heart disease: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (writing Committee to Revise the 1998 guidelines for the management of patients with valvular heart disease) developed in collaboration with the Society of Cardiovascular Anesthesiologists endorsed by the Society for Cardiovascular Angiography and Interventions and the Society of Thoracic Surgeons. Journal of the American College of Cardiology. 2006;48(3):e1-148.

110.   Habib G, Hoen B, Tornos P, Thuny F, Prendergast B, Vilacosta I, et al. Guidelines on the prevention, diagnosis, and treatment of infective endocarditis (new version 2009): the Task Force on the Prevention, Diagnosis, and Treatment of Infective Endocarditis of the European Society of Cardiology (ESC). Endorsed by the European Society of Clinical Microbiology and Infectious Diseases (ESCMID) and the International Society of Chemotherapy (ISC) for Infection and Cancer. European heart journal. 2009;30(19):2369-413.

111.   Mestres CA, Castro MA, Bernabeu E, Josa M, Cartana R, Pomar JL, et al. Preoperative risk stratification in infective endocarditis. Does the EuroSCORE model work? Preliminary results. European journal of cardio-thoracic surgery : official journal of the European Association for Cardio-thoracic Surgery. 2007;32(2):281-5.

112.   Rasmussen RV, Bruun LE, Lund J, Larsen CT, Hassager C, Bruun NE. The impact of cardiac surgery in native valve infective endocarditis: can euroSCORE guide patient selection? International journal of cardiology. 2011;149(3):304-9.

113.   Tleyjeh IM, Ghomrawi HM, Steckelberg JM, Hoskin TL, Mirzoyev Z, Anavekar NS, et al. The impact of valve surgery on 6-month mortality in left-sided infective endocarditis. Circulation. 2007;115(13):1721-8.

114.   Lalani T, Cabell CH, Benjamin DK, Lasca O, Naber C, Fowler VG, Jr., et al. Analysis of the impact of early surgery on in-hospital mortality of native valve endocarditis: use of propensity score and instrumental variable methods to adjust for treatment-selection bias. Circulation. 2010;121(8):1005-13.

115.   Kiefer T, Park L, Tribouilloy C, Cortes C, Casillo R, Chu V, et al. Association between valvular surgery and mortality among patients with infective endocarditis complicated by heart failure. Jama. 2011;306(20):2239-47.

116.   Leontyev S, Borger MA, Modi P, Lehmann S, Seeburger J, Doenst T, et al. Surgical management of aortic root abscess: a 13-year experience in 172 patients with 100% follow-up. The Journal of thoracic and cardiovascular surgery. 2012;143(2):332-7.

117.   . !!! INVALID CITATION !!! {}.

118.   Wilbring M, Tugtekin SM, Alexiou K, Matschke K, Kappert U. Composite aortic root replacement for complex prosthetic valve endocarditis: initial clinical results and long-term follow-up of high-risk patients. The Annals of thoracic surgery. 2012;94(6):1967-74.

119.   Barsic B, Dickerman S, Krajinovic V, Pappas P, Altclas J, Carosi G, et al. Influence of the timing of cardiac surgery on the outcome of patients with infective endocarditis and stroke. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2013;56(2):209-17.

120.   Meszaros K, Nujic S, Sodeck GH, Englberger L, Konig T, Schonhoff F, et al. Long-term results after operations for active infective endocarditis in native and prosthetic valves. The Annals of thoracic surgery. 2012;94(4):1204-10.

121.   Tsang VT, Brown KL, Synnergren MJ, Kang N, de Leval MR, Gallivan S, et al. Monitoring risk-adjusted outcomes in congenital heart surgery: does the appropriateness of a risk model change with time? The Annals of thoracic surgery. 2009;87(2):584-7.

122.   Madeira S, Rodrigues R, Tralhao A, Santos M, Almeida C, Marques M, et al. Assessment of perioperative mortality risk in patients with infective endocarditis undergoing cardiac surgery: performance of the EuroSCORE I and II logistic models. Interactive cardiovascular and thoracic surgery. 2016;22(2):141-8.

123.   Patrat-Delon S, Rouxel A, Gacouin A, Revest M, Flecher E, Fouquet O, et al. EuroSCORE II underestimates mortality after cardiac surgery for infective endocarditis. European journal of cardio-thoracic surgery : official journal of the European Association for Cardio-thoracic Surgery. 2016;49(3):944-51.

124.   Thourani VH, Kodali S, Makkar RR, Herrmann HC, Williams M, Babaliaros V, et al. Transcatheter aortic valve replacement versus surgical valve replacement in intermediate-risk patients: a propensity score analysis. Lancet (London, England). 2016;387(10034):2218-25.

125.   Arai T, Lefevre T, Hayashida K, Watanabe Y, O'Connor SA, Hovasse T, et al. Usefulness of a Simple Clinical Risk Prediction Method, Modified ACEF Score, for Transcatheter Aortic Valve Implantation. Circulation journal : official journal of the Japanese Circulation Society. 2015;79(7):1496-503.

126.   Balan P, Zhao Y, Johnson S, Arain S, Dhoble A, Estrera A, et al. The Society of Thoracic Surgery Risk Score as a Predictor of 30-Day Mortality in Transcatheter vs Surgical Aortic Valve Replacement: A Single-Center Experience and its Implications for the Development of a TAVR Risk-Prediction Model. The Journal of invasive cardiology. 2017;29(3):109-14.

127.   Ben-Dor I, Gaglia MA, Jr., Barbash IM, Maluenda G, Hauville C, Gonzalez MA, et al. Comparison between Society of Thoracic Surgeons score and logistic EuroSCORE for predicting mortality in patients referred for transcatheter aortic valve implantation. Cardiovascular revascularization medicine : including molecular interventions. 2011;12(6):345-9.

128.   Beohar N, Whisenant B, Kirtane AJ, Leon MB, Tuzcu EM, Makkar R, et al. The relative performance characteristics of the logistic European System for Cardiac Operative Risk Evaluation score and the Society of Thoracic Surgeons score in the Placement of Aortic Transcatheter Valves trial. The Journal of thoracic and cardiovascular surgery. 2014;148(6):2830-7.e1.

129.   Buchanan G, Chieffo A, Montorfano M, Latib A, Maisano F, Cioni M, et al. Does the Logistic EuroSCORE II Better Predict Outcomes Following Transcatheter Aortic Valve Implantation? Results from the Milan Registry. Journal of the American College of Cardiology. 2012;60(17 Supplement):B249.

130.   Conradi L, Silaschi M, Schnabel R, Seiffert M, Schön G, Diemert P, et al. Predictive reliability of logistic EuroSCORE II in patients undergoing transcatheter aortic valve implantation: assessment and comparison to classic systems of preoperative risk stratification. Journal of the American College of Cardiology. 2012;60(17 Supplement):B246.

131. D'Onofrio A, Gasparetto V, Rizzoli G, Napodano M, Tarantini G, Tessari C, et al. What is the best risk stratification model for patients undergoing transcatheter aortic valve implantation? . European heart journal. 2012;33:585-6.

132. Durand E, Borz B, Godin M, Tron C, Litzler PY, Bessou JP, et al. Performance analysis of EuroSCORE II compared to the original logistic EuroSCORE and STS scores for predicting 30-day mortality after transcatheter aortic valve replacement. The American journal of cardiology. 2013;111(6):891-7.

133. Haensig M, Holzhey DM, Borger MA, Schuler G, Shi W, Subramanian S, et al. Is the new EuroSCORE II a better predictor for transapical aortic valve implantation? European journal of cardio-thoracic surgery : official journal of the European Association for Cardio-thoracic Surgery. 2013;44(2):302-8; discussion 8.

134. Johansson M, Nozohoor S, Zindovic I, Nilsson J, Kimblad PO, Sjogren J. Prediction of 30-day mortality after transcatheter aortic valve implantation: a comparison of logistic EuroSCORE, STS score, and EuroSCORE II. The Journal of heart valve disease. 2014;23(5):567-74.

135. Piazza N, Wenaweser P, van Gameren M, Pilgrim T, Tzikas A, Otten A, et al. Relationship between the logistic EuroSCORE and the Society of Thoracic Surgeons Predicted Risk of Mortality score in patients implanted with the CoreValve ReValving system--a Bern-Rotterdam Study. American heart journal. 2010;159(2):323-9.

136. Sedaghat A, Sinning JM, Vasa-Nicotera M, Ghanem A, Hammerstingl C, Grube E, et al. The revised EuroSCORE II for the prediction of mortality in patients undergoing transcatheter aortic valve implantation. Clinical research in cardiology : official journal of the German Cardiac Society. 2013;102(11):821-9.

137. Silaschi M, Conradi L, Seiffert M, Schnabel R, Schon G, Blankenberg S, et al. Predicting Risk in Transcatheter Aortic Valve Implantation: Comparative Analysis of EuroSCORE II and Established Risk Stratification Tools. The Thoracic and cardiovascular surgeon. 2015;63(6):472-8.

138. Sinning J, Sedaghat A, Ghanem A, Hammerstingl C, Weber M, Schueler R, et al. The value of the German aortic valve score for the prediction of peri-and post-procedural mortality in TAVI patients. EuroIntervention : journal of EuroPCR in collaboration with the Working Group on Interventional Cardiology of the European Society of Cardiology. 2014;10:1774.

139. Sinnot Silva L, Bacelar Nunes Filho AC, Katz M, Abizaid A, Lima V, Carvalho LA, et al. Mortality prediction using five different risk scores for cardiac surgery in patients undergoing transcatheter aortic valve implantation (TAVI): A multicenter Brazilian registry. European heart journal. 2014;35:249-50.

140. Sirotina M, Weber J, Fox H, Hemmann K, Hofmann I, Moritz A, et al. Utility of conventional surgical risk scores in predicting outcome after transcatheter aortic valve replacement. Journal of the American College of Cardiology. 2013;61(10 Supplement):E1949.

141. Stahli BE, Tasnady H, Luscher TF, Gebhard C, Mikulicic F, Erhart L, et al. Early and late mortality in patients undergoing transcatheter aortic valve implantation: comparison of the novel EuroScore II with established risk scores. Cardiology. 2013;126(1):15-23.

142. Watanabe Y, Hayashida K, Lefevre T, Chevalier B, Hovasse T, Romano M, et al. Is EuroSCORE II better than EuroSCORE in predicting mortality after transcatheter aortic valve implantation? Catheterization and cardiovascular interventions : official journal of the Society for Cardiac Angiography & Interventions. 2013;81(6):1053-60.

143.   Wendt D, Thielmann M, Kahlert P, Kastner S, Price V, Al-Rashid F, et al. Comparison between different risk scoring algorithms on isolated conventional or transcatheter aortic valve replacement. The Annals of thoracic surgery. 2014;97(3):796-802.

144.   Zbronski K, Huczek Z, Puchta D, Paczwa K, Kochman J, Wilimski R, et al. Outcome prediction following transcatheter aortic valve implantation: Multiple risk scores comparison. Cardiology journal. 2016;23(2):169-77.

145.   Biancari F, Juvonen T, Onorati F, Faggian G, Heikkinen J, Airaksinen J, et al. Meta-analysis on the performance of the EuroSCORE II and the Society of Thoracic Surgeons Scores in patients undergoing aortic valve replacement. Journal of cardiothoracic and vascular anesthesia. 2014;28(6):1533-9.

146.   Hashmi IH, Hammad S, Rajagopal R, Croft D, More R, Rogers S, et al. Is SURTAVI risk model a step towards 'TAVI risk score': time to rethink? Heart (British Cardiac Society). 2013;99(suppl 2):A86-A7.

147.   Januzzi JL, Lewandrowski K, MacGillivray TE, Newell JB, Kathiresan S, Servoss SJ, et al. A comparison of cardiac troponin T and creatine kinase-MB for patient evaluation after cardiac surgery. JAmCollCardiol. 2002;39(9):1518-23.

148.   Muehlschlegel JD, Perry TE, Liu KY, Nascimben L, Fox AA, Collard CD, et al. Troponin is superior to electrocardiogram and creatinine kinase MB for predicting clinically significant myocardial injury after coronary artery bypass grafting. European heart journal. 2009;30(13):1574-83.

149.   Croal BL, Hillis GS, Gibson PH, Fazal MT, El-Shafei H, Gibson G, et al. Relationship between postoperative cardiac troponin I levels and outcome of cardiac surgery. Circulation. 2006;114(14):1468-75.

150.   Martin CB, Shaw AD, Gal J, Aravindan N, Murphy F, Royston D, et al. The comparison and validity of troponin I assay systems in diagnosing myocardial ischemic injury after surgical coronary revascularization. Journal of cardiothoracic and vascular anesthesia. 2005;19(3):288-93.

151.   Pegg TJ, Maunsell Z, Karamitsos TD, Taylor RP, James T, Francis JM, et al. Utility of cardiac biomarkers for the diagnosis of type V myocardial infarction after coronary artery bypass grafting: insights from serial cardiac MRI. Heart (British Cardiac Society). 2011;97(10):810-6.

152.   Mohammed AA, Agnihotri AK, van Kimmenade RR, Martinez-Rumayor A, Green SM, Quiroz R, et al. Prospective, comprehensive assessment of cardiac troponin T testing after coronary artery bypass graft surgery. Circulation. 2009;120(10):843-50.

153.   Apple FS, Collinson PO. Analytical characteristics of high-sensitivity cardiac troponin assays. Clin Chem. 2012;58(1):54-61.

154.   Giannitsis E, Kurz K, Hallermayer K, Jarausch J, Jaffe AS, Katus HA. Analytical validation of a high-sensitivity cardiac troponin T assay. ClinChem. 2010;56(2):254-61.

155.   Domanski MJ, Mahaffey K, Hasselblad V, Brener SJ, Smith PK, Hillis G, et al. Association of myocardial enzyme elevation and survival following coronary artery bypass graft surgery. Jama. 2011;305(6):585-91.

156.   Apple FS. A new season for cardiac troponin assays: it's time to keep a scorecard. ClinChem. 2009;55(7):1303-6.

157. Carrier M, Pellerin M, Perrault L, Solymoss BC, Pelletier LC. Troponin levels in patients with myocardial infarction after coronary artery bypass grafting. AnnThoracSurg. 2000;69:435-40.

158. Lim CC, Cuculi F, van Gaal WJ, Testa L, Arnold JR, Karamitsos T, et al. Early diagnosis of perioperative myocardial infarction after coronary bypass grafting: a study using biomarkers and cardiac magnetic resonance imaging. The Annals of thoracic surgery. 2011;92(6):2046-53.

159. Thygesen K, Alpert JS, White HD, on behalf of the Joint ESC-ACC-AHA-WHF Task Force for the Redefinition of Myocardial Infarction. Universal definition of myocardial infarction. EurHeart J. 2007;28(20):2525-38.

160. Miller WL, Garratt KN, Burritt MF, Lennon RJ, Reeder GS, Jaffe AS. Baseline troponin level: key to understanding the importance of post-PCI troponin elevations. European heart journal. 2006;27(9):1061-9.

161. Onorati F, De Feo M, Mastroroberto P, Cristodoro L, Pezzo F, Renzulli A, et al. Determinants and prognosis of myocardial damage after coronary artery bypass grafting. The Annals of thoracic surgery. 2005;79(3):837-45.

162. Cameron A, Davis KB, Green G, Schaff HV. Coronary bypass surgery with internal-thoracic-artery grafts - effects on survival over a 15-year period. NEnglJMed. 1996;334:216-9.

163. Antiplatelet Trialists' Collaboration. Collaborative overview of randomised trials of antiplatelet therapy - II: Maintenance of vascular graft or arterial patency by antiplatelet therapy. British Medical Journal. 1994;308:159-68.

164. Liakopoulos OJ, Choi YH, Haldenwang PL, Strauch J, Wittwer T, Dorge H, et al. Impact of preoperative statin therapy on adverse postoperative outcomes in patients undergoing cardiac surgery: a meta-analysis of over 30,000 patients. European heart journal. 2008;29(12):1548-59.

165. Ferguson TB, Jr., Coombs LP, Peterson ED. Preoperative beta-blocker use and mortality and morbidity following CABG surgery in North America. Jama. 2002;287(17):2221-7.

166. ElBardissi AW, Aranki SF, Sheng S, O'Brien SM, Greenberg CC, Gammie JS. Trends in isolated coronary artery bypass grafting: an analysis of the Society of Thoracic Surgeons adult cardiac surgery database. The Journal of thoracic and cardiovascular surgery. 2012;143(2):273-81.

167. Pylypchuk R, Wells S, Kerr A, Poppe K, Riddell T, Harwood M, et al. Cardiovascular disease risk prediction equations in 400 000 primary care patients in New Zealand: a derivation and validation study. Lancet (London, England). 2018;391(10133):1897-907.

168. Kerr A, Williams MJ, White H, Doughty R, Nunn C, Devlin G, et al. The All New Zealand Acute Coronary Syndrome Quality Improvement Programme: Implementation, Methodology and Cohorts (ANZACS-QI 9). N Z Med J. 2016;129(1439):23-36.