

Risk stratification in cardiac surgery

Neethu Krishna¹ · Praveen Kerala Varma¹

Received: 13 April 2015 / Revised: 27 June 2015 / Accepted: 30 June 2015 / Published online: 12 August 2015
© Indian Association of Cardiovascular-Thoracic Surgeons 2015

Abstract Preoperative risk scores are essential tools for risk assessment, cost–benefit analysis, and the study of therapy trends. Various scoring systems have been developed to predict mortality and morbidity after cardiac surgery. Risk stratification will inform patients and clinicians of the likely risk of death for a group of patients with a similar risk profile undergoing the proposed operation. This information is useful and should form part of the basis on which the patient and surgeon decide whether to proceed or not. Risk models were also applied for quality improvement programs comparing year-to-year outcomes, as well as allocation of healthcare resources through the prediction of length of stay and postoperative complication rates. In this review, we focus on the various risk score models used in adult and pediatric cardiac surgery.

Keywords Cardiac surgery · Risk Stratification · Risk Scores

Introduction

Preoperative risk scores are essential tools for risk assessment, cost–benefit analysis, and the study of therapy trends. Various scoring systems have been developed to predict mortality after adult heart surgery. Although all of these scoring systems are based on patient-derived data, such as age, gender, comorbidity, and so forth, there are considerable differences between scores with regard to their design and validity. As quality

control and cost–benefit analysis have gained new relevance in many countries, selection of appropriate scoring system for the evaluation of hospital performance has become an important issue [1]. Most of the scoring systems were primarily designed to predict mortality; however, postoperative morbidity has been acknowledged as the major determinant of hospital cost and quality of life after surgery [2].

One of the original aims for the development of cardiac risk models was risk adjustment, allowing fair comparison of treatment outcomes among different institutions or surgeons [3]. Risk stratification will inform patients and clinicians of the likely risk of death for a group of patients with a similar risk profile undergoing the proposed operation. This information is useful and should form part of the basis on which the patient and surgeon decide whether to proceed or not [4]. Risk models were also applied for quality improvement programs comparing year-to-year outcomes, as well as allocation of healthcare resources through the prediction of length of stay and postoperative complication rates [5].

The first widely used risk model, the Parsonnet score, was based on a retrospective analysis of data collected during the 1980s. Risk modelling has been significantly influenced by advances made in diagnostic and interventional technology. The advances in interventional cardiology have adversely changed the risk profile of patients presenting for cardiac surgery [6]. This resulted in development of newer risk scores and revalidation of existing risk scores.

Mortality and morbidity are the outcomes that are usually measured in any risk model. Mortality is the most widely used outcome measure as it is the least subjective of the outcome variables. Multivariate analysis is the cornerstone for assessing the outcome. The statistical technique commonly used for multivariate analysis is called regression analysis. Regression analysis builds a model based on dependency of

✉ Praveen Kerala Varma
varmapk@gmail.com

¹ Division of Cardiac Surgery, Amrita Institute of Medical Sciences, Kochi, Kerala 682 041, India

the outcome on a set of predictor variables otherwise called risk factors.

In this review, we aim to describe the various risk stratification scores used in adult and pediatric cardiac surgery.

Validation and discrimination of risk models

Performance of a risk model is evaluated by calibration and discrimination. The model is tested on the validation data set for calibration (by comparing the observed and predicted mortality or goodness of fit) and for discrimination [using the area under the receiver operating characteristic (ROC) curve]. Goodness of fit of the final model is tested using the Hosmer–Lemeshow statistic. The Hosmer–Lemeshow (HL) chi-square statistic measures the differences between the expected and observed outcomes over deciles of risk. If a model is well-calibrated, the *O/E* ratio should be close to 1; departures above or below are indicative of under-prediction and over-prediction, respectively. A well-calibrated model gives corresponding $P > 0.05$. Discrimination means how well the model differentiates a population that had an event from the one that did not. ROC curves are typically plotted to evaluate the performance of logistic regression models. In statistics, a ROC, or ROC curve, is a graphical plot that illustrates the performance of a binary classifier system as its discrimination threshold is varied. The curve is created by plotting the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings [7]. ROC curves were initially used in the Second World War by the US navy in radar-based operations to discriminate enemy ships or plane from friendly ones (hence the name) and since then have found extensive applications in industry and medicine. The curve typically plots sensitivity in the *Y* axis and 1 – specificity in the *X* axis from a dichotomous outcome. *C* statistic is equal to the area under the curve. A perfect discrimination gives a score of 1, and 0.5 denotes no discrimination (like toss of a coin for heads or tails). A value upward of 0.75 has good discriminatory ability (Fig. 1).

Risk assessment scores in adult cardiac surgery

Initial Parsonnet's score and modified Parsonnet's score

The Parsonnet score was developed in 1989 by Victor Parsonnet and is notable for having been a pioneer in systematic risk stratification in cardiac surgery and applicable to different populations [8]. This score is simple, additive, and grades the severity of illness of patients into five groups (Table 1) [8, 9]. This useful score has been rapidly taken up by several cardiac surgery teams, and other authors have confirmed its predictive value on hospital mortality and morbidity

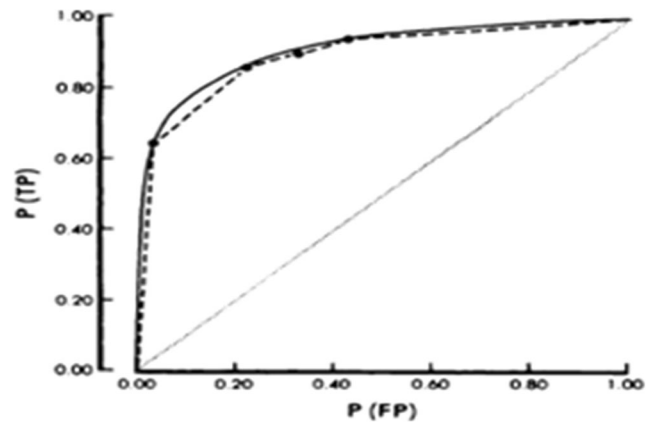


Fig. 1 ROC curve. *Dashed line* = empirical curve, *solid line* = smoothed (Gaussian-based) curve, and *straight line* = no discrimination

[10]. Two risk factors of the “initial Parsonnet’s score” are however imprecise, and their weights are arbitrarily chosen by the surgeon (catastrophic states, other rare circumstances). Thus, the reliability of the initial Parsonnet’s score decreases when these two risk factors are present. This original score was later modified, including 30 new risk factors according to the SUMMIT system [11, 12]. These 30 new risk factors take the place of the two imprecise risk factors of the initial score, and this new score is referred to as the “modified Parsonnet’s score” [13].

Improvements in anesthesia, modern-day intensive care, extracorporeal circulation, myocardial preservation, and surgical therapy advances have contributed to lower mortality rates than those predicted by Parsonnet’s score. Initial Parsonnet’s score overestimates mortality has been criticized as it does not take into account the quality of a patient’s coronary arteries in predicting operative risk and probably overestimates the risk of age. There have been several other computerized attempts at predicting operative risk more accurately, but this is always at the expense of simplicity and practicality of use. Parsonnet’s score has therefore become a familiar and workable tool for cardiologist and cardiac surgeon alike [14].

EuroSCORE

The European System for Cardiac Operative Risk Evaluation (EuroSCORE) identifies a number of risk factors which help to predict mortality from cardiac surgery [1]. The predicted mortality (in percent) is calculated by adding the weights assigned to each factor. Since its initial publication in 1999, EuroSCORE has been widely used in Europe and elsewhere and has been the subject of several studies [15, 16]. EuroSCORE is a simple, objective, and up-to-date system for assessing heart surgery, soundly based on one of the largest databases in European cardiac surgical history [15].

Table 1 Description of the initial Parsonnet's score [9]

Risk factors	Assigned weight
Female gender	1
Morbid obesity (≥ 1.5 , ideal weight)	3
Diabetes	3
Hypertension (systolic blood pressure >140 mmHg)	3
Ejection fraction	
≥ 50 %	0
30–49 %	2
<30 %	4
Age	
70–74	7
75–79	12
≥ 80	20
Reoperation	
First	5
Second	10
Preoperative intra-aortic balloon pump	20
Left ventricular aneurysm	5
Emergency surgery following PTCA or catheterization complications	10
Dialysis dependency (peritoneal dialysis or hemodialysis)	10
Catastrophic states (acute structural defect, cardiogenic shock, acute renal failure)	10–50
Other rare circumstances (paraplegia, pacemaker dependency, severe asthma, congenital heart disease in adult)	2–10
Mitral surgery	5
Mitral surgery and pulmonary artery pressure ≥ 60 mmHg	8
Aortic surgery	5
Aortic surgery and pressure gradient >120 mmHg	7
CABG at the time of valve surgery	2

CABG coronary artery bypass graft, PTCA percutaneous transluminal coronary angioplasty

There are two extremes in the selection of a risk stratification system. Accuracy can be achieved by assessing a large number of risk factors for an individual patient and comparing the findings with the results of a large database such as EuroSCORE. Such a system should provide very accurate risk assessment for small subgroups of patients. This approach required the gathering of large amounts of patient data and complex statistical operations. It is impossible to implement without sophisticated information technology which is not yet available to all hospitals. On the other hand, very simple models relying on one or two risk factors (such as age and sex, for example) are also possible. This approach would have some use for the overall assessment of a hospital's performance but is unlikely to be useful for risk assessment for an individual patient and is likely to perpetuate a reluctance to operate on high-risk patients. A compromise must be reached so that the system recognizes common risk factors and is able to provide some degree of risk prediction yet remains simple enough to use at the point of delivery of care [15]. EuroSCORE satisfies these requirements.

Most EuroSCORE risk factors are derived from the clinical status of the patient. Only four risk factors are related to the operation, and these are factors that are difficult to influence through subtle variation in surgical decision-making [15].

The logistic EuroSCORE was found more suitable for individual risk prediction in very high-risk patients. Using the same risk factors as the additive model, the logistic regression version of the score (the "logistic EuroSCORE") can be calculated. For a given patient, the "logistic EuroSCORE" which is the predicted mortality according to the logistic regression equation can be achieved with the following formula:

$$\text{Predicted mortality} = \frac{e^{\left(\beta_0 + \sum \beta_i X_i\right)}}{1 + e^{\left(\beta_0 + \sum \beta_i X_i\right)}}$$

where e is the natural logarithm = 2.718281828... β_0 is the constant of the logistic regression equation = -4.789594 β_i is the coefficient of the variable X_i in the logistic regression

equation $X_i = 1$ if a categorical risk factor is present and 0 if it is absent. For age, $X_i = 1$ if patient age < 60 ; X_i increase by one point per year thereafter. Hence, for age 59 or less $X_i = 1$, age 60 $X_i = 2$, age 61 $X_i = 3$, and so on [16].

The logistic model was more discriminatory than the additive in tracing extended intensive care unit (ICU) stay [4].

EuroSCORE II

The EuroSCORE, worldwide used as a model for prediction of mortality after cardiac surgery, was renewed. On 3 October 2011, the EuroSCORE II was launched at the 2011 European Association for Cardiothoracic Surgery (EACTS) meeting in Lisbon, and the online calculator (www.EuroSCORE.org) has been updated to use this new risk stratification model [17].

EuroSCORE II, an update of the logistic EuroSCORE model, uses similar methodology but is derived from a more current data set and refined to incorporate evidence-based improvements and to reflect better current cardiac surgical practice [17]. This score reduces the overestimation of the calculated risk by the initial EuroSCORE [18]. EuroSCORE predicts 30-day mortality, 1-year mortality, and costs in cardiac surgery.

Society of Thoracic Surgeons (STS) score

The Society of Thoracic Surgeons National Cardiac Database (STS NCD) was created in 1989, and it has become the largest clinical database of its kind. The primary aim for the development of the STS model was the support of national quality improvement programs. Now, it is also used for research focusing on improvement of patient care and outcome [5, 19]. The STS NCD is unparalleled in terms of its size and comprehensiveness: data were collected prospectively from more than 950 participating centers in the USA. The STS NCD now also includes more than 3.6 million surgical procedures [20]. STS risk models for various cardiac procedures have been developed since 1999 and have undergone periodic revisions [20, 21]. A wide variety of endpoints are included in some of the models calculating risk for isolated coronary artery bypass grafting, valve surgery, or combined surgeries [21].

STS score includes 50 clinical variables of preoperative variables. The predictive performance of the STS algorithms is in general comparable with other systems and remains the most widely used model in the USA [21]. STS score also allows for the calculation of postoperative morbidity [5].

ASCERT score

Most survival prediction models for coronary artery bypass grafting surgery are limited to in-hospital or 30-day endpoints. So, a long-term survival model was developed using

data from the Society of Thoracic Surgeons Adult Cardiac Surgery Database and Centers for Medicare and Medicaid Services [22].

The American College of Cardiology Foundation, the STS, and the Duke Clinical Research Institute were collaborated on a comparative effectiveness study (American College of Cardiology Foundation–Society of Thoracic Surgeons Collaboration on the Comparative Effectiveness of Revascularization Strategies [ASCERT]) of coronary artery bypass graft (CABG) and percutaneous coronary interventions (PCI), funded by the National Heart, Lung, and Blood Institute of the National Institutes of Health. The first aim of the ASCERT study is to develop novel, long-term mortality risk prediction models for CABG and PCI to estimate the time-dependent effect of preoperative patient factors on medium- and long-term mortality after CABG [23]. This provides valuable information for shared decision-making, comparative effectiveness research, quality improvement, and provider profiling [22].

SYNTAX score

The SYNTAX score (SS), based on the anatomical characteristics of coronary artery disease, is recommended by practice guidelines to decide between PCI with drug-eluting stent or CABG surgery in patients with unprotected left main coronary artery stenosis or three-vessel disease. The score has been criticized for considering only anatomical variables, without taking into account other factors that may influence the results of both procedures [24]. Syntax score is very useful for understanding prognosis of left main coronary disease patients.

As the SS was initially validated for patients with native CAD, it cannot be implemented in patients with CABG. To help address this issue, the CABG SS was developed. This score can be calculated by computing first the baseline SS of native vessels and then subtracting points on the basis of graft functionality. Despite the limited power of the study, it suggested a trend toward higher all cause death and major adverse cardiac events (MACE) in patients with high CABG SS. One major limitation of this score is that it does not take into consideration the type of graft used [25].

SS ideally should be calculated by a “heart team” rather than individually. SS when associated with clinical variables like age, creatinine level, and EF (ACEF), “Clinical Syntax Score” was superior in predicting major adverse cardiac events (MACE), and similar predicting mortality as EuroSCORE, it offers additional advantage in prediction of ischemic endpoints [26].

Therefore, the SYNTAX score II (SS II) was developed using baseline features and 4-year follow-up information recorded in the SYNTAX study. In addition to SS and the presence or absence of left main coronary artery disease, this novel score considers age, gender, creatinine clearance, left ventricular ejection fraction, chronic obstructive lung disease, and

peripheral vascular disease. The predictive accuracy of the score for mortality was validated in the study population and in a multinational registry [24].

The SS II allowed for the individualized assessment of long-term mortality in patients with LM/multi-vessel CAD undergoing either PCI or CABG, compared to the grouping of risk (low, intermediate, high) with the anatomical SS. The SS II was developed in the randomized SYNTAX Trial and validated in the Drug-eluting stent for left main coronary artery disease (DELTA) registry [25].

Amphiascore

Amphiascore was a predictive model for major adverse outcomes after CABG and/or heart valve operation in a large cohort of patients from a single institute in the Netherlands. It was created in Amphia Hospital (a teaching hospital at Breda, the Netherlands) which has kept an expanding database for more than 10 years with pre, per, and postoperative characteristics of all patients who underwent cardiac surgery [27].

This score was developed depending on the three outcomes such as,

1. In-hospital death, defined as death during hospitalization in the Amphia hospital or in one of the affiliated hospitals
2. MACE defined as in-hospital death or perioperative myocardial infarction (Q-wave or non-Q-wave with creatinine kinase MB fraction (CK-MB). 50 mg/l in the year 1997 or CK-MB. 100 mg/l from 1998 onwards) or ventricular tachycardia/fibrillation
3. Extended length of stay (ELOS) defined as intensive care length of stay of at least 3 days or in-hospital death [27]

Amphiascore performs in discriminating patients with respect to in-hospital death. This model for predicting major adverse cardiac events and extended length of stay on intensive care may be useful tools in categorizing patients in various subgroups of risk for postoperative morbidity [27].

CABDEAL

The CABDEAL model for risk stratification has been developed solely for the prediction of morbidity. The CABDEAL model was developed using a Bayesian approach, and it includes seven preoperative risk factors which are assigned point values: renal dysfunction, advanced age, obesity, diabetes mellitus, emergency surgery, arrhythmia and/or unstable angina or previous AMI, and chronic obstructive lung disease [28]. The maximum possible risk score is 10. A score of 3 or more is associated with increased morbidity, the outcome criteria for which were in the original model the following: at least one severe complication postoperatively and postoperative length of stay (POS) of >11 days, or at least two mild

complications postoperatively, or death during the hospital stay. Due to changes in treatment policies, the POS limit for increased morbidity was lowered to 8 days. Severe postoperative complications included reoperation due to bleeding or low cardiac output, mediastinitis, pneumonia, or prolonged ventilator support (>36 h) or ICU stay longer than 3 days, stroke (cerebrovascular accident), acute renal failure requiring dialysis or a more than 50 % increase in serum creatinine level (compared to preoperative values), severe cardiac failure requiring inotropic support or intra-aortic balloon pump insertion or perioperative myocardial infarction (MI), severe arrhythmia (ventricular fibrillation or asystole), or death. Mild complications included superficial wound infection (leg or sternum), atrial fibrillation, or other mild supraventricular arrhythmia [28].

Cleveland clinic score

It is also termed as Higgins score [5]. This model is developed using logistic regression analysis. The Cleveland model is designed to predict both morbidity and mortality. The score level for significantly increased mortality (>10 %) in the Cleveland model is 6. For increased morbidity, the score level is 4 in the Cleveland model [28]. Outcome criteria for increased morbidity in the Cleveland model comprised one of the following: central nervous system complication (stroke), cardiac complication (perioperative MI, need for mechanical assistance device, intra-aortic balloon pump), acute renal failure requiring dialysis, serious infection (mediastinitis, sepsis), return to surgery due to bleeding, prolonged mechanical ventilation (>3 days), or death [28]. Cleveland Clinic score has one of the highest discriminating powers after EuroSCORE (for CABG only).

French score

This score was highly predictive for mortality and severe morbidity. A national database was developed as a part of this score, which includes anonymous information from two thirds of all cardiac surgery cases. Nationwide results for France allow each center freely to assess its results. This self-assessment approach is the most accurate way of quality of care assessment [29].

Magovern score

Clinical risk score (CRS) is based entirely on preoperative data, and it reliably predicts morbidity and mortality for patients undergoing CABG. This score was derived from the Allegheny General Hospital's cardiothoracic surgery database. The database was implemented in July 1991, with prospective data collection on all patients undergoing cardiac surgery [30]. This model identifies patients at high risk for

postoperative morbidity and mortality. The model was developed from the experience of one institution. It has been validated at the Allegheny General Hospital but not at other institutions [30]. The most powerful predictors were those associated with emergency operation and depressed cardiac function. Other factors reflect comorbid disease processes, such as chronic obstructive pulmonary disease, renal insufficiency, peripheral vascular disease, diabetes, low serum albumin, and anemia. Finally, age, gender, and low body mass were also significant predictors. Obesity was not a predictor of morbidity or mortality [30]. Magovern score can also predict 1-year mortality as well as 30 days.

New York's Cardiac Surgery Reporting System (NYS) score

New York's Cardiac Surgery Reporting System was used to develop an in-hospital and 30-day logistic regression model for patients undergoing CABG surgery in 2009, and this model was converted into a simple linear risk score that provides estimated in-hospital and 30-day mortality rates for different values of the score. The accuracy of the risk score in predicting mortality was tested. This score was also validated by applying it to 2008 New York CABG data. Subsequent analyses evaluated the ability of the risk score to predict complications and length of stay [31]. There are seven risk factors comprising the score, with risk factor scores ranging from 1 to 5, and the highest possible total score is 23. The risk score is a simple way of estimating short-term mortality that accurately predicts mortality in the year the model was developed as well as in the previous year. Perioperative complications and length of stay are also well predicted by the risk score [31].

Northern New England score (NNE)

Northern New England Score was developed from a prospective regional study which was conducted to identify factors associated with in-hospital mortality among patients undergoing isolated CABG. A prediction rule was developed and validated based on the data collected [32]. Variables used to construct the regression model of in-hospital mortality included age, sex, body surface area, presence of comorbid disease, history of CABG, left ventricular end-diastolic pressure, ejection fraction score, and priority of surgery. The model significantly predicted the occurrence of in-hospital mortality [32].

Ontario province risk score

A multicenter population-based study was conducted to develop and validate a risk index for mortality, ICU length of stay, and postoperative length of stay after cardiac surgery using data from all nine adult cardiac surgery institutions in Ontario. A six-variable risk index (age, sex, left ventricular

function, type of surgery, urgency of surgery, and repeat operation) was developed using logistic regression analysis to predict in-hospital mortality, ICU stay in days, and postoperative stay in days after cardiac surgery. It is also termed as Provincial Adult Cardiac Care Network (PACCN) [33].

Mortality, ICU length of stay, and postoperative length of stay after cardiac surgery can be predicted using a simple six-variable risk index. Thus, the index has multiple potential applications, including comparing patient outcomes and resource use among different surgeons and hospitals, counseling patients about the risks of cardiac surgery, and use in patient and staff scheduling when resources are limited [33]. Hospitals are given their risk-adjusted outcomes so that they can evaluate their relative performance with the goal of continuous quality improvement. Use of an additive model allows clinicians to see how well they are performing relative to others at different levels of patient risk (e.g., low, medium, or high risk), and it provides a summary measure (i.e., the mean risk score) of a hospital's case mix severity. The major difference between this model and others that have been developed is its lack of inclusion of comorbid diseases [33].

Pons risk score

A risk stratification model was developed to assess open heart surgery mortality in Catalonia (Spain) in order to use risk-adjusted hospital mortality rates as an approach to analyze quality of care. A risk stratification model through a metacentric, prospective, and exhaustive collection of data in all types of open heart procedures was developed [34]. The main variable analyzed was surgical mortality, which was defined as that occurring during the 30 days after the intervention or during hospitalization irrespective of the length-of-stay (hospital-to-hospital transfer was not considered discharge). Risk-adjusted surgical mortality was the outcome studied as an approach to hospital quality of care [34]. A score was generated for each patient using the weight of the model. Patients were stratified into categories depending on the score. Several cut-off points were examined in order to determine the best association of the score with surgical mortality. Finally, five risk levels were selected depending on the score: level 1 or low risk (0–10), level 2 or fair risk (11–15), level 3 or high risk (16–20), level 4 or very high risk (21–30), and level 5 or extremely high risk (≥ 31) [34].

Toronto risk score (TRS)

The TRS is a valid measure of acuity that can identify patients who are at high risk of experiencing an adverse effects (AE) and having prolonged length of stay after any cardiac surgery procedure, capture changes in acuity over time, and allow for continuous quality performance evaluation [35]. The primary binomial outcome for this model was any postoperative AE

following cardiac surgery. AE was defined as any of the following: operative death, a perioperative myocardial infarction defined by electrocardiogram and enzymatic criteria, low cardiac output syndrome (systolic blood pressure less than 90 mmHg and cardiac index less than 2.1 L/min/M² lasting longer than 15 min despite adequate preload), a perioperative stroke, new postoperative renal failure (defined as the need for any form of dialysis), or deep sternal wound infection [35].

UK national score

In the UK, a model has recently been produced to predict mortality following CABG (also known as UK Society of Cardiothoracic Surgeons score) [36].

It is easy to implement, objective, and accurate predictor of observed mortality, allowing comparison between surgeons and units [36].

Veterans administration (VA) score

The VA Continuous Improvement in Cardiac Surgery Study was initiated in 1987 to develop risk-adjusted outcome models of continuous quality improvement and to evaluate the quality of all cardiac surgical procedures performed in VA medical centers [37]. The data included preoperative risk factors, surgical data, postoperative death and complications data, and length of stay. In the VA, 30-day surgical death was defined as death from any cause occurring within the first 30 days after surgery or death at a later time occurring as a direct result of a perioperative complication. This broad-based definition also includes all index hospitalization deaths [38]. The models have been developed for two groups of patients: those undergoing CABG only and those undergoing a valve operation or other cardiac procedures. Procedural designation for patients undergoing aortic valve replacement, mitral valve replacement, great vessel repair, and other procedures has been added to the multivariate analysis to identify risk factors specific for each of these procedures [38]. The preoperative risk variables are entered into a backward stepwise logistic regression analysis with surgical death as a dependent variable. The expected surgical death rate for each patient is calculated from this logistic regression model and compared with the observed surgical death rate [38].

AusSCORE

A new model was developed for predicting 30-day mortality after isolated CABG for the Australian population. The Australian Society of Cardiac and Thoracic Surgeons (ASCTS) database has prospectively collected information about adult patients having cardiac surgery in six public hospitals in Victoria since June 2001. The ASCTS database contains information on patient risk factors (including preoperative cardiac

status and previous interventions), intra-operative details (including the procedure performed, myocardial protection, and procedural durations), and postoperative outcomes. The index outcome was mortality, defined as death within 30 days post-operatively [39]. The risk factors in the AusSCORE are as follows: age, NYHA class, urgency of procedure, ejection fraction estimate, previous cardiac surgery, hypercholesterolemia (lipid-lowering treatment), peripheral vascular disease, and cardiogenic shock [39].

Ambler score

This is the first risk model that predicts in-hospital mortality for aortic and/or mitral valve patients with or without concomitant CABG. Based on a large national database of heart valve patients, this model has been evaluated successfully on patients who had valve surgery during a subsequent time period. It is simple to use, includes routinely collected variables, and provides a useful tool for patient advice and institutional comparisons. This risk model provides a simple, useful tool for risk stratification for most patients undergoing valve surgery [40].

German Aortic Valve Score

A scoring system was developed to predict mortality in aortic valve procedures in adults. German Aortic Valve Score (German AV Score) is based upon the comprehensive data pool mandatory by law in Germany [41]. It is well known that a predictive model works best in the setting where it was developed; therefore, the German AV Score fits well to the patient population in Germany. It was designed for fair and reliable outcome evaluation. It allows comparison of predicted and observed mortality for conventional aortic valve surgery and trans-catheter aortic valve implantation in low-, moderate-, and high-risk groups. Thus, it enables primarily a risk-adjusted benchmark of outcome and fosters the efforts for continuous improvement of quality in aortic valve procedures [41]. This score fairly predicts risk (absolute) for TAVI. However, it does not provide additional prognostic information.

Observant score

A simple score developed using pre-procedural variables, for prediction of 30-day mortality after trans-catheter aortic valve replacement (TAVR). The risk score was built from the (TAVR) cohort of the Observational Study of Appropriateness, Efficacy and Effectiveness of AVR-TAVR Procedures for the Treatment of Severe Symptomatic Aortic Stenosis (OBSERVANT) study. Briefly, OBSERVANT is a national, observational, prospective, cohort study that enrolled patients with aortic stenosis undergoing TAVR or surgical aortic valve replacement at 95 Italian centers [42].

The principal findings of this study, in which a new computational tool specific to patients undergoing TAVR (OBSERVANT score) was created and internally validated, are the following: (1) early mortality of TAVR at 30 days can be predicted using seven clinical variables that are routinely available before the intervention, (2) renal dysfunction (defined as glomerular filtration rate (GFR) <45 mL/min) is the single most powerful independent predictor of 30-day mortality, and (3) the OBSERVANT model, when validated against an independent internal data set, provided better discrimination, goodness of fit, and global accuracy for 30-day mortality than the logistic EuroSCORE, which is more complex (17 variables) and does not incorporate factors specific to the inherent risk of patients undergoing TAVR [42].

Risk scores in pediatric cardiac surgery

Aristotle score

Aristotle the Greek philosopher proposed the Consensus Theory. Consensus theory holds that truth is whatever is agreed upon, or in some versions, might come to be agreed upon, by some specified group.

The analysis of congenital heart surgery outcomes is challenging owing to the large number of surgical procedures that vary in complexity. One method that has been proposed for complexity-adjusted outcomes analysis is known as the Aristotle Basic Complexity Score (ABC score). The ABC score expresses the case complexity of congenital heart surgery procedures based on three components: the potential for mortality, the potential for morbidity, and the technical difficulty of the procedure [43]. The Aristotle project, involving a panel of expert surgeons, started in 1999 and included 50 pediatric surgeons from 23 countries. The complexity was based on the procedures as defined by the STS/EACTS International Nomenclature and was undertaken in two steps: the first step was establishing the basic score, which adjusts only the complexity of the procedures. It is based on three factors: the potential for mortality, the potential for morbidity, and the anticipated technical difficulty. A questionnaire was completed by the 50 centers. The second step was the development of the Comprehensive Aristotle Score, which further adjusts the complexity according to the specific patient characteristics. It includes two categories of complexity factors, the procedure-dependent and independent factors. After considering the relationship between complexity and performance, the Aristotle Committee is proposing that performance = complexity × outcome [44].

The utility of the ABC score depends on its ability to correctly classify procedures according to their potential for morbidity, mortality, and technical difficulty. Although the difficulty of a procedure is inherently subjective and difficult to validate, the accuracy of the ABC score with respect to mortality and morbidity can be objectively determined for procedures with adequate sample size [43]. The ABC score generally discriminates between low-risk and high-risk congenital procedures making it a potentially useful covariate for case mix adjustment in congenital heart surgery outcome analysis [43].

RACH'S score

Under the leadership of Kathy Jenkins, M.D. (Children's Hospital, Boston, MA, USA) and colleagues, the Risk Adjustment in Congenital Heart Surgery (RACHS-1) method was developed to adjust for baseline case mix differences in comparisons of discharge mortality among groups of patients undergoing pediatric and congenital cardiac surgery. The RACHS-1 method was created using a combination of judgment-based and empirical methodology. It is one of the first widely accepted complexity adjustment tools developed in this field [45].

Initially, an 11-member nationally representative panel of pediatric cardiologists and cardiac surgeons grouped cardiac surgical procedures into six risk categories based on expected discharge mortality (category 1 [lowest risk] to category 6 [highest risk]), although functionally, findings have shown that category 5 has too few cases for accurate estimates of mortality rates. The categories then were refined using empirical data from two large data sets: one from the Pediatric Cardiac Care Consortium (PCCC) and the other generated from statewide hospital discharge databases. In addition to risk group, the RACHS-1 method incorporates age at surgery, prematurity, presence of a major non-cardiac structural anomaly, and whether multiple surgical procedures were performed simultaneously or not. RACHS-1 system discriminates better at the higher end of complexity [45].

Heart transplantation risk scores

CARRS score

The CARRS scoring model is developed using five pre-transplant risk factors (C for prior CVA, A for albumin <3.5 mg/dL, R for retransplant, R for glomerular filtration rate <40 mL/min, and S for >2 prior sternotomies with 2 points for each, except renal with 1 point) sorted out high and low survival groups [46].

A prognostic risk score (CARRS) derived from these factors stratified survival post-heart transplantation (post-HTx) in

high-risk (3+ points) versus low-risk (0–2 points) patients (87.9 versus 52.9 % at 1-year post-HTx; 65.9 versus 28.4 % at 5-year post-HTx; $P < 0.001$). Low-risk alternate patients had post-HTx survival comparable with regular patients [47].

The creation of the CARRS score was based on the results of the univariable and multivariable proportional hazards risk analysis. The multivariable analysis failed to identify multiple independent predictors and could not score the covariables by relative hazard. Prior cerebral vascular accident, albumin < 3.5 mg/dL, re-HTx, renal dysfunction (GFR < 40 mL/min), and > 2 prior sternotomies were associated with poor survival after heart transplantation (Htx). Univariable analysis and Kaplan–Meier analysis help to score predictors with a hazards ratio > 2 and a pronounced early survival effect with 2 points and GFR < 40 mL/min with 1 point (attributable to lower impact compared with the other factors). Significant univariable predictors with > 15 % missing data or negligible hazards, as well as intra-operative and donor risk factors, were also not included. Stratification of high- and low-point values was varied according to survival predictive power before a final inflection point was set at 0 to 2 points for low risk and 3 to 9 points for high risk [47].

The risk stratification using the non-invasive CARRS score allows identification of patients with unacceptably high mortality after HTx. Among patients previously accepted for alternate donor listing, application of the CARRS score identifies patients with unacceptably high mortality after HTx and those with a survival similar to regularly listed patients [47].

IMPACT score

IMPACT score is a recipient risk index developed for predicting short-term mortality after orthotopic heart transplantation (OHT). This model utilized United Network for Organ Sharing (UNOS) data to develop a novel quantitative recipient risk score for use in OHT [48].

From the final model, a 50-point recipient risk score (Index for Mortality Prediction After Cardiac Transplantation [IMPACT]) was created approximating the magnitude of relative odds of 1-year mortality and applied independently to all members of the derivation and validation sets. Cumulative survival was estimated using the Kaplan–Meier method, with censoring for those individuals lost to follow-up or alive at the end of study time (administratively censored). All means are presented with standard deviations, medians with interquartile ranges, and odds ratios with 95 % confidence intervals. Statistical analyses were performed with STATA software (v9.2 SE; StataCorp LP, College Station, TX) [48].

This score was a novel internally validated OHT recipient risk score, which is highly predictive of 1-year mortality. This risk index may prove valuable for patient prognosis, organ allocation, and research stratification in OHT [48].

Conclusions

Risk scores measure cardiac mortality but do not assess quality of care. The comparisons of operative mortality among centers are meaningless without risk adjustments derived from case mix. However, risk scores are useful tools to assess costs related to patient severity. Risk scores measure the risk of care not the quality of care. Overall risk scores are an essential tool in cardiac surgery practice for risk assessment, decision-making, and consent.

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

Funding This study did not receive funding or grants from any agencies.

Ethical approval This article does not contain any studies with human participants performed by any of the authors.

Informed consent Since no human subjects were involved in the study, informed consent was not obtained.

References

1. Geissler HJ, Holz P, Marohl S, et al. Risk stratification in heart surgery: comparison of six score systems. *Eur J Cardiothorac Surg.* 2000;17:400–6.
2. Higgins TL. Quantifying risk and assessing outcome in cardiac surgery. *J Cardiothorac Vasc Anesth.* 1998;12:330–40.
3. Shahian DM, Blackstone EH, Edwards FH, et al. Cardiac surgery risk models: a position article. *Ann Thorac Surg.* 2004;78:1868–77.
4. Messaoudi N, De Cocker J, Stockman BA, Bossaert LL, Rodrigus IE. Is EuroSCORE useful in the prediction of extended intensive care unit stay after cardiac surgery? *Eur J Cardiothorac Surg.* 2009;36:35–9.
5. Prins C, Jonker ID, Botes L, Smit FE. Cardiac surgery risk-stratification models. *Cardiovasc J Afr.* 2012;23:160–4.
6. Nilsson J, Algotsson L, Hoglund P, Luhrs C, Brandt J. Comparison of 19 preoperative risk stratification models on open-heart surgery. *Eur Heart J.* 2006;27:867–74.
7. Hanley JA, Mc-Neil BJ. The meaning and use of the area under a receiver operating characteristics (ROC) curve. *Radiology.* 1982;143:29–36.
8. 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;79:13–112 (suppl 1).
9. Gabrielle F, Roques F, Michel P, et al. Is the Parsonnet's score a good predictive score of mortality in adult cardiac surgery: assessment by a French multicentre study. *Eur J Cardiothorac Surg.* 1997;11:406–14.
10. Nashef SAM, Carey F, Silcock MM, Oomen PK, Levy RD, Jones MT. Risk stratification for open heart surgery: trial of the Parsonnet system in a British hospital. *Br Med J.* 1992;305:1066–7.
11. Clark RE. It is time for a national cardiothoracic surgical data base. *Ann Thorac Surg.* 1989;48:755–6.

12. Hattler BG, Madia C, Johnson C, et al. Risk stratification using the Society of Thoracic Surgeons program. *Ann Thorac Surg.* 1994;52:1348–52.
13. Kacila M, Tiwari KK, Granov N, Omerbasic E, Straus S. Assessment of the initial and modified Parsonnet score in mortality prediction of the patients operated in the Sarajevo Heart Center. *Bosn J Basic Med Sci.* 2010;10(2):165–8.
14. Lawrence DR, Valencia O, Smith EE, Murday A, Treasure T. Parsonnet score is a good predictor of the duration of intensive care unit stay following cardiac surgery. *Heart.* 2000;83:429–32.
15. Nashef SAM, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R. European system for cardiac operative risk evaluation (EuroSCORE). *Eur J Card Thorac Surg.* 1999;16:9–13.
16. Roques F, Michel P, Goldstone AR, Nashef SA. The logistic EuroSCORE. *Eur Heart J.* 2003;24:882–3.
17. Nashef SA, Roques F, Sharples LD, et al. EuroSCORE II. *Eur J Cardiothorac Surg.* 2012;41:734–45.
18. Noyez L, Kievit PC, van Swieten HA, de Boer MJ. Cardiac operative risk evaluation: the EuroSCORE II, does it make a real difference? *Neth Heart J.* 2012;20:494–8.
19. Welke KF, Ferguson Jr TB, Coombs LP, et al. Validity of the Society of Thoracic Surgeons national adult cardiac surgery database. *Ann Thorac Surg.* 2004;77:1137–9.
20. Shahian DM, O'Brien SM, Filardo G, et al. The Society of Thoracic Surgeons 2008 cardiac surgery risk models: part 1—coronary artery bypass grafting surgery. *Ann Thorac Surg.* 2009;88:S2–S22.
21. Granton J, Cheng D. Risk stratification models for cardiac surgery. *Semin Cardiothorac Vasc Anesth.* 2008;12:167–74.
22. Shahian DM, O'Brien SM, Sheng S, et al. Predictors of long-term survival after coronary artery bypass grafting surgery. *Circulation.* 2012;125:1491–500.
23. Klein LW, Edwards FH, DeLong ER, Ritzenthaler L, Dangas GD, Weintraub WS. ASCERT: the American College of Cardiology Foundation—the Society of Thoracic Surgeons collaboration on the comparative effectiveness of revascularization strategies. *J Am Coll Cardiol Interv.* 2010;3:124–6.
24. Farooq V, van Klaveren D, Steyerberg EW, et al. SYNTAX score II: a useful tool to decide type of revascularization therapy. *Lancet.* 2013;381:639–50.
25. Yadav M, Palmerini T, Caixeta A, et al. Prediction of coronary risk by SYNTAX and derived scores synergy between percutaneous coronary intervention with Taxus and cardiac surgery. *J Am Coll Cardiol.* 2013;62:1219–30.
26. Garg S, Samo G, Garcia-Garcia HM. ARTS-II investigators. A new tool for the risk stratification of patients with complex. Coronary artery disease: the clinical SYNTAX score. *Circ Cardiovasc Interv.* 2010;3:317–26.
27. Huijskes RVHP, Rosseel PMJ, Tijssen JGP. Outcome prediction in coronary artery bypass grafting and valve surgery in the Netherlands: development of the Amphiascore and its comparison with the EuroSCORE. *Eur J Cardiothorac Surg.* 2003;24:741–9.
28. Kurki TS, Jarvinen O, Kataja MJ, Laurikka J, Tarkka M. Performance of three preoperative risk indices; CABDEAL, EuroSCORE and Cleveland models in a prospective coronary bypass database. *Eur J Cardiothorac Surg.* 2002;21:406–10.
29. Roques F, Gabrielle F, Michel P, De Vincentiis C, David M, Baudet E. Quality of care in adult heart surgery: proposal for a self-assessment approach based on a French multicenter study. *Eur J Cardiothorac Surg.* 1995;9:433–9.
30. Magovern JA, Sakert T, Magovern GJ, et al. A model that predicts morbidity and mortality after coronary artery bypass graft surgery. *J Am Coll Cardiol.* 1996;28:1147–53.
31. Hannan EL, Farrell LS, Wechsler A, et al. The New York Risk score for in-hospital and 30-day mortality for coronary artery bypass graft surgery. *Ann Thorac Surg.* 2013;95:46–54.
32. O'Connor GT, Plume SK, Olmstead EM, et al. Multivariate prediction of in-hospital mortality associated with coronary artery bypass graft surgery. Northern New England Cardiovascular Disease Study Group. *Circulation.* 1992;85:2110–8.
33. Tu JV, Jaglal SB, Naylor CD. Multicenter validation of a risk index for mortality, intensive care unit stay, and overall hospital length of stay after cardiac surgery. Steering committee of the provincial adult cardiac care network of Ontario. *Circulation.* 1995;91:677–84.
34. Pons JMV, Granados A, Espinas JA, Borrás JM, Martín I, Moreno V. Assessing open heart surgery mortality in Catalonia (Spain) through a predictive risk model. *Eur J Cardiothorac Surg.* 1997;11:415–23.
35. Ivanov J, Borger MA, Rao V, David TE. The Toronto risk score for adverse events following cardiac surgery. *Can J Cardiol.* 2006;22:221–7.
36. Bridgewater B, Neve H, Moat N, Hooper T, Jones M. Predicting operative risk for coronary artery surgery in the United Kingdom: a comparison of various risk prediction algorithms. *Heart.* 1998;79:350–5.
37. Grover FL, Johnson RR, Marshall G, Hammermeister KE. Factors predictive of operative mortality among coronary artery bypass subsets. *Ann Thorac Surg.* 1993;56:1296–306. **discussion 1306–1307.**
38. Grover FL, Shroyer AL, Hammermeister K, et al. A decade's experience with quality improvement in cardiac surgery using the Veterans Affairs and Society of Thoracic Surgeons national databases. *Ann Surg.* 2001;234:464–72.
39. Reid C, Billah B, Dinh D, et al. An Australian risk prediction model for 30-day mortality after isolated coronary artery bypass: the AusSCORE. *J Thorac Cardiovasc Surg.* 2009;138:904–10.
40. Ambler G, Omar RS, Royston P, et al. Generic, simple risk stratification model for heart valve surgery. *Circulation.* 2005;112:224–31.
41. Kottling J, Schiller W, Beckmann A, et al. German aortic valve score: a new scoring system for prediction of mortality related to aortic valve procedures in adults. *Eur J Cardiothorac Surg.* 2013;43:971–7.
42. Capodanno D, Barbanti M, Tamburino C, et al. A simple risk tool (the OBSERVANT score) for prediction of 30-day mortality after transcatheter aortic valve replacement. *Am J Cardiol.* 2014;113:1851–8.
43. O'Brien SM, Jacobs JP, Clarke DR, et al. Accuracy of the Aristotle basic complexity score for classifying the mortality and morbidity potential of congenital heart surgery operations. *Ann Thorac Surg.* 2007;84:2027–37.
44. Lacour-Gayet F, Clarke D, Jacobs J, et al. The Aristotle score: a complexity-adjusted method to evaluate surgical results. *Eur J Cardiothorac Surg.* 2004;25:911–24.
45. Jacobs JP, Jacobs ML, Lacour-Gayet FG, et al. Stratification of complexity improves the utility and accuracy of outcomes analysis in a multi-institutional congenital heart surgery database: application of the risk adjustment in congenital heart surgery (RACHS-1) and Aristotle systems in the Society of Thoracic Surgeons (STS) congenital heart surgery database. *Pediatr Cardiol.* 2009;30:1117–30.
46. Copeland JG. Perspective on a risk score for cardiac transplantation. *Circ Heart Fail.* 2013;6:362–3.
47. Schulze PC, Jiang J, Yang J, et al. Preoperative assessment of high risk candidates to predict survival after heart transplantation. *Circ Heart Fail.* 2013;6:527–34.
48. Weiss ES, Allen JG, Amaoutakis GJ, et al. Creation of a quantitative recipient risk index for mortality prediction after cardiac transplantation (IMPACT). *Ann Thorac Surg.* 2011;92:914–22.