ELSEVIER

European Journal of Cardio-thoracic Surgery 35 (2009) 760—768

EUROPEAN JOURNAL OF CARDIO-THORACIC SURGERY

www.elsevier.com/locate/ejcts

Risk-prediction for postoperative major morbidity in coronary surgery☆

Pedro E. Antunes, José Ferrão de Oliveira, Manuel J. Antunes *

Cardiothoracic Surgery, University Hospital, Coimbra, Portugal

Received 13 June 2008; received in revised form 14 October 2008; accepted 22 October 2008; Available online 25 February 2009

Abstract

Objective: Analysis of major perioperative morbidity has become an important factor in assessment of quality of patient care. We have conducted a prospective study of a large population of patients undergoing coronary artery bypass surgery (CABG), to identify preoperative risk factors and to develop and validate risk-prediction models for peri- and postoperative morbidity. **Methods:** Data on 4567 patients who underwent isolated CABG surgery over a 10-year period were extracted from our clinical database. Five postoperative major morbidity complications (cerebrovascular accident, mediastinitis, acute renal failure, cardiovascular failure and respiratory failure) were analysed. A composite morbidity outcome (presence of two or more major morbidities) was also analysed. For each one of these endpoints a risk model was developed and validated by logistic regression and bootstrap analysis. Discrimination and calibration were assessed using the under the receiver operating characteristic (ROC) curve area and the Hosmer—Lemeshow (H—L) test, respectively. **Results:** Hospital mortality and major composite morbidity were 1.0% and 9.0%, respectively. Specific major morbidity rates were: cerebrovascular accident (2.5%), mediastinitis (1.2%), acute renal failure (5.6%), cardiovascular failure (5.6%) and respiratory failure (0.9%). The risk models developed have acceptable discriminatory power (under the ROC curve area for cerebrovascular accident [0.715], mediastinitis [0.696], acute renal failure [0.778], cardiovascular failure [0.710], respiratory failure [0.787] and composite morbidity [0.701]). The results of the H—L test showed that these models predict accurately, both on average and across the ranges of patient deciles of risk. **Conclusions:** We developed a set of risk-prediction models that can be used as an instrument to provide information to clinicians and patients about the risk of postoperative major morbidity in our patient population undergoing isolated CABG.

Keywords: Coronary artery bypass surgery; Morbidity; Predictive models; Postoperative risk-adjusted morbidity

1. Introduction

Evaluation of patient outcomes has become increasingly accepted as one important step to assess and improve quality of patient care. Because differences in outcomes may result from disease severity, effectiveness of treatment, or chance [1,2], and most outcome studies are observational rather than randomised, risk adjustment is necessary to account for case-mix. In this context, risk-prediction models play an important role in current cardiac surgical practice where they may be used to assess the impact of specific predictors on outcome, to aid in patient counselling and treatment selection, to profile provider quality, and to serve as the basis for continuous quality improvement [3].

Risk models for mortality after cardiac surgery are widely used, but the application to populations other than those for which they were developed may not be adequate. In a previously published study, we developed and validated a risk model for in-hospital mortality in patients submitted to isolated coronary artery bypass graft (CABG) procedures in our institution [4]. But although operative mortality is obviously the most important clinical endpoint, mortality alone is no longer considered sufficient to assess patient outcomes. It is clearly recognised that other non-fatal postoperative complications can significantly impact patients' functional status and quality of life [5]. Therefore, identification of factors and calculation of risk-adjusted morbidity rates for CABG procedures may provide valuable insights on areas to focus for improved quality of care.

In this work, we aimed at identifying the preoperative risk factors for perioperative morbidity in patients submitted to isolated CABG, using a detailed dataset with information on a large group of patients undergoing this procedure at our institution, and to develop and validate risk-prediction models for the most important postoperative morbidity complications.

2. Materials and methods

2.1. Study design and population

We performed a retrospective cohort study of patients undergoing CABG during a 10-year period, from January 1992

☆ Presented at the 22nd Annual Meeting of the European Association for Cardio-thoracic Surgery, Lisbon, Portugal, September 14—17, 2008.

* Corresponding author. Address: Cirurgia Cardiotorácica, Hospitais da Universidade, 3000 Coimbra, Portugal. Tel.: +351 239400418; fax: +351 239829674.

E-mail address: antunes.cct.huc@sapo.pt (M.J. Antunes).

1010-7940/$ — see front matter
doi:10.1016/j.ejcts.2008.10.046

through December 2001. Only patients undergoing isolated first operations and reoperations were included. Patients undergoing CABG combined with heart valve repair or replacement, resection of a ventricular aneurysm, or other surgical procedures were excluded. After exclusions, the material for this study's risk modelling analyses consisted of 4567 patient records. There were 4030 men (88.2%) and 537 women and the mean age was 60.6 ± 9.2 years (median 62 years). The mean number of grafts per patient was 2.8 ± 0.8 and the mean cardiopulmonary bypass time was 63.3 ± 22.9 min. The left internal mammary artery (IMA) alone and double AMI were used in 99.5% and 23.4% of the study population, respectively. All operations were performed without cardioplegia, under hypothermic ventricular fibrillation or empty beating heart, a technique described in detail in previous reports from our institution [6,7]. The mean hospital stay was 7.9 ± 6.4 days (minimum, 0 days; maximum, 204 days; 25% quartile, 6 days; median, 7 days; 75% quartile, 7 days), and, in general, survivors were discharged to their home. No use was made of aftercare institutions.

2.2. Data collection

Data had been collected prospectively from each patient on a standardised written form by the respective surgeon, and validated and inputted in a computerised database by one of us (PEA). Supervision for data consistency was performed by the project co-ordinator (MJA) and aggregate outputs were periodically cross-checked against an independent clinical database. The data collection instruments included questions regarding demographic characteristics, preoperative risk factors, previous interventions, preoperative cardiac status, cardiac catheterisation results, intraoperative management and postoperative complications.

2.3. Morbidity endpoints

Five major postoperative morbidity complications, considered either life threatening or potentially resulting in permanent functional disability, were selected. They appear to be the most uniformly reported by other investigators: cerebrovascular accident, mediastinitis, acute renal failure, cardiovascular failure and respiratory failure. A composite morbidity outcome (presence of more than one major morbidity) was also considered an endpoint for this study. Definitions of these variables are presented in Appendix B. All morbidity complications were analysed as events occurring during hospital stay, unlimited in time, with exception of mediastinitis which was analysed as a 30-day event.

2.4. Methods of analysis

More than 50 preoperative patient variables were available in the database, of which 23 potential risk factors were chosen, based on univariate screening, clinical knowledge and previous research. The risk factors selected for analysis are listed and defined in Appendix C and include: age, gender, body mass index, body surface area, hypertension, diabetes mellitus, recent smoking, peripheral vascular disease, cerebrovascular disease, renal failure, serum creatinine level, anaemia, chronic obstructive pulmonary disease (COPD), cardiomegaly, recent myocardial infarction, unstable angina, Canadian Cardiovascular Society (CCS) class II/III, non-elective surgery, previous cardiac surgery, left ventricular dysfunction, left main disease, three-vessel disease and need for intra-aortic balloon pumping.

The entire database was initially used to develop the predictive logistic models. Univariate screening of all model-eligible risk factors was performed using unpaired Student's t test or the Mann–Whitney test for numeric variables, and the χ^2 test or Fisher's exact test for categorical variables. Multicollinearity among variables was obviated by using only one of a set of variables with a correlation coefficient greater than 0.5 in the regression analysis. Variables with a p value lower than 0.2 by univariate analysis were used as independent variables for further analysis.

A multivariate stepwise logistic regression analysis was then performed for each of the six dependent morbidity groups: (1) cerebrovascular accident (2) mediastinitis (3) acute renal failure (4) respiratory failure (5) cardiovascular failure and (6) the composite morbidity. Because of the relatively small effective sample size of some events (respiratory failure and mediastinitis), a p value less than 0.1 was selected for variable retention in the final regression model. Bootstrap analysis was used in combination with logistic regression analysis to select the final set of risk factors included in the model. In the bootstrap procedure, 200 samples of 4567 patients (the same number of observations as the original database) were sampled with replacement. A stepwise logistic regression analysis was applied to every bootstrap sample. If the predictors occurred in more than 50% of the bootstrap models, they were judged to be reliable and were retained in the final model. Unreliable variables, if present, were removed from the final model. Thus, the risk factors were not only identified as statistically significant by traditional analysis, but also occurred the most frequently in the bootstrap analysis. The tables of risk factors include frequency of occurrence from multivariable bootstrap modelling, as well as conventional magnitude and certainty of association.

Finally, we validated the risk-prediction model internally by randomly drawing 200 samples, each containing 100% of the total number of subjects. The risk-prediction model was applied to each sample to calculate an individual sample area under the ROC curve and then the mean and standard error of the mean, with 95% confidence intervals (95% CI), for all 200 ROC values.

2.5. Model performance

Two different properties were used to evaluate the predictive accuracy of the models: calibration and discrimination. Calibration, which measures the ability of the model to assign the appropriate risk, was evaluated by the Hosmer–Lemeshow (H–L) goodness-of-fit method. The H–L χ^2 statistics measures the differences between expected and observed outcomes over deciles of risk. A statistically non-significant result (p value ≥ 0.05) suggests that the model predicts accurately on average [8]. In order to get more insight into the model performance across the ranges of patient deciles of risk, we have plotted the observed and expected events rates in these risk groups. Accurate

Table 1
Demographic and clinical characteristics ($n = 4567$).

Risk factors[a]	
Mean age (years)	60.7 ± 9.3
Mean body mass index (kg/m^2)	26.1 ± 2.2
Mean body surface area (cm^2)	177.9 ± 14.0
Female gender	11.8
Diabetes mellitus	22.6
Hypertension	57.0
Recent smoking	11.5
Peripheral vascular disease	10.3
Cerebrovascular disease	5.2
Anaemia	3.9
Renal failure	2.1
Chronic pulmonary disease	3.3
Cardiomegaly	11.6
Recent myocardial infarct	5.1
Unstable angina	6.8
Angina CCS class III or IV	40.0
Previous cardiac surgery	1.7
Left main disease	16.5
Non-elective surgery	6.6
Left ventricular dysfunction	13.3
Three-vessel disease	75.3
Intra-aortic balloon pump	0.6

CCS, Canadian Cardiovascular Society.
[a] Values are expressed in percentage, unless specified otherwise.

predictions of events within each of these risk groups would suggest that the risk model is suitable for patient advice for all (low- to high-risk) patients.

Discrimination, which measures the model's ability to differentiate among those who have or have not suffered an event, was evaluated by analysis of the area under the ROC curve [9]. If the area is greater than 0.7, it can be concluded that the model has an acceptable discriminatory power and, consequently, may be used to rank patients into treatment groups to facilitate management [10].

3. Results

3.1. Risk profile for study population and outcomes

This generally male CABG population (11.8% female) was predominantly noted to have triple-vessel disease (75.3%). The risk profile for the CABG study population is shown in Table 1.

The study population had hospital mortality and major composite morbidity rates of 1.0% and 9.0%, respectively. The specific major morbidity rates were: cerebrovascular accident, 2.5%; mediastinitis, 1.2%; acute renal failure, 5.6%; cardiovascular failure 5.6%; and respiratory failure, 0.9%. The composite morbidity rate was 9.0%.

Table 2
Logistic regression risk models for CABG postoperative major morbidity. (A) Cerebrovascular accident, (B) mediastinitis, (C) renal failure, (D) respiratory failure, (E) cardiovascular failure and (F) composite morbidity.

Model	Risk factor	Coefficient	p value	Bootstrap frequency (%)	Odds ratio	95% CI (OR)	
A	Age (per one year increase)	0.033	0.004	85.4	1.034	1.011	1.057
	Female sex	0.575	0.020	67.5	1.778	1.096	2.884
	Peripheral vascular disease	0.919	<0.001	90.2	2.507	1.603	3.921
	Cerebrovascular disease	1.048	<0.001	92.2	2.851	1.683	4.832
	LV dysfunction	0.677	0.003	65.5	1.969	1.259	3.080
	Non-elective surgery	0.641	0.027	60.2	1.899	1.077	3.347
	Constant	−6.238					
B	BSA (per each cm^2)	0.029	0.003	91.2	1.029	1.010	1.048
	Recent smoking	0.813	0.012	85.3	2.255	1.194	4.259
	Diabetes	0.597	0.037	58.4	1.817	1.035	3.189
	Cardiomegaly	0.741	0.022	53.2	2.099	1.115	3.951
	Constant	−9.980					
C	Age (per one year increase)	0.210	0.05	60.5	1.022	1.000	1.044
	Serum creatinine (per 0.1 mg/dl increase)	2.477	<0.001	98.5	11.908	7.409	19.139
	Constant	−6.915					
D	Age(per one year increase)	0.050	0.013	80.4	1.051	1.011	1.093
	Recent smoking	1.187	0.002	75.8	3.278	1.536	6.995
	Peripheral vascular disease	1.206	0.001	60.9	3.339	1.662	6.708
	COPD	2.042	<0.001	57.7	7.704	3.519	16.864
	Anaemia	1.210	0.015	65.7	3.355	1.263	8.908
	Constant	−8.605					
E	Angina CCS class III or IV	0.279	0.038	75.6	1.322	1.016	1.722
	Previous cardiac surgery	0.967	0.005	52.1	2.629	1.332	5.189
	LV dysfunction	0.377	0.025	68.5	1.458	1.048	2.029
	Non-elective surgery	0.659	0.001	63.0	1.933	1.291	2.896
	Constant	−3.083					
F	Peripheral vascular disease	0.458	0.002	69.5	1.581	1.183	2.113
	Cerebrovascular disease	0.399	0.047	72.3	1.491	1.004	2.212
	Angina CCS class III or IV	0.292	0.007	70.5	1.340	1.083	1.657
	Previous cardiac surgery	0.630	0.050	52.8	1.878	.998	3.535
	COPD	0.813	<0.001	63.2	2.254	1.469	3.460
	LV dysfunction	0.398	0.003	68.7	1.489	1.140	1.945
	Non-elective surgery	0.493	0.005	56.5	1.638	1.157	2.319
	Constant	−2.680					

BSA, body surface area; CCS, Canadian Cardiovascular Society; COPD, chronic obstructive pulmonary disease; LV, left ventricle.

3.2. Model results

Six different risk models were developed, one for each of the five morbid events and one for the composite morbidity, and the final logistic model results are presented in Table 2. This table summarises, for each one of the constructed models, the variables used and their frequency of occurrence (%) in bootstrap analyses, regression coefficients, odds ratio and associated *p* values. All the risk factors included in each model occurred in more than 50% of the bootstrap samples, indicating reliability.

The various preoperative risk factors influenced differently each of the five morbid events. The variables found to have the most significant impact in the occurrence of morbidity were: for *cerebrovascular accident*: age (OR = 1.034 per one year increase), female sex (OR = 1.778), peripheral vascular disease (OR = 2.507), cerebrovascular disease (OR = 2.851), LV dysfunction (OR = 1.969) and non-elective surgery (OR = 1.899); for *mediastinitis*: BSA (OR = 1.029 per each cm^2 increase), recent smoking (OR = 2.255), diabetes (OR = 1.817) and cardiomegaly (OR = 2.099); for *acute renal failure*: age (OR = 1.022 per one year increase) and serum creatinine (OR = 11.908 per 0.1 mg/dl increase); for *respiratory failure* were: age (OR = 1.051 per one year increase), recent smoking (OR = 3.278), peripheral vascular disease (OR = 3.339), COPD (OR = 7.704) and anaemia (OR = 3.355); and for *cardiovascular failure*: angina CCS class III/IV (OR = 1.322), previous cardiac surgery (OR = 2.629), LV dysfunction (OR = 1.458) and non-elective surgery (OR = 1.933).

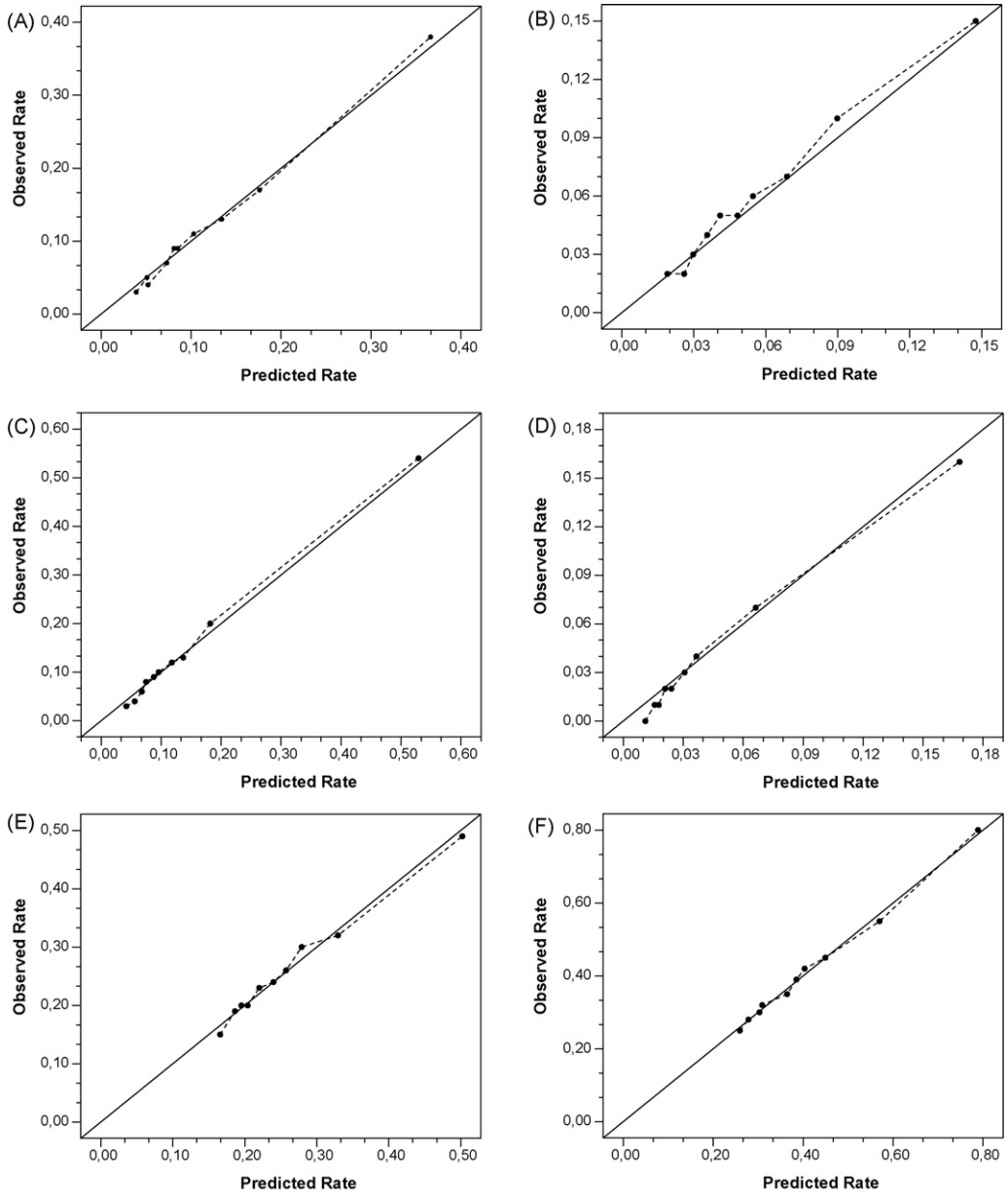

Fig. 1. Model calibrations (A) cerebrovascular accident (B) mediastinitis (C) renal failure (D) respiratory failure (E) cardiovascular failure and (F) composite morbidity.

Table 3
Risk model performance metrics.

	AUC (95% CI)	H–L test	H–L test *p* Value
Final cerebrovascular accident model	0.715 (0.661, 0.749)	6.084	0.638
Final mediastinitis model	0.696 (0.637, 0.745)	24.9	0.401
Final renal failure model	0.778 (0.738, 0.818)	11.692	0.165
Final respiratory failure model	0.787 (0.713, 0.861)	6.220	0.623
Final cardiovascular failure model	0.710 (0.683, 0.740)	6.14	0.979
Final composite morbidity model	0.701 (0.688, 0.721)	11.110	0.196

H–L, Hosmer–Lemeshow.

In the final composite morbidity outcome, peripheral and cerebral vascular disease, age, angina CCS class III/IV, previous cardiac surgery, COPD, LV dysfunction and non-elective surgery were the variables which were found to have the most impact.

For all the risk models, the H–L goodness-of-fit test was not statistically significant ($p \geq 0.05$). Fig. 1A–F demonstrates the calibration of the models, i.e., how well the predicted event rates match the observed rates among patient decile risk groups. As noted by the close agreement between these results, these models appear to be relatively accurate across the ranges of patient risk subgroups. These results indicate that the models predict accurately, both on average and across the ranges of patient deciles of risk, hence it is suitable for use in all (low- to high-risk) patients. The risk models developed demonstrate an acceptable discriminatory power (area under the ROC curve for cerebrovascular accident [0.715], mediastinitis [0.696], acute renal failure [0.778], cardiovascular failure [0.710], respiratory failure [0.787] and composite morbidity outcome [0.701]). The details of model performance metrics are provided in Table 3.

4. Discussion

For many years, operative mortality was the sole criterion used for evaluation of patient outcomes, and many published studies have analysed the mortality of cardiac operations, but studies concentrated on the analysis of perioperative morbidity and its influence on global early and late results are much fewer. Although operative mortality is obviously the most deleterious clinical endpoint when analysing surgical results, this is no longer considered sufficient to assess surgical/patient outcomes. It is clearly recognised that other non-fatal postoperative complications can significantly impact not only the perioperative period but also the patient's quality of life, and may often constitute serious threats to the longer-term survival, functional capability, and overall well-being after CABG or any other cardiac surgical procedure. Therefore, identification of risk factors for increased perioperative morbidity and analysis of risk-adjusted morbidity for CABG procedures may provide valuable information which may subsequently be used to improve quality of care.

The Society of Thoracic Surgeons (STS) Quality Measurement Task Force has recently identified a group of measures to serve as the basis for comprehensive assessment of the quality of adult cardiac surgery [5]. Eleven individual measures of coronary artery bypass grafting quality were selected within four domains: (1) perioperative medical care, (2) operative care, (3) risk-adjusted operative mortality and (4) postoperative risk-adjusted major morbidity, the latter defined as the risk-adjusted occurrence of any of the following: renal failure, deep sternal wound infection, re-exploration, stroke, or prolonged ventilation/intubation. We could identify only one other study, published by Shroyer and colleagues, conducted to develop separate risk-prediction models for major morbidity events [11]. Using part of the large national experience captured in the STS database, these authors examined five postoperative CABG complications (stroke, renal failure, reoperation within 24 h after CABG, prolonged postoperative ventilation, and mediastinitis) and, for each of these, developed risk-prediction models.

However, risk models cannot be uniformly applied to different population groups, as the experience with the Parsonnet, EuroSCORE and the STS models for operative mortality has demonstrated. In a previously published study, we described the development of our own local risk-prediction model for mortality after CABG in our population, which predicted outcomes better than those widely used models [4]. In the current study, complementary outcome measures for risk-adjusted major morbid postoperative complications were developed. Five morbid events, either life threatening or potentially resulting in permanent functional disability, which appeared to be some of the most uniformly reported complications [12–16], were analysed: cerebrovascular accident, mediastinitis, renal failure, respiratory failure and cardiovascular failure. A model for composite morbidity (association of two or more major morbidity events) was also developed. Except for mediastinitis, which was evaluated as a 30-day event, the morbidity complications were analysed as events occurring during hospital stay, unlimited in time. Although it represents one of the most widely reported metrics to assess post-operative complications after CABG, it may be a too short interval for the evaluation of the true early risk. Nevertheless, and in the context of the present study, we believe that the more important issue is the ability to measure and validate it conveniently and accurately.

The main goal of our study was to identify the preoperative risk factors and to develop and validate risk-prediction models which could be used as instruments to provide information to clinicians and patients about the risk of major postoperative morbidity in our patient population undergoing isolated CABG surgery. To ensure the accuracy and usefulness of such model, many factors are essential, including selection of an appropriate clinical database,

inclusion of critical variables, and proper model development and validation. The data used for this study were taken from our clinical database, which was created and is in use since the beginning of the surgical activity in this department, in 1988. In our database, some of the variables selected for analysis (ejection fraction, haematocrit, cardiothoracic ratio) were codified as categorical instead of continuous variables, which may constitute one limitation to the model building process. The risk models were developed by means of logistic regression and bootstrap analyses, which are the techniques most commonly used for risk modelling.

However, the resultant models are useful only if they reliably predict outcomes for patients by determining significant risk factors associated with the particular outcome. A problem might arise from this dependence on risk factor analysis. Different investigators evaluating the same predictors through regression analyses may obtain heterogeneous results because of methodological discrepancies and inadvertent biases introduced in the statistical elaboration [17]. Bootstrap analysis is a simulation method for statistical inference, which, if applied to regression analysis, can provide variables that have a high degree of reproducibility and reliability as risk-factors for a given outcome. We also used bootstrap analysis to internally validate the model. This methodology was recently proposed as a breakthrough method for internal validation of surgical regression models [18]. The main advantage of this technique is that the entire dataset can be used for building a more robust model, especially in moderate-size databases and for rare outcomes [19].

However, as discussed above, the risk model created may not be used in other patient populations. Local models need to be developed for use in each particular population.

Overall, the set of risk models developed in this study performed acceptably. The Hosmer—Lemeshow test was not statistically significant for all the models and demonstrated good calibration across the ranges of patient risk subgroups. These results indicate that the models predict accurately, both on average and across the ranges of patient deciles of risk, hence, are suitable for use in all (low- to high-risk) patients. The risk models developed also demonstrate an acceptable discriminatory power and, consequently, may be used to rank patients into treatment groups to facilitate management (under the ROC curve areas $\geq$0.7). However, the demographic and clinical characteristics presented in Table 1 may place the study population in a low-risk profile, which means that any inference of the models, namely the validity of the stability over the spectrum of risk, must be reduced to the centre where it was developed, possibly limiting the applicability to other centres.

The recognition of specific risk factors for each particular morbid event and for global morbidity permits preoperative modulation of these factors, aimed at decreasing perioperative morbidity. As our knowledge of these correlations in our specific conditions increased, we have applied measures to improve on preoperative risk factors and observed a significant decrease of morbidity rates in our patient population, which may have been a form of practical validation of the model created but have also contributed to decrease the level of morbidity.

This may be one of the weaknesses of this study. As it evolved, it influenced its own results. One other weakness is that formal external auditing was not available for our data. However, systematic supervision and periodical cross-checking, as described, captured (and corrected) a number of erroneous data inputs, however small enough to render the whole dataset satisfactorily reliable. Additionally, although we have performed a prospectively designed study, the data were collected from a single-centre database, which carries the risk that any inference may be reduced to the centre where it was developed, possibly limiting the applicability to others, as discussed above.

In conclusion, we developed a set of risk-prediction models that can be used as instruments to provide information to clinicians and patients about the risk of postoperative major morbidity in our patient population awaiting CABG surgery. It has also served to apply corrective measures which permitted improvement of our own results. Naturally, it is for our own use and is not intended for use in other patient populations. But it may facilitate comparisons with the results of other centres. Other groups are encouraged to create their own models using the methodologies applied here.

References

[1] Daley J. Criteria by which to evaluate risk-adjusted outcomes programs in cardiac surgery. Ann Thorac Surg 1994;58:1827—35.
[2] Tu JV, Sykora K, Naylor CD. Assessing the outcomes of coronary artery bypass graft surgery how many risk factors are enough? Steering Committee of the Cardiac Care Network of Ontario. J Am Coll Cardiol 1997;30:1317—23.
[3] Ferraris VA, Ferraris SP. Risk stratification and comorbidity. In: Cohn LH, Edmunds Jr LH, editors. Cardiac surgery in the adult. 3rd ed., McGraw-Hill: New York; 2008. p. 199—246.
[4] Antunes PE, Eugénio L, Ferrão de Oliveira J, Antunes MJ. Mortality risk prediction in coronary surgery: a locally developed model outperforms external risk models. Interact Cardiovasc Thorac Surg 2007;6:437—41.
[5] Shahian DM, Edwards FH, Ferraris VA, Haan CK, Rich JB, Normand SL, DeLong ER, O'Brien SM, Shewan CM, Dokholyan RS, Peterson ED. Society of Thoracic Surgeons Quality Measurement Task Force. Quality measurement in adult cardiac surgery: part 1 — conceptual framework and measure selection. Ann Thorac Surg 2007;83:S3—12.
[6] Antunes PE, Oliveira JF, Antunes MJ. Non-cardioplegic coronary surgery in patients with severe left ventricular dysfunction. Eur J Cardiothorac Surg 1999;16:331—6.
[7] Antunes MJ, Bernardo JB, Oliveira JM, Fernandes LE, Andrade CM. Coronary artery bypass surgery with intermittent aortic cross-clamping. Eur J Cardiothorac Surg 1992;6:189—94.
[8] Hosmer DW, Lemeshow S. Applied logistic regression, 2nd ed., New York: John Wiley and Sons; 2000.
[9] Grunkemeier GL, Ruyun J. Receiver operating characteristic curve analysis of clinical risk models. Ann Thorac Surg 2001;72:323—6.
[10] Omar RZ, Ambler G, Royston P, Eliahoo J, Taylor KM. Cardiac surgery risk modelling for mortality: a review of current practice and suggestions for improvement. Ann Thorac Surg 2004;77:2232—7.
[11] Shroyer AL, Coombs LP, Peterson ED, Eiken MC, De Long ER, Chen A, Ferguson Jr TB, Grover FL, Edwards FH. The Society of Thoracic Surgeons: 30-day operative mortality and morbidity risk models. Ann Thorac Surg 2003;75:1856—64.
[12] Antunes PE, Prieto D, Ferrão de Oliveira J, Antunes MJ. Renal dysfunction after myocardial revascularization. Eur J Cardiothorac Surg 2004;25:597—604.
[13] Simon C, Luciani R, Capuano F, Miceli A, Roscitano A, Tonelli E, Sinatra R. Mild and moderate renal dysfunction: impact on short-term outcome. Eur J Cardiothorac Surg 2007;32:286—90.
[14] Antunes PE, Ferrão de Oliveira J, Antunes MJ. Predictors of cerebrovascular events in patients subjected to isolated coronary surgery. The

importance of aortic cross-clamping. Eur J Cardiothorac Surg 2003;23:328–33.
[15] Magovern JA, Sakert T, Benckart DH, Burkholder JA, Liebler GA, Magovern Sr GJ, Magovern Jr GJ. A model for predicting transfusion after coronary artery bypass grafting. Ann Thorac Surg 1996;61:27–32.
[16] Baker RA, Hallsworth LJ, Knight JL. Stroke after coronary artery bypass grafting. Ann Thorac Surg 2005;80:1746–50.
[17] Naftel DC. Do different investigators sometimes produce different multivariable equations from the same data? J Thorac Cardiovasc Surg 1994;107:1528–9.
[18] Blackstone EH. Breaking down barriers. Helpful breakthrough statistical methods you need to understand better. J Thorac Cardiovasc Surg 2001;122:430–9.
[19] Grunkemeier GL, Wu Y. Bootstrap resampling methods. Something for nothing? Ann Thorac Surg 2004;77:1142–4.

Appendix A. Conference discussion

Dr P. Kappetein (*Rotterdam, Netherlands*): Obviously you have done an enormous amount of work. You have a huge database with patients and several risk factors that you took into account to see whether these were predictors for the complications that you mentioned.

The point is that the more variables you have and the more end points you consider, the more likely that you will find something that is statistically significant.

For example, left ventricular dysfunction was a predictor for left ventricular failure, and, of course, that is consistent with other studies.

These risk models become more and more complex. A good example is the paper from Mr Lee which was published in JAMA. He showed that with 14 different risk factors, you can estimate the likelihood whether an individual dies within 4.2 years. The model shows a C statistic of 0.83. When you only consider age and apply this to a population of people between 50 and 80 years, the C statistic is 0.79. So all these other 13 variables add only 4% to the accuracy of the model.

Wouldn't it just be enough to look at age alone to predict complication rates, or do you still think that there is some kind of value in looking at all different risk factors?

Dr Antunes: As I mentioned in the data analysis, we performed a stepwise forward regression analysis and, in conjunction, we applied the bootstrap method. The risk factors included in those models appear in more than 50% of the bootstrap samples, so there is some reliability in those factors.

If we look at the model of renal failure, there are only two variables, age and serum creatinine level. But the relative contribution of the serum creatinine level accounts for more than 95% of the model.

So I'm concluding that the risk factors that were included in the model were carefully chosen.

Dr Kappetein: Okay.

Dr Antunes: I come as a coauthor of this paper just to clarify this. There were initially 50 variables. Univariate analysis isolated 23. Those ones underwent multivariate analysis, and this isolated the significant factors for each particular outcome. They came out as the statistically significant variables for the development of this model.

I just also should draw the attention to the fact that we do not intend to propose this as a model for everybody. It has to be validated with other populations. But it is probably an encouragement for each institution to develop its own model because each one of us has a different population, and it is important for evaluation of the quality of our own work that we keep track of our own results. That is the intention of this model, nothing else.

Dr B. Buxton (*Melbourne, Australia*): Could you translate the data that you presented to daily use, say, in the ICU? Could it be a practical model? For instance, could you predict the outcome of respiratory failure in certain patients before the surgery?

Dr Antunes: Yes, sure. That is exactly the aim, that we can tell our own patient by analysing his variables, his age, his preoperative serum creatinine, that the chance of this patient having a mediastinitis is, for example, X. If this X is too high, we may think that if we manage to preoperatively drop his creatinine level, then we lower the risk of having a mediastinitis, and that's very important. It is important for information of the patient, of the physician who refers the patients, and for our own information.

As Dr Pedro Antunes developed this work over 10 years, we incorporated this knowledge into our own practice and saw our own incidence of morbidity coming down quite dramatically.

So the learning curve is not just the technical improvement but also this ability to modulate the risk factors of the patient.

I also draw attention to, as far as we know, there being only one work, and that's from the STS database, which is specifically developed for predicting morbidity. All the risk models that you know are worked out for mortality. This one is specifically for morbidity. And as the mortality has come down, it is becoming more and more important for us to also analyse morbidity. In fact, a patient that dies is probably not a problem. Unfortunately for him, the problems have ended. When you develop a specific morbidity, it may be a problem for a long time.

Dr Buxton: Are there any other comments?

Dr Kappetein: Yes, I have some comments. I think you will agree that the more variables you take into account, the more end points, the more likely you will find a factor.

And, for example, left ventricular dysfunction is of course predictive for cardiovascular failure like dying is predicted for mortality.

And so I think it's always very important to first select the kind of variables that you want to put in your model and then run the model.

Dr Antunes: Yes, but in the composite morbidity model that we have here, which is the most important, finally you only have to fit in seven variables. I don't know of any risk model that has less than seven variables.

Dr Buxton: Perhaps we could continue this discussion outside. Any other comments before we close?

The thing that is common to all of these papers is that they're quite complicated. There is a lot of mathematic modelling, and there are a large number of data. And I wonder whether as surgeons we should take a refresher course in the bootstrapping and the various logistic models that we need to run our practices.

Appendix B. Definition of the morbidity endpoints

Cerebrovascular accident	Global or focal neurological deficit lasting less (transient ischemic attack) or more than 24 h (reversible ischemic neurological deficit; stroke)
Mediastinitis	At least one of the following: (1) an organism isolated from culture of mediastinal tissue or fluid; (2) evidence of mediastinitis seen during operation; (3) one of the following conditions: chest pain, sternal instability, or fever ($>38\ ^{\circ}C$), in combination with either purulent discharge from the mediastinum or an organism isolated from blood culture or culture of mediastinal drainage
Acute renal failure	At least one of the following: (1) a postoperative serum creatinine (Scr) level $\geq$2.1 mg/dl plus an increase in the Scr level $\geq$0.9 mg/dl from preoperative to maximum postoperative values if preoperative Scr $\leq$2.0 mg/dl; (2) an increase in the Scr level $\geq$1.5 mg/dl if baseline Scr $\geq$2.0 mg/dl; (3) a new requirement for dialysis
Respiratory failure	Indicate whether the patient had pulmonary insufficiency requiring postoperative ventilator support for $>$48 h or tracheostomy, or both
Cardiovascular failure	At least one of the following: acute myocardial infarction; intra-aortic balloon pump; ventricular assist device; cardiac arrest due to sustained ventricular fibrillation/tachycardia; inotropic drugs for $>$24 h

Appendix C. Definition of preoperative variables

Diabetes	History of diabetes treated with oral agents or insulin
Hypertension	Blood pressure exceeding 140/90 mmHg, or a history of high blood pressure, or the need of antihypertensive medications
Renal failure	Creatinine >2.0 mg/dl and no dialysis dependency
Recent smoking	Up to less than four weeks of surgery
Anaemia	Haematocrit ≤34%
Cardiomegaly	Cardiothoracic ratio >0.50 on a chest X-ray-film
Chronic pulmonary obstructive disease	Patient requires pharmacologic therapy for the treatment of chronic pulmonary compromise, or patient has a FEV1 <75% of predicted value
Peripheral vascular disease	Claudication, either with exertion or at rest; amputation for arterial insufficiency; aorto-iliac occlusive disease reconstruction; peripheral vascular bypass surgery, angioplasty or stent; documented abdominal aorta aneurism, repair or stent; or non-invasive carotid test with >75% occlusion
Cerebrovascular disease	Unresponsive coma >24 h, CVA, RIND or TIA
Recent myocardial infarction	<30 days
Unstable angina	Preoperative use of iv nitrates until arrival in the anaesthetic room
Previous cardiac surgery	Previous surgery requiring opening of the pericardium
Left ventricular dysfunction	Ejection fraction <40%
Non-elective surgery	Urgent or emergent surgery
Intra-aortic balloon pump	Preoperative intra-aortic balloon pump for haemodynamic reasons

Editorial comment

Predicting morbidity after coronary surgery

Keywords: Cardiac surgery; Morbidity; Risk model

Antunes and co-workers [1] present a risk model for the prediction of major morbidity (stroke, mediastinitis and organ failure) after coronary surgery. The specialty of cardiac surgery has led the field of risk assessment in relation to operative mortality and those interested in this field have at their service a number of risk models which allow the prediction of mortality with a reasonable degree of accuracy. The authors correctly state that evaluation of clinical outcomes should no longer be confined to operative mortality, which is now closely monitored in most units with robust quality control mechanisms. Other outcomes of interest to the patient, the provider and the purchaser of health care are those of major morbidity and those that deal with long-term survival and quality of life. For such outcomes to be useful in quality monitoring, they should be risk-adjusted. This paper tackles the subject of one group of outcomes, namely major morbidity, to see if it can be predicted by risk models.

The models were developed using data from over 4500 patients who underwent coronary bypass grafting. The endpoints were well defined and logistic regression was used to identify the risk factors that were then used to compile five individual risk models, one for each of the major morbidities studied and one composite risk model for the development of any of the major morbidities. The factors identified are intuitively appropriate and the models have generally acceptable discriminatory power, with the renal and respiratory failure models performing better than the other three models and better than the composite model.

That morbidity risk prediction is useful is beyond doubt. Two questions remain to be addressed. The first concerns the method by which such predictions are made, and the second the extent of usefulness of such predictions.

In this paper, the authors have recommended different models for different complications. The attraction of this approach is obvious: it focuses the selection of the appropriate risk factors for the outcome of interest (for example, serum creatinine is only a predictor for renal failure, whereas peripheral vascular disease predicts both respiratory failure and stroke and age predicts stroke, respiratory and renal failure). The disadvantage of such a method is the potential proliferation of risk models that may be difficult to use at the bedside. The authors have constructed *de novo* morbidity-specific models, whereas other workers have adapted, supplemented or simply transposed mortality models for the prediction of morbidity with reasonable success [2–6]. The attraction of this approach is in its simplicity, as most surgeons are familiar with at least one mortality model, but simplicity may be counterbalanced by the loss of specificity.

Finally, although the stated aim of the paper is to provide tools for assessing the quality of care, an additional potential benefit of progress in this field is the attractive possibility of providing tools for improving the efficiency of service and resource usage. Major morbidity is costly in its use of scarce resources such as critical care beds, dialysis facilities and rehabilitation services. Accurate and discriminating risk