A Simple Predictive Model of Prolonged Intensive Care Unit Stay after Surgery for Acquired Heart Valve Disease

Jianping Xu*, Yipeng Ge*, Shengshou Hu, Yunhu Song, Hansong Sun, Ping Liu

Cardiovascular Institute and Fuwai Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Bei Jing, China

Background and aim of the study: **The study aim was to construct a simple model (the Fuwai risk score) to predict prolonged intensive care unit (ICU) stay after surgery to treat acquired heart valve disease.**
Methods: **Data were collected retrospectively from 2,218 consecutive patients who underwent surgery for acquired heart valve disease. Prolonged ICU stay was defined as ≥5 days. A simple logistic score was calculated using the logistic coefficient, and the additive score by odds ratio. The Fuwai risk score, EuroSCORE and Parsonnet score were applied to predict a prolonged ICU stay and mortality. A C statistic (receiver operating characteristic curve) was used to test discrimination of the models. Calibration was assessed with a Hosmer-Lemeshow goodness-of-fit statistic.**
Results: **The simple logistic model of the Fuwai risk score showed very good discriminatory ability (C statistic 0.76) and calibration (Hosmer-Lemeshow, p = 0.25) in predicting prolonged ICU stay, while the additive algorithm had good discriminatory ability (C statistic 0.75) but poor calibration (p <0.001). The additive algorithm greatly underestimated the risk for high-risk patients. The Fuwai risk score showed good discriminatory ability, but poor calibration in predicting mortality. Neither the EuroSCORE nor Parsonnet score was superior to the Fuwai risk score.**
Conclusion: **The logistic algorithm of the Fuwai risk score is a simple, objective, convenient and accurate scoring system which may be used to predict prolonged ICU stay after surgery to treat acquired heart valve disease.**

The Journal of Heart Valve Disease 2007;16:109-115

The costs of valvular surgery can vary enormously between patients with an uncomplicated recovery and those who suffer severe postoperative complications. Today, whilst the costs of intensive care unit (ICU) stay are extremely high, prolonged ICU stay may also lead to a shortage of ICU beds and result in the cancellation of operations. The ability to predict prolonged ICU stay would facilitate the decisions to operate, to allocate resources, and to estimate costs (1). In addition, long-term survival analyses have demonstrated a significantly lower survival among patients with longer ICU stay (2). The longer-term survivors had a poor functional state and a reduced quality of life (3,4). Faced with these problems, a simple, objective and accurate score system capable of predicting prolonged ICU stay after cardiac surgery would be a major asset. Hence, the aims of the present study were to: (i) develop such a risk score system; and (ii) evaluate the validation of the EuroSCORE and Parsonnet score among Chinese patients with acquired heart valve disease.

*Contributed equally to this work.

Address for correspondence:
Jianping Xu, Adult Cardiac Surgery Center, Fuwai Hospital, Bei Li Shi Lu Number 167, Bei Jing 100037, China
e-mail: cardsurgfw@yahoo.com

Clinical material and methods

Patients

Between January 2004 and January 2006, data from 2,218 consecutive patients operated on at Fuwai Hospital, Beijing, China for acquired heart valve disease were collected retrospectively. Patients with concomitant thoracic aortic surgery were excluded. The study was approved by the Ethics Committee of the hospital.

All data were entered into a computerized database. In order to ensure the highest possible quality of data entry, all data were entered twice independently by two operators, and any discrepancies checked and corrected. The EuroSCORE and Parsonnet score were each calculated according to previously published reports (5-7).

Criteria for discharge from the ICU

Criteria for ICU discharge included cardiovascular stability, no need for respiratory assistance, evidence of adequate renal function with normal serum electrolyte levels, and evidence of adequate neuropsychological function. At this point the need for intensive monitoring was no longer necessary, and the patient could be cared for in a general unit.

Statistical analysis

All tests were two-sided and significant if the p-value was <0.05. The variables were analyzed using Student's *t*-test for normally distributed continuous variables, and a chi-square test and Fisher's exact probability test (where appropriate) for categorical variables. Factors significantly associated with a prolonged ICU stay by univariate analysis were entered into a logistic regression.

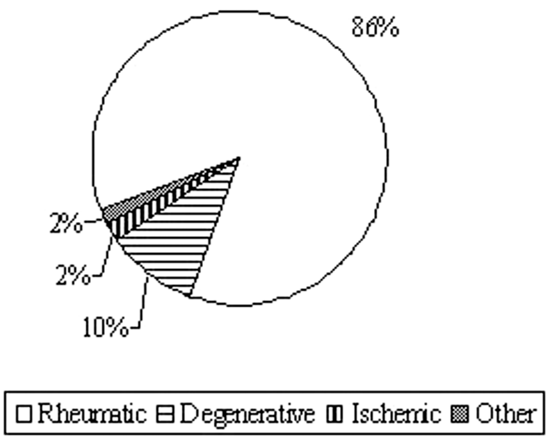

Figure 1: Causes of valve dysfunction.

Table I: Main clinical characteristics and perioperative data of patients (n = 2,193).

Variable	ICU ≥5 days (n = 345)	ICU <5 days (n = 1,848)	p-value
Preoperative factors			
Age (years)*	53.29 ± 11.02	48.47 ± 10.93	<0.001
Male gender	177 (51.3)	856 (46.32)	0.09
Body mass index (kg/m²)*	22.73 ± 3.42	23.28 ± 3.37	0.005
NYHA class IV	32 (9.28)	73 (3.95)	<0.001
Smoking history	84 (24.35)	330 (17.86)	0.005
No sinus rhythm	209 (60.58)	992 (53.68)	0.02
Diabetes	15 (4.35)	60 (3.25)	0.30
Renal dysfunction	1 (0.29)	11 (0.59)	0.70
Hypertension	80 (23.19)	369 (19.97)	0.17
History of stroke	24 (6.96)	70 (3.79)	0.008
History of MI	14 (4.06)	21 (1.14)	<0.001
Ischemic MR	15 (4.35)	17 (0.92)	<0.001
Previous cardiac surgery	42 (12.17)	105 (5.68)	<0.001
Active endocarditis	1 (0.29)	19 (1.03)	0.35
Extracardiac arteriopathy	51 (14.78)	112 (6.06)	<0.001
Pulmonary dysfunction	118 (34.20)	272 (14.72)	<0.001
Pulmonary hypertension	66 (19.13)	225 (12.18)	0.001
Mean aortic gradient ≥120 mmHg	8 (2.32)	20 (1.08)	0.07
LVEF (%)*	57.93 ± 10.12	60.68 ± 8.25	<0.001
CTR*	0.62 ± 0.09	0.57 ± 0.07	<0.001
Serum albumin (g/l)*	40.46 ± 4.65	40.89 ± 4.01	0.11
Hemoglobin (g/l)*	136.46 ± 22.15	140.08 ± 37.79	0.09
White blood cell count ($\times 10^9$/l)*	7.00 ± 7.46	6.72 ± 3.82	0.61
Serum creatinine (µmol/l)*	79.60 ± 17.31	78.84 ± 18.14	0.47
Intraoperative factors			
Combined with CABG	60 (17.39)	120 (6.49)	<0.001
Double valve surgery	135 (39.13)	563 (30.47)	0.002
CPB time (min)*	121.20 ± 53.13	95.93 ± 39.50	<0.001
Aortic cross-clamp time (min)*	84.04 ± 37.19	70.48 ± 33.24	<0.001

*Values are mean ± SD.
Values in parentheses are percentages.
CABG: Coronary artery bypass grafting; CPB: Cardiopulmonary bypass; CTR: Cardiothoracic ratio; LVEF: Left ventricular ejection fraction; MI: Myocardial infarction; MR: Mitral regurgitation.

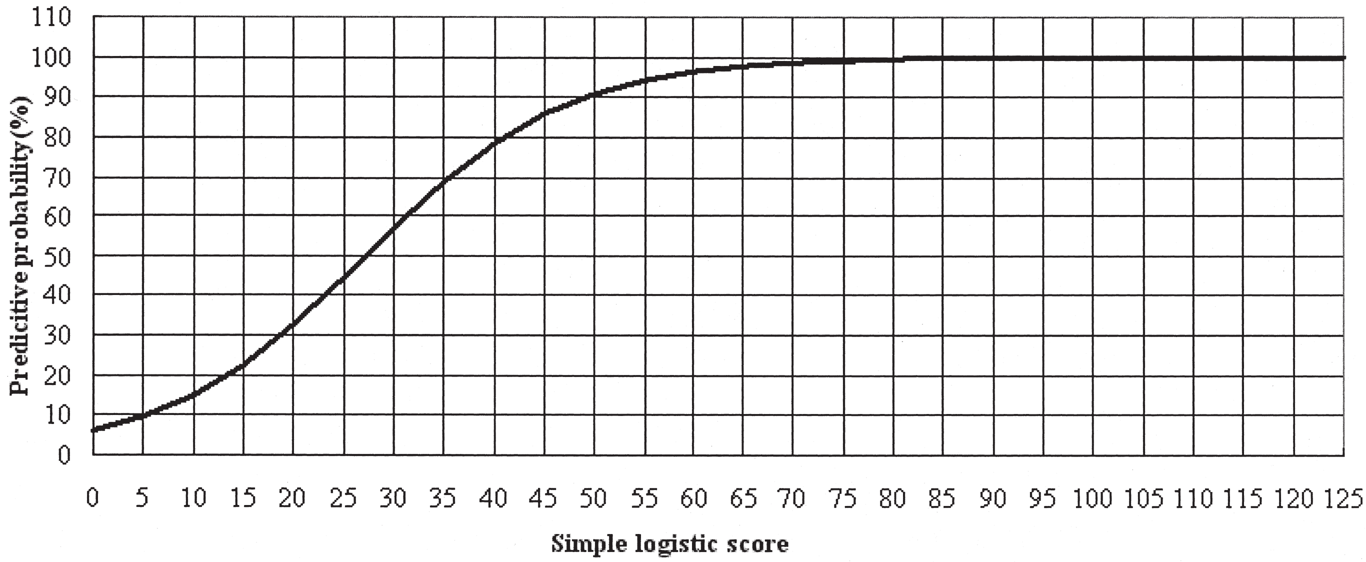

Figure 2: The simple logistic score in relation to predictive probability.

Construction of the simple logistic score

Regression coefficients were used to calculate a prognostic score (S) and a predicted probability (P). The logistic formula was calculated as: $S = B_0 + B_1 \times X_1 + B_2 \times X_2 + \ldots + B_i \times X_i$, where B_0 is the constant of the logistic regression equation and B_i is the coefficient of the variable X_i in the logistic regression, $X_i = 1$ if a categorical risk factor is present, and 0 if it is absent, and $P = e^S/(1 + e^S)$. Because the regression coefficients are always small and have too many decimal places, it is not convenient for the surgeons to calculate S and to predict probability. In the present study, the aim was to construct a simplified logistic model which was simple, accurate and convenient for clinical use. Thus, $B_i \times 10$ was defined as simple logistic score (S_i), and the new formula was calculated as: $S = B_0 + (S_{logistic}/10)$,. A figure was then constructed to relate $S_{logistic}$ and the resulting probability (P). If $S_{logistic}$ is known by the surgeon, the resulting probability can be rapidly and conveniently calculated.

Establishment of the additive model

The odds ratios (OR) were used to construct the additive model.

The discriminatory power of the model was considered excellent if the area under the receiver operating characteristic (ROC) curve was >0.80, very good if >0.75, and good if >0.70 (8). Calibration of the model was assessed with the Hosmer-Lemeshow goodness-of-fit statistic (9); a p-value >0.05 indicated acceptable calibration. The logistic regression was performed using SPSS software (version 13.0). Areas under the ROC curves were compared using MedCalc software (version 8.2.0.3). The algorithm was based on the method developed by Hanley and McNeil (10).

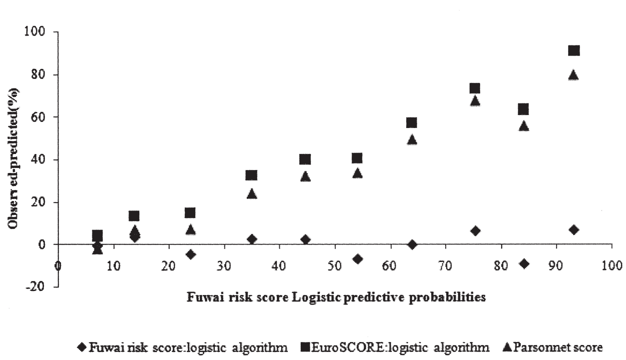

Figure 3: Differences between predicted and observed probabilities by three risk estimation methods.

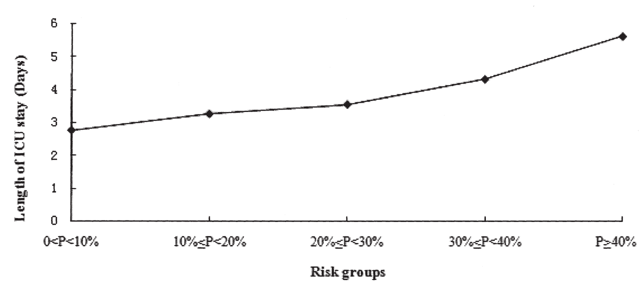

Figure 4: Lengths of ICU stay for different risk groups.

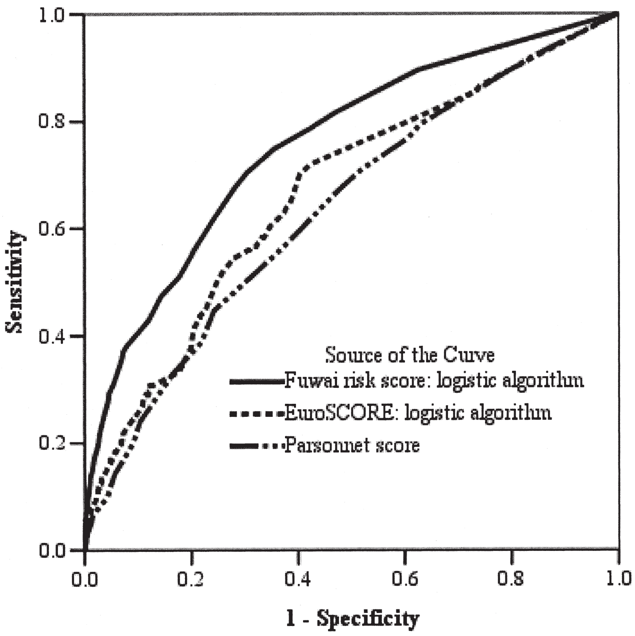

Figure 5: Discriminatory power of three risk models: Prediction of prolonged ICU stay.

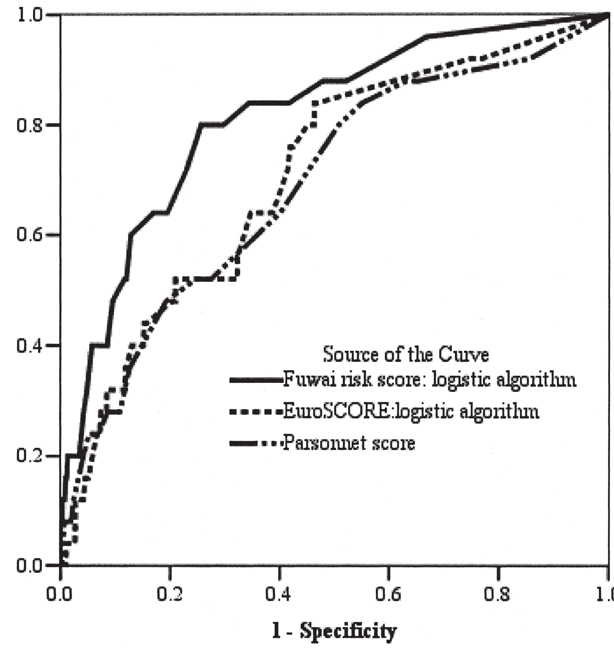

Figure 6: Discriminatory power of three risk models: Prediction of mortality.

Table II: Definitions of important variables.

Variable	Definition
Pulmonary dysfunction	Maximal voluntary ventilation observed/predicted<71%, or vital capacity observed/predicted <71%, or forced expiratory volume in 1 s/forced vital capacity <61% was predicted as preoperative pulmonary dysfunction
Diabetes	Diet-controlled, requiring oral therapy or insulin or had a documented diagnosis of diabetes, regardless of duration of disease
Hypertension	Systolic blood pressure ≥140 mmHg, or requiring antihypertensive medication.
History of stroke	Diagnosis of stroke involves a medical history and a physical examination. Even these patients have no severely sequelae affecting ambulation or day-to-day functioning after the stroke.
History of MI	MI confirmed by angiography.
Previous cardiac surgery	Any previous cardiac surgery requiring opening the pericardium, but excluding surgery during the current hospitalization.
Extracardiac arteriopathy	Any one or more of the following: claudication, carotid occlusion or stenosis, previous or planned intervention on the abdominal aorta, limb arteries or carotids.
Ischemic MR	Ischemic MR (IMR) caused by coronary artery disease. All IMR patients had a previous MI 30 days or more before hospital admission and exhibited functional IMR. Patients with papillary muscle infarction causing papillary muscle rupture or elongation were excluded, as were those with angiographic or echocardiographic findings that demonstrated rheumatic, infectious, or degenerative (myxomatous) mitral disease.
Pulmonary hypertension	Systolic pulmonary artery pressure ≥60 mmHg.
Renal dysfunction	Serum creatinine >200 mmol/l, or established dialysis treatment.
Double valve surgery	Mitral valve surgery combined with concomitant aortic valve surgery, mitral and aortic valve surgery combined with concomitant tricuspid valve plasty, mitral and aortic surgery combined with concomitant coronary artery bypass grafting

MI: Myocardial infarction; MR: Mitral regurgitation.

J Heart Valve Dis
Vol. 16. No. 2
March 2007

Results

An ICU stay ≥5 days was defined as 'prolonged', this being the 90th percentile of admission for these patients. The 10th percentile of the left ventricular ejection fraction was 50%; the 90th percentile of the cardiothoracic ratio was 0.68, and its 95th percentile 0.72. The 95th percentile of the cardiopulmonary bypass (CPB) time was 185 min.

A total of 25 patients died in the hospital, 11 without a prolonged ICU stay. Among these patients, one patient could not be weaned from CPB, four patients died from low cardiac output syndrome, three died from ventricular fibrillation, two from multiple organ dysfunction syndrome, and one patient died from ventricular rupture.

A total of 345 patients required prolonged ICU stay. The main clinical characteristics and perioperative data are shown in Table I. The definitions of important variables are listed in Table II. The main causes of valve dysfunction are shown in Figure 1, and the risk factors, regression coefficient, simple logistic score, OR, and additive scores are listed in Table III. The

Table III: The Fuwai risk score: Simple logistic score and additive score.

Risk factors	p-value	Logistic coefficient	Logistic score	Additive score
Patient-related factors				
Age ≥65 years	<0.001	1.1068368	11	3
Pulmonary dysfunction	<0.001	0.6335934	6	2
History of stroke	0.04	0.5601759	6	2
Previous cardiac surgery	0.001	0.7372775	7	2
Cardiac-related factors				
Ischemic MR	0.02	0.9968896	10	3
40% <LVEF ≤50%	0.001	0.6961856	7	2
LVEF ≤40%	<0.001	1.7241158	17	5
0.68 ≤CTR <0.72	<0.001	1.0041032	10	2
CTR ≥0.72	<0.001	1.0445035	11	3
Pulmonary hypertension	0.01	0.4356828	4	1
Surgery-related factors				
Combined with CABG	0.009	0.5684678	6	2
Double valve surgery	0.01	0.3389950	3	1
Repeat CPB support	0.001	1.4373544	14	4
CPB time ≥185 min	<0.001	1.1312184	11	3
Constant		-2.7237197		

CABG: Coronary artery bypass grafting; CTR: Cardiothoracic ratio; CPB: Cardiopulmonary bypass; LVEF: Left ventricular ejection fraction; MR: Mitral regurgitation.

Table IV: Discriminatory power and calibration of different models.

Variable	Statistic	Fuwai risk score		EuroSCORE		Parsonnet score
		Logistic algorithm	Additive algorithm	Logistic algorithm	Additive algorithm	
ICU ≥5 days	*C* statistic	0.76 (0.74-0.77)	0.75 (0.74-0.77)	0.67 (0.65-0.69)*	0.66 (0.64-0.68)*	0.63 (0.61-0.65)*
	Goodness of fit	p = 0.25	p <0.001	p <0.001	p <0.001	p <0.001
Mortality	*C* statistic	0.82 (0.80-0.83)	0.82 (0.80-0.83)	0.71 (0.69-0.73)†	0.69 (0.67-0.71)††	0.69 (0.67-0.71)†††
	Goodness of fit	p <0.001	p = 0.003	p <0.001	p <0.001	p <0.001

*p <0.001 versus Fuwai risk score logistic model and additive model.
†p = 0.04 versus Fuwai risk score logistic model; p = 0.03 versus Fuwai risk score additive model.
††p = 0.01 versus Fuwai risk score logistic model and additive model.
†††p = 0.03 versus Fuwai risk score logistic model and additive model.

$S_{logistic}$ and related probability are presented in Figure 2; for example, if a patient had only two risk factors (previous cardiac surgery and double valve surgery), then $S_{logistic} = S_{previous\ cardiac\ surgery} + S_{double\ valve\ surgery} = 7 + 3 = 10$. The data in Figure 2 show that the resultant predictive probability is 15%.

Observed probabilities compared well with predicted probabilities for the logistic algorithm (Fig. 3). Moreover, patients with higher risks had longer ICU stay (Fig. 4). However, the EuroSCORE logistic algorithm and Parsonnet score greatly underestimated the risk for high-risk patients (Fig. 3).

The Fuwai risk score showed very good discriminatory ability in predicting a prolonged ICU stay (C statistic: logistic 0.76; additive 0.75); the discriminatory power was also good for predicting mortality (Table IV). The simple logistic algorithm of the Fuwai risk score showed good calibration (Hosmer-Lemeshow, $p = 0.25$) for predicting a prolonged ICU stay, whilst calibration of the additive algorithm was poor (Hosmer-Lemeshow, $p < 0.001$). The additive algorithm greatly underestimated the risk for high-risk patients, and this contributed to its poor calibration. The Fuwai risk score did not show satisfactory calibration when predicting mortality. Neither the EuroSCORE nor Parsonnet score were superior to Fuwai risk score for predicting either a prolonged ICU stay (Fig. 5) or mortality (Fig. 6).

Discussion

Although the EuroSCORE and Parsonnet score were originally designed to predict mortality, it has been reported that both systems can be used to predict prolonged ICU stay following cardiac surgery. For example, Lawrence et al. (11) suggested that the Parsonnet score could be used to predict an ICU stay <24 h, while Nilsson et al. (12) claimed that the additive EuroSCORE algorithm could predict an ICU stay of more than 2 days after open-heart surgery. Previously, the Ontario Province Risk score had been used to predict ICU stays ≥6 days, but its area under the ROC curve was only 0.66 (13). Today, surgeons pay much more attention to patients with prolonged ICU stay, because this not only involves greater costs in terms of healthcare resources but the patients also have a poor long-term prognosis (2-4). A prior study conducted in Chinese patients at the Fuwai Hospital (14) identified the risk factors for prolonged ICU stay, but was unable to establish a complete scoring system and the additive model was not constructed; neither was validation of the Parsonnet score tested.

Previously, an attempt was made to modify the EuroSCORE for this type of patient (15), but this approach was not used widely and the present authors were reluctant concerning its use. The results of a previous study and the present investigation showed that the risk factors for prolonged ICU stay differed greatly from the factors of EuroSCORE. In theory, even if the EuroSCORE were to be modified it is reasonable to speculate that it would still not provide good results, and the development of a totally new system was deemed necessary.

The logistic equation requires computer facilities for its solution; hence, it is not convenient for clinical use and many surgeons are reluctant to use this system. The additive EuroSCORE algorithm is easier and more convenient to apply, and has been widely used during the past decade, although an increasing number of investigators have shown it to be inferior to its logistic counterpart (15-18). The results of the present study showed the additive algorithm greatly to underestimate the risk for high-risk patients, thereby confirming the results of other studies (15-18). Jin et al. suggested that, when possible, the logistic equation might be used to predict mortality, and should always be used in both clinical and research environments (15). The logistic score developed in the present study was not only very accurate but also simple to use when predicting a prolonged ICU stay, without any need for computer facilities to solve complex logistic equations. Thus, it is hoped that increasing numbers of surgeons accept this simple logistic algorithm in the future.

When compared to the findings of other studies, the discriminatory ability of the logistic and additive EuroSCORE algorithms was relatively disappointing in predicting mortality (19-21). As with the Parsonnet score, the results were similar to those reported from other Asian countries, such as Saudi Arabia and Japan (21,22). Today, the EuroSCORE has gained wide popularity as a risk-stratifying tool, with many groups reporting it to serve as a simple, objective, and robust scoring system. However, in Australia, Yap and colleagues found that the EuroSCORE did not accurately predict the outcome of patients undergoing cardiac surgery (23), which suggested that it might not be suitable for use outside Europe. The epidemiology of heart disease among the Chinese population differs from that of European and American populations, with valvular disease being more common than coronary artery disease. Originally, the EuroSCORE was designed for predicting mortality for all types of cardiac surgery, and the Parsonnet score for surgery of acquired adult heart disease. However, when these models were applied to patients with acquired heart valve disease, their predictive ability may be reduced. For example, unstable angina, emergency, and postinfarct septal rupture are rare circumstances in patients with acquired heart valve disease.

Study limitations

Because circulatory arrest, preoperative shock, post infarct ventricular septal defect and papillary muscle rupture are very rare in patients with acquired heart valve disease, these factors were not included in this study, and hence the number of patients with these factors was insufficient to obtain a stable and accurate statistical result. A too-small number of patients with preoperative renal dysfunction and active endocarditis might also suggest that these risk factors were not associated with a prolonged ICU stay. Although the single institution nature of the study might be considered an additional limitation, cardiac surgeons are invited to test the new system in their hospitals, not only as an overall measure but also in individual patients.

In conclusion, a new scoring system to predict prolonged ICU stay following surgery to treat acquired heart valve disease has been established. The logistic algorithm of the Fuwai risk score provides a simple, objective scoring system with satisfactory predictive ability. To the present authors' knowledge, this is the first study to demonstrate the construction of a new scoring system to predict the outcome of cardiac surgery in Asia. Although developed using information from a Chinese database, the Fuwai risk score may also be of value to patients in other Asian countries such as India and Vietnam.

References

1. Ferraris VA, Ferraris SP, Singh A. Operative outcome and hospital cost. J Thorac Cardiovasc Surg 1998;115:593-602
2. Hein OV, Birnbaum J, Wernecke K, et al. Prolonged intensive care unit stay in cardiac surgery: Risk factors and long-term-survival. Ann Thorac Surg 2006;81:880-885
3. Hellgren L, Stahle E. Quality of life after heart valve surgery with prolonged intensive care. Ann Thorac Surg 2005;80:1693-1698
4. Bashour CA, Yared JP, Ryan TA, et al. Long-term survival and functional capacity in cardiac surgery patients after prolonged intensive care. Crit Care Med 2000;28:3847-3853
5. Roques F, Michel P, Goldstone AR, et al. The logistic EuroSCORE. Eur Heart J 2003;24:881-882
6. Nashef SA, Roques F, Michel P, et al. European system for cardiac operative risk evaluation (EuroSCORE). Eur J Cardiothorac Surg 1999;16:9-13
7. Parsonnet V, Dean D, Bernstein AD. A method of uniform stratification of risk for evaluating the results of surgery in acquired adult heart disease. Circulation 1989;79:I3-I12
8. Swets JA. Measuring the accuracy of diagnostic systems. Science 1988;240:1285-1293
9. Lemeshow S, Hosmer DW, Jr. A review of goodness of fit statistics for use in the development of logistic regression models. Am J Epidemiol 1982;115:92-106
10. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983;148:839-843
11. Lawrence DR, Valencia O, Smith EE, et al. Parsonnet score is a good predictor of the duration of intensive care unit stay following cardiac surgery. Heart 2000;83:429-432
12. Nilsson J, Algotsson L, Hoglund P, et al. EuroSCORE predicts intensive care unit stay and costs of open heart surgery. Ann Thorac Surg 2004;78:1528-1534
13. Tu JV, Jaglal SB, Naylor CD. Multicenter validation of a risk index for mortality, intensive care unit stay, and overall hospital length of stay after cardiac surgery. Circulation 1995;91:677-684
14. Xu J, Ge Y, Pan S, et al. A preoperative and intraoperative predictive model of prolonged intensive care unit stay for valvular surgery. J Heart Valve Dis 2006;15:219-224
15. Jin R, Grunkemeier GL. Additive vs. logistic risk models for cardiac surgery mortality. Eur J Cardiothorac Surg 2005;28:240-243
16. Gogbashian A, Sedrakyan A, Treasure T. EuroSCORE: A systematic review of international performance. Eur J Cardiothorac Surg 2004;25:695-700
17. Karthik S, Srinivasan AK, Grayson AD, et al. Limitations of additive EuroSCORE for measuring risk stratified mortality in combined coronary and valve surgery. Eur J Cardiothorac Surg 2004;26:318-322
18. Zingone B, Pappalardo A, Dreas L. Logistic versus additive EuroSCORE. A comparative assessment of the two models in an independent population sample. Eur J Cardiothorac Surg 2004;26:1134-1140
19. Nashef SA, Roques F, Hammill BG, et al. Validation of European System for Cardiac Operative Risk Evaluation (EuroSCORE) in North American cardiac surgery. Eur J Cardiothorac Surg 2002;22:101-105
20. Roques F, Nashef SA, Michel P, et al. Does EuroSCORE work in individual European countries? Eur J Cardiothorac Surg 2000;18:27-30
21. Kawachi Y, Nakashima A, Toshima Y, et al. Risk stratification analysis of operative mortality in heart and thoracic aorta surgery: Comparison between Parsonnet and EuroSCORE additive model. Eur J Cardiothorac Surg 2001;20:961-966
22. Syed AU, Fawzy H, Farag A, et al. Predictive value of EuroSCORE and Parsonnet scoring in Saudi population. Heart Lung Circ 2004;13:384-388
23. Yap CH, Reid C, Yii M, et al. Validation of the EuroSCORE model in Australia. Eur J Cardiothorac Surg 2006;29:441-446