Journal of Cardiology 68 (2016) 135–140

Contents lists available at ScienceDirect

Journal of Cardiology

journal homepage: www.elsevier.com/locate/jjcc

Original article

Comparison of modern risk scores in predicting operative mortality for patients undergoing aortic valve replacement for aortic stenosis

Hironobu Yamaoka (MD)[1], Kenji Kuwaki (MD, PhD)[1,*], Hirotaka Inaba (MD, PhD), Taira Yamamoto (MD, PhD), Tomoko S. Kato (MD, PhD), Shizuyuki Dohi (MD, PhD), Satoshi Matsushita (MD, PhD), Atsushi Amano (MD, PhD, FJCC)

Department of Cardiovascular Surgery, Juntendo University, Tokyo, Japan

ARTICLE INFO

Article history:
Received 20 January 2015
Received in revised form 1 August 2015
Accepted 24 August 2015
Available online 26 September 2015

Keywords:
EuroSCORE II
STS score
Ambler score
Japan score
Aortic valve replacement

ABSTRACT

Background: The aim of our study was to examine and compare the predictive value of operative mortality of the European System for Cardiac Operative Risk Evaluation (EuroSCORE) II, the Society of Thoracic surgeons (STS) score, the Ambler score, and the Japan score in patients undergoing aortic valve replacement (AVR) for aortic stenosis (AS).

Methods: A total of 406 patients undergoing AVR with or without coronary artery bypass grafting (CABG, $n = 139$) at our institution from August 2002 to December 2013 were enrolled in this cohort study. Accuracy of calibration and discrimination performance of four risk scores was assessed in the overall patient population and quartiles of each risk score.

Results: Operative mortality was 3.4% ($n = 14$). The mean EuroSCORE II, STS score, Ambler score, and the Japan score were 3.1%, 4.9%, 5.7%, and 3.2%, respectively, giving observed-to-expected (O/E) ratio of 1.09, 0.69, 0.59, and 1.06. The *C*-statistics for operative mortality were 0.704 (EuroSCORE II), 0.781 (STS score), 0.709 (Ambler score), and 0.771 (Japan score). In the last quartile, the EuroSCORE II demonstrated excellent calibration (O/E ratio = 0.97) and discrimination (*C*-statistic = 0.904), and the STS score (O/E ratio = 0.86, *C*-statistic = 0.779) and the Japan score (O/E ratio = 1.14, *C*-statistic = 0.80) showed reasonable correlation. However, the risk calibration by the Ambler score in the last quartile was unacceptable (O/E ratio = 0.60).

Conclusions: The EuroSCORE II and the Japan score showed superior ability of calibration, but the STS score and the Ambler score overestimated the risk. However, the discrimination power was similar among the four risk scores. In the last quartile of risk, the EuroSCORE II gave an excellent predictive performance.

Introduction

Aortic valve replacement (AVR) has been the gold standard treatment for severe aortic stenosis (AS) [1]. In recent years, the prevalence of AS is increasing in the elderly population, many of whom have several comorbidities that increase surgical risk [2]. Recent studies have demonstrated that nearly 30% of the severe AS patients either refused or were not proposed for AVR due to the high-risk of operative mortality and morbidity [3,4]. The advent of transcatheter aortic valve implantation (TAVI) as a less invasive alternative to surgical AVR for high-risk patients has increased awareness of accurate preoperative risk assessment by surgical risk score to decide proper treatment strategy. Although a number of risk scoring models for the prediction of perioperative mortality after cardiac surgery have been developed, the most widely used risk scores are the EuroSCORE (European System for Cardiac Operative Risk Evaluation), the STS (Society of Thoracic Surgeons) score, and the Ambler score.

The EuroSCORE was developed in 1999, and a new version of the original EuroSCORE, named EuroSCORE II, was published in 2012 to improve the score's performance in contemporary cardiac surgery [5,6]. The STS score was published in 2008, and specifically considers the type of valve surgery with or without coronary artery bypass grafting [7,8]. The Ambler score is another risk model

* Corresponding author at: Department of Cardiovascular Surgery, Juntendo University, 2-1-1, Hongo, Bunkyo-ku, Tokyo 113-8421, Japan.
Tel.: +81 3 3813 3111; fax: +81 3 3815 5228.
E-mail address: kuwakikj@yahoo.co.jp (K. Kuwaki).
[1] These two authors contributed equally to this study.

http://dx.doi.org/10.1016/j.jjcc.2015.08.017
0914-5087/

specifically developed for prediction of operative mortality after heart valve surgery in 2005 [9]. The Japan score is also an important risk algorithm [10]. These four risk scores are important tools in the assessment of surgical risk for heart valve surgery. We, therefore, examined the performance and accuracy of the EuroSCORE II, STS score, Ambler score, and Japan score in predicting operative mortality in patients with AS who underwent AVR.

Methods

A total of 406 patients who underwent AVR for either pure AS or predominant AS at Juntendo University Hospital between August 2002 and December 2013 were included in this study. Patients were excluded from the study if they required concomitant mitral and/or tricuspid valve surgery, but those with concomitant coronary artery bypass grafting (CABG) were included. Baseline preoperative patient characteristics of this study are presented in Table 1. This study was approved by the Medical Ethics Committee of Juntendo University.

Operative mortality was defined as death within 30 days of surgery or death at any time before discharge from the hospital. Mean follow-up was 4.6 years (range, 0.3–12.3 years), and cumulative follow-up was 1780 patient-years. Four patients were lost to follow-up, and these patients were excluded from analysis of late survival. Completeness of follow-up was 99%. The EuroSCORE II (http://www.euroscore.org/calc.html), the STS score (http://riskcalc.sts.org/stswebriskcalc273), Ambler score (http://www.vcl.ac.uk/statistics/research/riskmodel/index.html), and Japan score (https://center6.umin.ac.jp/islet/jcvsd/index.html) were calculated using online calculators.

All surgical procedures were performed through a median sternotomy using cardiopulmonary bypass with systemic normothermia. Myocardial protection was achieved with combined antegrade and retrograde cold blood cardioplegia. The aortic valve prostheses used in this study included 296 bioprostheses (Carpentier-Edwards, CA, USA, $n = 225$; Mosaic, MN, USA, $n = 33$; Trifecta, MN, USA, $n = 31$; Mitroflow, Vancouver, Canada, $n = 7$) and 110 mechanical prostheses (St. Jude Medical, MN, USA, $n = 45$; OnX, TX, USA, $n = 40$; ATS, MN, USA, $n = 21$; CarboMedics, TX, USA, $n = 3$; Bicarbon, Saluggia, Italy, $n = 1$), and concomitant CABG was performed in 34% of the patients ($n = 139$). The average number of CABG distal anastomoses per patients was 2.2.

Table 1
Preoperative patient characteristics.

	($n = 406$)
Age, mean ± SD	71.6 ± 9.9
Male, *n* (%)	217 (53)
BMI, kg/m^2, mean ± SD	23.0 ± 3.3
Ejection fraction, %, mean ± SD	62.8 ± 13.7
NYHA class III–IV, *n* (%)	137 (34)
Dialysis, *n* (%)	55 (14)
Hypertension, *n* (%)	297 (73)
Diabetes mellitus, *n* (%)	112 (28)
History of stroke, *n* (%)	34 (8)
Peripheral vascular disease, *n* (%)	30 (7)
History of myocardial infarction, *n* (%)	30 (7)
Atrial fibrillation, *n* (%)	23 (6)
COPD, *n* (%)	35 (9)
Prior cardiac surgery, *n* (%)	14 (3)
Emergency or urgent surgery	7 (2)
EuroSCORE II, mean %	3.1
STS score, mean %	4.9
Ambler score, mean %	5.7
Japan score, mean %	3.2

BMI, body mass index; NYHA, New York Heart Association; COPD, chronic obstructive pulmonary disease.

Statistical analysis

Categorical variables were given as percentages and were compared between groups using a chi-squared test or Fisher's exact test. Continuous variables were reported as mean ± standard deviation and were compared between groups using Student's *t*-test or repeated measures ANOVA (analysis of variance). Kaplan–Meier curves and the log-rank test were used for survival analysis for long-term mortality during follow-up. A *p*-value <0.05 was considered statistically significant. A chi-square test or Fisher's exact test was used to compare the number of observed and predicted operative deaths for each score in the overall patient population and in the subgroups of patients who were divided into quartiles, each based on the risk score to assess the calibration performance. The Hosmer–Lemeshow test and Brier score were also calculated. Receiver operating characteristics analysis was used to calculate the *C*-statistic, and 95% confidence intervals for operative mortality to assess discrimination ability. Comparison between two *C*-statistics was performed by the Hanley and McNeil method [11] using the statistical package MEDCALC 12.2. For the rest of the analysis, SPSS version 18.0 (IBM, Armonk, New York, USA) was used.

Results

The observed operative mortality in this study was 3.4% (14 out of 406 patients) (Table 2). The causes of deaths included cardiac failure in 4 patients, gastrointestinal complication in 3, cerebral infarction in 2, fatal arrhythmia in 2, sepsis, pneumonia, and mediastinitis in 1 patient each. The mean predicted operative mortality by the EuroSCORE II, the STS score, Ambler score, and Japan score was 3.1%, 4.9%, 5.7%, and 3.2%, respectively (Table 3). In terms of observed/expected mortality ratio (O/E ratio), the EuroSCORE II (O/E ratio = 1.09) and the Japan score (O/E ratio = 1.06) demonstrated a slight but very acceptable underprediction. The STS score moderately overestimated the operative mortality (O/E ratio = 0.69), but it was not statistically significant. The Ambler score also overestimated the risk (O/E ratio = 0.59), and there was a strong tendency toward statistical significance ($p = 0.05$) between observed and Ambler predicted mortality (Table 3). The Ambler score also had the highest Brier score (0.0318) when compared with the EuroSCORE II (Brier score = 0.0297), the STS score (Brier score = 0.0289), and the Japan score (Brier score = 0.0296), although no statistically significant difference was found between the four risk scores ($p = 0.17$).

Observed and expected operative mortality in quartiles for the four risk scores is presented in Fig. 1. The EuroSCORE II mildly to moderately underestimated the risk for the first to third quartiles, but it made a good prediction for the highest quartile (O/E ratio = 0.97). The STS score overestimated the risk in all quartiles. The Ambler score underestimated the mortality for the first and

Table 2
Operative procedures and outcome.

Valve type implanted	
Bioprosthetic valve, *n* (%)	296 (72.9)
Mechanical valve, *n* (%)	110 (27.1)
Effective orifice area index	
≤0.85	42 (10.3)
>0.85	364 (89.7)
Aortic cross-clamp time, min, mean ± SD	110 ± 44
Cardiopulmonary bypass time, min, mean ± SD	130 ± 46
Concomitant CABG, *n* (%)	139 (34)
Early deaths, *n* (%)	14 (3.44)
Survival at 5 years, %	92.5 ± 1.4

CABG, coronary artery bypass grafting.