Comparison of original EuroSCORE, EuroSCORE II and STS risk models in a Turkish cardiac surgical cohort†

Ayse Gul Kunta,*, Murat Kurcepeh, Mete Hidiroglu, Levent Cetin, Aslihan Kucuk, Vedat Bakuy, Ahmet Ruchan Akar and Erol Sener

INTRODUCTION

Scoring systems are an essential part of current cardiac surgical practice in assessing operative mortality and morbidity. Changes in cardiac surgery case mix, surgical techniques and clinical outcomes continuously lead investigators to improve and modify currently available risk-stratification systems. All improvements and modifications require further validation tests for different patient populations globally.

A European System for Cardiac Operative Risk Evaluation (EuroSCORE) additive (standard) model from eight European countries was developed between 1995 and 1999 [1, 2] and has gained wide acceptance in Europe, North America [3] and Asia [4, 5]. However, this model generally overestimates mortality in low-risk patients (EuroSCORE ≤6) and underestimates it in high-risk patient groups (EuroSCORE >13) [6]. The logistic EuroSCORE model was developed to improve the predictive performance in high-risk patients [7, 8]. The investigators of the EuroSCORE have continued their work creating a new version called EuroSCORE II [9, 10] to improve the accuracy of the original EuroSCORE for estimating the risk of death after cardiac surgery. The Society of Thoracic Surgeons (STS) database was first established in 1989 to assess operative mortality after cardiac surgery in adults [11-13]. This risk-stratification model has been widely used in North America and comprised over 40 clinical parameters [14].

EuroSCORE and the STS risk algorithms have been compared for coronary bypass grafting [15, 16]. However, similar comparisons...
between the EuroSCORE, EuroSCORE II and STS in the patient populations for coronary bypass grafting are limited. We, therefore, aimed to compare these three risk models on a prospectively collected data from the Turkish cardiac surgical population stored in the TurkoSCORE database.

MATERIALS AND METHODS

Data collection

Turkish patients (n = 428) undergoing isolated CABG at the Ataturk Education and Research Hospital between June 2004 and March 2012 were extracted from the TurkoSCORE database. Patients who underwent isolated coronary artery bypass grafting and older than 70 years were included in this study. Patients were excluded from the study if concomitant other cardiac procedures were performed. A total of 546 variables for each patient data set including demographic and administrative (n = 33), preoperative risk factors and medications (n = 77), preoperative evaluations (n = 112), laboratory findings including genetic risk factors (n = 60), intraoperative (n = 160), postoperative data and complications (n = 59), follow-up (n = 32), and mortality and morbidity (n = 13) have been recorded in this database [5]. For all patients already having additive and logistic EuroSCORE and STS values, the new version of EuroSCORE was applied to the original data set. Observed and predicted mortalities were compared for the additive and logistic EuroSCORE, EuroSCORE II and STS risk calculator.

Receiver operating characteristic curve analysis

The area under the receiver operating characteristic (ROC) curve (AUC) is the value used to compare different risk-estimation methods [16]. Each risk-estimation method provides an ordering for the patient population based on the risk of mortality. Then, a ROC curve is obtained based on this ordering. Each ordering results in an AUC value between 0.5 and 1.0 where the 0.5 AUC value indicates a random ranking of the patients being alive and deceased. The higher the AUC value, the better the performance of the risk-assessment method. In this study, AUC values were calculated for additive and logistic EuroSCORE, EuroSCORE II and STS models to compare the predictive performance of these risk-assessment methods.

Statistical analysis

In prevalence analysis, categorical variables were presented as numbers and/or percentages, and continuous variables presented as mean ± standard deviation (SD). In ROC analysis, a non-parametric method was employed to calculate standard error for area under ROC (AUC) and to provide 95% confidence intervals (CIs). One-sample t-test was employed during the comparison of observed and predicted mortality rates. A two-sided P-value of <0.05 was considered to indicate statistical significance between the observed mortality and the predicted mortality rates. Data were analysed using SPSS 16.0 (SPSS, Inc., Chicago, IL, USA) for Windows (Microsoft Corp., Redmond, WA, USA).

RESULTS

Patients' characteristics

The clinical features of the study group are reported in Table 1. The mean patient age was 74.5 years at the time of surgery. The segment of age between 70–74 years was 13.1% in the Turkish population (n = 8018) and 17.9% in the European population.

| Table 1: Prevalence of risk factors in EuroSCORE, EuroSCORE II and STS risk-stratification algorithms |
|-----------------------------------------------|----------------|----------------|----------------|
| Age (years)                                   | √              | √              | √              |
| Gender (female)                               | √              | √              | √              |
| Chronic pulmonary disease                     | √              | √              | √              |
| Extracardiac arteriopathy                     | √              | √              | √              |
| Neurological dysfunction                      | √              | √              | √              |
| Previous cardiac surgery                      | √              | √              | √              |
| Creatinine >200 µmol/l                        | √              | √              | √              |
| Critical preoperative state                   | √              | √              | √              |
| Unstable angina                               | √              | √              | √              |
| Left ventricular function                     | √              | √              | √              |
| Acute myocardial infarction                   | √              | √              | √              |
| Myocardial infarct >8-21 days                 | √              | √              | √              |
| Pulmonary hypertension                        | √              | √              | √              |
| Pulmonary hypertension—moderate               | √              | √              | √              |
| Emergency                                     | √              | √              | √              |
| Diabetes control—insulin                      | √              | √              | √              |
| NHYA class 1                                  | √              | √              | √              |
| Hypertension                                  | √              | √              | √              |
| Three-vessel disease                          | √              | √              | √              |

STS: Society of Thoracic Surgeons.
The discrepancies between observed and expected mortality rates were apparent in our patient population. For the entire cohort, actual hospital mortality was 7.9% (n = 245 132) [2, 5, 13]. Patients’ outcomes

The discrepancies between observed and expected mortality rates were apparent in our patient population. For the entire cohort, actual hospital mortality was 7.9% (n = 245 132) [2, 5, 13].

DISCUSSION

This study evaluates the validity and applicability of the current gold standard risk-prediction models namely additive EuroSCORE, logistic EuroSCORE, EuroSCORE II and STS risk calculator in a patient population from a single institution. The results of our study suggest that EuroSCORE II significantly underestimated the mortality risk for Turkish cardiac patients, whereas additive and logistic EuroSCORE and STS risk calculators were well calibrated in this cohort. Although the discriminatory ability of the models was very satisfactory, we are unable to identify whether the failure of the EuroSCORE II model in our study is due to factors unique to Turkish surgical patients or factors newly added in the model.

EuroSCORE has gained wide acceptance in Europe and also in Turkey in predicting in-hospital mortality in cardiac surgery. In fact, the use of the EuroSCORE model in the adult Turkish cardiac surgical population is obligatorily practiced by the national health authority and Turkish Social Security Agency. Additive and logistic EuroSCORE were introduced in 1999 and 2003, respectively [1, 2, 7, 8]. The additive EuroSCORE is designed by using the \( \beta \) coefficients as weights for each risk factor, whereas the logistic EuroSCORE model aimed to improve the predictive performance of high-risk patients with a desire to use the full logistic equation of EuroSCORE rather than the approximation. However, the present study analysed the performance of the additive and logistic EuroSCORE models on a Turkish cardiac surgical cohort that was dependent on the original risks.
databases. In agreement with the EuroSCORE study group, the additive EuroSCORE model predicted very well in our patient cohort.

Nashef et al. [1] analysed 14,799 patients and found that the observed mortality was 4.37 ± 5.06 and predicted mortality was 4.72 ± 4.95 for the entire cohort. The group identified the superiority of the logistic EuroSCORE in predicting mortality in high-risk cardiac surgical patients [7]. They reported that the logistic model was closer to actual mortality than the additive model [4.84% (4.72–4.94) vs 4.21 (4.21–4.26)]. They realized that most of the difference was due to high-risk patients where actual mortality was 11.18% and the predicted was 7.8% (additive) and 11.2% (logistic). However, we previously showed that the original EuroSCORE risk models overestimated mortality at all risk subgroups in a Turkish cardiac surgical population (n = 9443) from 2005 to 2010 [5].

Subsequently, Nashef et al. [10] updated the EuroSCORE risk model on 22,381 consecutive patients undergoing major cardiac surgery in 154 hospitals in 43 countries over a 12-week period (May–July 2010) to compare with the original 1995 EuroSCORE database. Recently developed EuroSCORE II included additional surgical risk predictors such as poor mobility, diabetes on insulin, New York Heart Association, Canadian Cardiovascular Society Class 4 angina and weight of the intervention; and excluded neurological dysfunction, unstable angina and post-infarct septal rupture. They validated EuroSCORE II on a data subset of 5553 patients (actual mortality: 4.2% and predicted: 4.0%), and very good discrimination was maintained with an area under the ROC curve of 0.81. EuroSCORE II also improved on the original logistic EuroSCORE, though mainly for combined aortic valve replacement and CABG cases in a cohort of patients from UK [17]. In contrast to recent reports, prospectively collected data from the Society for Cardiothoracic Surgery in Great Britain and Ireland database revealed that the model is poorly calibrated for isolated coronary artery bypass grafting and in both the highest- and lowest-risk patients [18]. In our study, the EuroSCORE II significantly underestimated mortality risk in our elderly cardiac cohort (observed: 7.9% and predicted: 1.7%). Recently introduced predictors in EuroSCORE II may have an impact on underestimated mortality risk in our Turkish surgical cohort. The information regarding poor mobility secondary to musculoskeletal dysfunction could have been missed because of the retrospective study design. However, there was no incidence of missing data in this surgical cohort.

The first version of The Society of Thoracic Surgeons National Adult Cardiac Surgery Database (STS NCD) was developed nearly two decades ago. The STS risk calculator predicts morbidity factors such as stroke, renal dysfunction, reoperation, prolonged ventilation, deep sternal wound infection and the length of hospital stay. All STS risk models have undergone periodic revisions. The STS risk model on coronary artery bypass grafting originated from 774,881 isolated CABG procedures from 819 centres performed between 2002 and 2006. Nilsson et al. [15] compared EuroSCORE and STS risk algorithms in a cohort of 4497 CABG operations from Sweden. The actual 30-day mortality was 1.9% (95% CI 1.5–2.3%) in their cohort. The investigators reported that the area under the ROC curve was 0.84 (95% CI 0.80–0.88) for EuroSCORE and 0.71 (95% CI 0.66–0.77) for STS; the discriminatory power (area under the ROC curve) was significantly larger for EuroSCORE compared with STS (P < 0.00005). They concluded that the additive EuroSCORE algorithm had a significantly better discriminatory power to predict 30-day mortality than the STS risk algorithm for patients undergoing coronary artery bypass in the Swedish population. In our Turkish cardiac cohort, the original EuroSCORE and STS risk models were well calibrated.

The major limitations of this study are the sample size and single-centre design, which limit generalized conclusions for the Turkish cardiac population.

CONCLUSION

In our practice, the original additive and logistic EuroSCORE risk models and STS risk calculator estimated mortality very well in
the cardiac Turkish population compared with EuroSCORE II. Ethnicity, seasonal variations and single-centre study should be kept in mind.

Funding

This study was supported by the Turkish Society of Cardiovascular Surgery.

Conflict of interest: none declared.

REFERENCES