Validation of the EuroSCORE risk models in Turkish adult cardiac surgical population

Ahmet Ruchan Akar a, *, 1, Murat Kurtcephe b, 1, Erol Sener c, Cem Alhan d, Serkan Durdu a, Ayse Gul Kunt c, Halil Altay G¨uvenir b, The working Group for the Turkish Society of Cardiovascular Surgery and Health Committee supported this study.

Department of Cardiovascular Surgery, Ankara University School of Medicine, Ankara, Turkey
Bilkent University, Engineering Faculty, Department of Computer Engineering, Ankara, Turkey
Department of Cardiovascular Surgery, Ankara Ataturk Hospital, Ankara, Turkey
Department of Cardiovascular Surgery, Acibadem University, Istanbul, Turkey

Received 21 July 2010; received in revised form 30 December 2010; accepted 5 January 2011; Available online 20 February 2011

Abstract
Objective: The aim of this study was to validate additive and logistic EuroSCORE risk models on Turkish adult cardiac surgical population. Methods: EuroSCORE project involves a reliable web-based database to build up Turkish risk stratification models. Current patient population consisted of 9443 adult patients who underwent cardiac surgery between 2005 and 2010. However, the additive and logistic EuroSCORE models were applied to only 8018 patients whose EuroSCORE determinants were complete. Observed and predicted mortalities were compared for low-, medium-, and high-risk groups. Results: The mean patient age was 59.5 years (±12.1 years) at the time of surgery, and 28.6% were female. There were significant differences (all p < 0.001) in the prevalence of recent myocardial infarction (23.5% vs 9.7%), moderate left ventricular function (29.9% vs 25.6%), unstable angina (9.8% vs 8.0%), chronic pulmonary disease (13.4% vs 3.9%), active endocarditis (3.2% vs 1.1%), critical preoperative state (9.0% vs 4.1%), surgery on thoracic aorta (3.7% vs 2.4%), extracardiac arteriopathy (8.6% vs 11.3%), previous cardiac surgery (4.1% vs 7.3%), and other than isolated coronary artery bypass graft (CABG; 23.0% vs 36.4%) between Turkish and European cardiac surgical populations, respectively. For the entire cohort, actual hospital mortality was 1.96% (n = 157; 95% confidence interval (CI), 1.70—2.32). However, additive predicted mortality was 2.98% (p < 0.001 vs observed; 95%CI, 2.90—3.00), and logistic predicted mortality was 3.17% (p < 0.001 vs observed; 95%CI, 3.03—3.21). The predictive performance of EuroSCORE models for the entire cohort was fair with 0.757 (95%CI, 0.707—0.800) AUC value (area under the receiver operating characteristic, AUC) for additive EuroSCORE, and 0.760 (95%CI, 0.721—0.800) AUC value for logistic EuroSCORE. Observed hospital mortality for isolated CABG was 1.23% (n = 75; 95%CI, 0.95—1.51) while additive and logistic predicted mortalities were 2.87% (95%CI, 2.82—2.93) and 2.89% (95%CI, 2.80—2.98), respectively. AUC values for the isolated CABG subset were 0.768 (95%CI, 0.707—0.830) and 0.766 (95%CI, 0.705—0.828) for additive and logistic EuroSCORE models. Conclusion: The original EuroSCORE risk models overestimated mortality at all risk subgroups in Turkish population. Remodeling strategies for EuroSCORE or creation of a new model is warranted for future studies in Turkey.

Keywords: Cardiac surgery; Risk prediction model; Mortality; EuroSCORE; TurkoSCORE

1. Introduction

Risk stratification and prediction models are mandatory tools for assessing quality of care, medical decision making, and facilitating patient counseling and consenting. The additive European System for Cardiac Operative Risk Evaluation (EuroSCORE) model [1,2] is based on one of the largest and most complete European data collected between September and November 1995 from 128 hospitals in eight European countries including UK, Germany, France, Italy, Spain, Finland, Sweden, and Switzerland. In 2003, the logistic EuroSCORE model was developed to improve the predictive performance in high-risk patients [3,4]. Subsequent studies from European [5—7], North American [8], and Japanese [9] cohorts have provided compelling evidence about EuroSCORE risk stratification systems’ validation in predicting early mortality widely.
across the different cardiac surgical populations. However, important epidemiological differences in the national cohorts of patients have also been reported in the EuroSCORE database [10]. Furthermore, accurate predictions using the additive and logistic EuroSCORE models have recently failed in Australian [11] and Chinese [12] cardiac surgical practice.

EuroSCORE has also been widely used in Turkey. The national health authority, Social Security Institution stipulates the use of standard EuroSCORE model in adult Turkish cardiac surgical practice, although there is relatively scarce previous work testing EuroSCORE model in the Turkish population [13]. Thus, there has been debate as to whether additive or logistic EuroSCORE can be generalized to adult Turkish cardiac patients since the calibration (agreement between predicted probabilities and observed outcome frequencies) and discriminatory power (ability to distinguish between patients with and without the outcome) values and clinical performance for Turkish population are missing. We, therefore, aimed to validate both the additive and logistic EuroSCORE models on a prospectively collected data from Turkish cardiac surgical population stored in the TurkoSCORE database.

2. Materials and methods

2.1. Project setup

Turkish Society of Cardiovascular Surgery set up a working group to build up an adult Cardiac Surgery Database and risk model algorithm with a final target of availability for integrative data sets such as the European Association of Cardiothoracic Surgery [1—4] and the Society of Thoracic Surgeons (STS) National Adult Cardiac Surgery Databases [14—16]. A comprehensive set of variables and definitions have been designed by the Ankara University (ARA, and SD) to include parallel variables to both the EuroSCORE and the STS National Databases as well as for future multicenter clinical

<table>
<thead>
<tr>
<th>Variables</th>
<th>EuroSCORE definition [1,2]</th>
<th>TurkoSCORE definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>In years at last birthday</td>
<td>In years, at time of surgery</td>
</tr>
<tr>
<td>Gender</td>
<td>Female</td>
<td>Patient’s sex at birth as either male or female</td>
</tr>
<tr>
<td>Chronic pulmonary disease</td>
<td>Long-term use of bronchodilators or steroids for lung disease</td>
<td>Patient has a FEV1 &lt; 80% of predicted value; patient has a FEV1/FVC &lt; 70%; use of bronchodilators or steroids for lung disease</td>
</tr>
<tr>
<td>Extracardiac arteriopathy</td>
<td>Any one or more of: claudication, carotid occlusion or &gt;50% stenosis, previous or planned intervention on abdominal aorta, limb or carotids</td>
<td>Claudication either with exertion or at rest; amputation for vascular arterial insufficiency; aorto-iliac occlusive disease reconstruction; vascular reconstruction, bypass surgery, or percutaneous intervention for peripheral arterial disease; documented aortic aneurysm with or without repair; non-invasive carotid test with &gt;70% stenosis; previous intervention on carotids</td>
</tr>
<tr>
<td>Neurological disease</td>
<td>Disease severely affecting ambulation or day-to-day functioning</td>
<td>Unresponsive coma &gt;24 h; recent CVA within two weeks of the surgical procedure; any neurological disease affecting ambulation such as myasthenia gravis; TIA and RIND excluded.</td>
</tr>
<tr>
<td>Previous cardiac surgery</td>
<td>Requiring opening of the pericardium</td>
<td>Previous cardiac surgery with cardiopulmonary bypass</td>
</tr>
<tr>
<td>Serum creatinine</td>
<td>&gt;0.2 mmol/l preoperatively</td>
<td>&gt;2.26 mg/dL (closest to the date and time prior to surgery)</td>
</tr>
<tr>
<td>Active endocarditis</td>
<td>Patient still under antibiotic treatment for endocarditis at the time of surgery</td>
<td>Currently being treated for endocarditis at the time of surgery; positive blood cultures; vegetation on echocardiography; prosthetic valve endocarditis</td>
</tr>
<tr>
<td>Critical preoperative state</td>
<td>Any one or more of the following: ventricular tachycardia or fibrillation or aborted sudden death, preoperative cardiac massage, preoperative ventilation before arrival in the anesthetic room, preoperative inotropic support, intra-aortic balloon counterpulsation or preoperative acute renal failure (anuria or oliguria &lt;10 ml/h)</td>
<td>Any one of more of the following: ventricular tachycardia or fibrillation or aborted sudden death, preoperative cardiac massage, preoperative ventilation before arrival in the anesthetic room, preoperative inotropic support, intra-aortic balloon counterpulsation or preoperative acute renal failure (anuria or oliguria &lt;10 ml/h)</td>
</tr>
<tr>
<td>Unstable angina</td>
<td>Rest angina requiring i.v. nitrates until arrival in the anesthetic room</td>
<td>CCS 4C and CCS 4D</td>
</tr>
<tr>
<td>LV dysfunction</td>
<td>Moderate or LVEF 30—50%; Poor or LVEF &lt; 30%</td>
<td>Moderate or LVEF 35—49%; poor or LVEF 20—34%; severely reduced or LVEF &lt; 20%</td>
</tr>
<tr>
<td>Recent myocardial infarction</td>
<td>&lt;90 days</td>
<td>&lt;90 days</td>
</tr>
<tr>
<td>Pulmonary hypertension</td>
<td>Systolic PA pressure &gt; 60 mmHg</td>
<td>Systolic PA pressure &gt; 60 mmHg</td>
</tr>
<tr>
<td>Emergency</td>
<td>Carried out on referral before the beginning of the next working day</td>
<td>Surgery required within 24 h following referral</td>
</tr>
<tr>
<td>Other than Isolated CABG</td>
<td>Major cardiac procedure other than or in addition to CABG</td>
<td>Major cardiac procedure other than or in addition to CABG</td>
</tr>
<tr>
<td>Surgery on thoracic aorta</td>
<td>For disorder of ascending, arch or descending aorta</td>
<td>Surgery on thoracic aorta</td>
</tr>
<tr>
<td>Post-infarct septal rupture</td>
<td>Ventricular septal rupture as a complication of myocardial infarction</td>
<td></td>
</tr>
</tbody>
</table>

EuroSCORE, European system for cardiac operative risk evaluation; FEV1, forced expiratory volume in one second; LVEF, left ventricular ejection fraction; CABG, coronary artery bypass graft; CVA, cerebral vascular accident (symptoms >24 h after onset); RIND, reversible ischemic neurologic deficit; TIA, transient ischemic attack; CPR, cardiopulmonary resuscitation; CCS, Canadian Cardiovascular Society.
research and integration. The present study obtained the approval of the Institutional Review Board according to the Declaration of Helsinki. The Department of Computer Engineering of Bilkent University constructed a reliable, user-friendly web-based application enabling access from different departments (MK, and AT) and took the responsibility of working on machine learning systems (AG, and MK) on this database to learn risk models [17].

2.2. Data collection methods and definitions

Data collection methods and definitions are available from http://turkoscore.cs.bilkent.edu.tr/Turkoscore/Login.aspx. 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), intra-operative (n = 160), postoperative data and complications (n = 59), follow-up (n = 32), and mortality and morbidity (n = 13) have been recorded in this database. In brief, Table 1 summarizes the comparative definitions of EuroSCORE and TurkoSCORE models. Additive and logistic EuroSCORE values were calculated for each patient using the EuroSCORE calculator available online (http://www.euroscore.org) integrated to our database. Initially, our data set consisted of 9443 adult cardiac surgical patients from 1 January 2005 to 30 May 2010. However, patients were excluded from the study if any one of the following exclusion criteria was met: (1) patients with missing >1 EuroSCORE predictors, (2) patients with missing in-hospital mortality data, and (3) centers with irregular data input despite regular warnings. In accordance with the above-mentioned criteria, 1425 patients were excluded from the patient cohort.

2.3. Receiver operating characteristic (ROC) curve analysis

The AUC is defined as the area under the receiver operating characteristic curve. The 0.5 AUC value indicates a random ranking of the patients being alive and dead. An increasing value of AUC from 0.5 toward 1.0 indicates increasing distinctiveness and better discrimination of the patients’ status. AUC values were generated for additive and logistic EuroSCORE models to test discrimination and to describe performance and accuracy [18]. The results were presented with 95% confidence intervals.

2.4. Statistics

Categorical variables are presented as numbers and/or percentages, and continuous variables are presented as mean ± SD. The significance of differences in proportions between the 1995 EuroSCORE data set and TurkoSCORE data set was determined by the chi-square test. Normally distributed continuous variables were compared using the unpaired t-test. Standard statistical tests were used to calculate 95% confidence intervals. A two-sided p value of less than 0.05 was considered to indicate statistical significance. Data were analyzed using SPSS 16.0 (SPSS Inc., Chicago, IL, USA) for Windows (Microsoft Corp., Redmond, WA, USA).

3. Results

3.1. Patients’ characteristics

National cardiac surgery data by the Turkish Ministry of health revealed in 2009 that 66 105 adult open-heart surgeries and 5328 congenital cardiac surgical operations were performed in Turkey. The proportion of cardiac surgery workload according to institutions was as follows: 61% were performed in private hospitals, 27% in public hospitals, and, finally, 12% in university hospitals (personal communication). The proportion of cardiac surgical workload according to the type of institution in our cohort was similar to overall national statistics. The final study cohort comprised of 8018 patients from three institutions, and was compared with the 1995 EuroSCORE data set. Overall, 85% of the data set was complete. However, there was no incidence of missing in-hospital mortality data in the final cohort. The clinical features of the study group and the EuroSCORE patient population are reported in Table 2. On average, the Turkish cardiac patients were younger than the European counterparts. Moreover, the Turkish population was more likely to have chronic pulmonary disease, active endocarditis, critical preoperative state, unstable angina, moderate left ventricular function, recent myocardial infarction, and surgery on thoracic aorta. Fewer patients in Turkish population have extracardiac arteriopathy, previous cardiac surgery, and surgery other than isolated coronary artery bypass graft (CABG) surgery compared to European population. The prevalence of female sex, neuro-
logical disease, serum creatinine >0.2 mol l⁻¹, pulmonary hypertension, emergency surgery, and post-infarct septal rupture was similar between the two populations.

3.2. Patients’ outcomes

The discrepancies between observed and expected mortality rates were apparent in our patient population. For the entire cohort, observed hospital mortality was 1.96% (n = 157; 95%CI, 1.70–2.32). However, additive predicted mortality was 2.98% (p < 0.001 vs observed; 95%CI, 2.90–3.00), and logistic predicted mortality was 3.17% (p < 0.001 vs observed; 95%CI, 3.03–3.21). Observed hospital mortality for isolated coronary bypass surgery was 1.23% (n = 75; 95%CI, 0.95–1.51), additive predicted mortality was 2.87% (95%CI, 2.82–2.93) and logistic predicted mortality was 2.89% (95%CI, 2.80–2.98). The excluded patients had an overall in-hospital mortality of 2.16% (p = 0.609 vs study cohort), which avoids any significant clinical bias. As shown in Tables 3 and 4, both scoring systems overestimated mortality at each risk tertiles for the entire cohort and isolated CABG surgery, respectively.

3.3. ROC analysis

The ROC curves for the entire cohort and CABG subset are given in Figs. 1 and 2, respectively. The predictive

![Fig. 1. Receiver operator curve of the additive and logistic EuroSCORE models for adult cardiac surgery analyzed from TurkoSCORE dataset (n = 8018).](image1)

![Fig. 2. Receiver operator curve of the additive and logistic EuroSCORE models for isolated CABG subset analyzed from TurkoSCORE dataset (n = 6171).](image2)
performance of EuroSCORE models for the entire cohort was fair with 0.757 (95%CI, 0.717—0.797) AUC value for additive EuroSCORE, and 0.760 (95%CI, 0.721—0.800) AUC value for logistic EuroSCORE. AUC values for the isolated CABG subset were 0.768 (95%CI, 0.707—0.830) and 0.766 (95%CI, 0.705—0.828) for additive and logistic EuroSCORE models.

4. Discussion

Risk-adjusted quality monitoring and control have a paramount importance in the current cardiac surgical practice [6,19]. Over the last decade, professional and public interest support the use of risk-standardized outcomes in Turkish healthcare system. The EuroSCORE risk models, which include 17 independent variables and consider 30 days’ operative mortality, are the most commonly used risk prediction models in Turkish adult cardiac surgical practice. Thus, the first step was to carefully evaluate the validity and applicability of these gold-standard risk prediction models. The results of our study suggest that additive and logistic EuroSCORE risk models overestimated mortality at all risk subgroups in the Turkish cohort. Second, we observed significant differences in preoperative patient characteristics between the Turkish and 1995 EuroSCORE data sets that represent a temporal change in operative case mix. Furthermore, substantial changes in the operative techniques and perioperative care after the development of the risk score should also be kept in mind.

Additive and logistic EuroSCOREs were introduced in 1999 and 2003, respectively [1,3]. The additive EuroSCORE is simple, user friendly, and easily calculated at the bedside without specialized equipment; however, it tends to underestimate risk in high-risk groups [1,2,4]. Additive scoring is designed by using the \( \beta \) coefficients as weights for each risk factor. Logistic EuroSCORE model aimed to improve the predictive performance of high-risk patients with a desire of using the full logistic equation of EuroSCORE rather than the approximation. It is worth noting, however, that the present study analyzed the performance of the additive and logistic EuroSCORE models on a Turkish cohort that was independent from the original databases.

In agreement with previous reports [7,11,20], the additive and logistic EuroSCORE models overpredict mortality in our patient cohort. Yap et al. [11] analyzed 8331 patients based on the Australasian Society of Cardiac and Thoracic Surgeons patient database and found that the additive and logistic EuroSCORE models predicted mortality were 5.31%, and 8.76%, respectively, while the observed mortality was 3.20% for the entire cohort and 2.00% for the CABG subset. Bhatti et al. [20] analyzed prospectively collected British data including 9995 patients from ‘North West Quality Improvement Programme in Cardiac Interventions’. The investigators demonstrated that the discrimination of the logistic EuroSCORE was good with ROC curve area of 0.79 for all cardiac surgeries, but overpredicted in-hospital mortality [20]. The predictive performances of both EuroSCORE models for our patient population were fair, reaching to 0.76. Recently, D’Errigo et al. [7] analyzed 30 610 isolated CABG interventions based on the ‘Italian CABG Outcome Project’ and showed that their observed mortality was 2.54% significantly lower than the 6.27% predicted mortality by the logistic EuroSCORE. As a consequence, the investigators suggested the use of recalibration coefficient of 0.4 for logistic EuroSCORE in the Italian population.

Recently, Ranucci et al. [21] suggested recalibration of the logistic EuroSCORE in high-risk patients using adjusted model with different correction factors (0.4 for logistic EuroSCORE between 5.1 and 6.0, and 0.6 for logistic EuroSCORE between 6.1 and 25). In another study, the investigators demonstrated that a mortality risk score could be developed on the basis of a very limited number of risk factors such as age, creatinine, and left ventricular ejection fraction (ACEF) in elective cardiac operations with an accuracy equivalent to or even better than more complex models, with good calibration and satisfying clinical performance [22]. Nissinen et al. [23] from Finland supported institutionally derived modifications to improve the accuracy of EuroSCORE. Current cardiac surgical mortality in UK has been running at approximately 0.6 of EuroSCORE prediction [24], whereas the situation in the Turkish patient population changes between 0.49 and 0.67 of EuroSCORE prediction for different risk subgroups.

The major limitations of this study are the lack of contribution of all centers at the national level and the exclusion of certain cases due to missing values, which limits significant conclusions. However, we made every attempt to ensure avoidance of any selection bias due to human interference in the exclusion procedure. The present study cohort involves all university, government, and private hospitals, which reflects the variability of current adult cardiac surgical practice in Turkey. Furthermore, data accuracy was assured by cross-questioning within the database and regular internal audits.

5. Conclusion

In our practice, the original EuroSCORE risk models overestimated mortality at all risk subgroups in Turkish population. Efforts must be undertaken at a national level to promote and support the adjustment of logistic EuroSCORE model or the development of a new risk prediction model with better calibration, discrimination, and clinical performance.

Acknowledgments

The authors thank all the surgical teams who contributed patients to the study and The Working Group for the Turkish Society of Cardiovascular Surgery and Turkish Ministry of Health; Serap Aytuk Aka, Umit Ozurda, Mustafa Pac, Mehmet Ali Ozatik, Murat Sargon, Batuhan Ozay, Sahin Senay, Cagin Zaim, Cagdas Baran, Kemal Esref Erdogan, Zeki Catav, Bahadir Inan, Ifran Sencan, Orhan Koc, Gunseli Cubukuoglu Deniz, Arin Dogan, Guletek Bayraktar, and Umit Kervan. We sincerely thank Aysen Tunca for her initial work on TurkoSCORE database and Leyla Yigit PhD. for her statistical review of the manuscript.
Statement of responsibility: The authors had full access to the TurkoSCORE data and take full responsibility for their integrity.

References


