The logistic EuroSCORE predicts the hospital mortality of the thoracic aortic surgery in consecutive 327 Japanese patients better than the additive EuroSCORE

Takahiro Nishida*, Munetaka Masuda, Yukihiro Tomita, Shigehiko Tokunaga, Yosihisa Tanoue, Akira Shiose, Shigeki Morita, Ryuji Tominaga

Department of Cardiovascular Surgery, Kyushu University Graduate School of Medical Sciences, 3-1-1 Maidashi, Higashi-ku, Fukuoka 812-8582, Japan

Received 3 June 2006; received in revised form 8 July 2006; accepted 14 July 2006

Abstract

**Objective:** Risk stratification in thoracic aortic surgery is a topic of major interest. Recent studies have shown the European System for Cardiac Operative Risk Evaluation (EuroSCORE) to be an extremely useful and reliable risk stratification score and also a good indicator of quality of care in cardiac surgery. The purpose of this study was to evaluate the significance of the additive and logistic EuroSCOREs in patients undergoing surgery on the thoracic aorta in Japan. **Methods:** We calculated the predicted mortality according to the additive and logistic EuroSCORE algorithms in 327 consecutive patients who underwent surgery of the thoracic aorta during a 30-year period (between 1976 and 2005). We compared the score validity between the two algorithms and also evaluated the score validity for the patients who underwent thoracic aortic surgery. The score validity was assessed by calculating the area under the receiver operating characteristic (ROC) curve. **Results:** The overall in-hospital mortality was 13%. The area under the ROC curve was satisfactorily high for the additive (0.68, 0.73, 0.73) as well as the logistic EuroSCORE (0.69, 0.74, 0.75) in the patients who underwent thoracic aortic surgery during 30-, 20-, and 10-year periods, respectively. The actual mortality was 7% (Group 1; an additive EuroSCORE of 3—6), 16% (Group 2; 7—11), and 37% (Group 3; >12). The mortality expected by the additive and logistic EuroSCORE in the three different risk groups were (5%, 9%, 19%) and (5%, 14%, 43%), respectively. Namely, the mortality expected by the logistic EuroSCORE perfectly matched with the actual mortality in any of the three risk groups. In contrast, the mortality expected by the additive EuroSCORE tended to dissociate when the number of risks increased. Significant difference was observed between the observed mortality and the mortality expected by the additive EuroSCORE algorithm in the high-risk group (p = 0.0473). **Conclusions:** Although both the additive and the logistic EuroSCORE reliably predicted the overall operative mortality for thoracic aortic surgery in 327 Japanese patients, the logistic EuroSCORE better matched with the actual mortality in the operative risk especially in the high-risk group.

© 2006 Published by Elsevier B.V.

**Keywords:** Risk stratification; Mortality; Thoracic aorta; Aneurysm

1. Introduction

The European System for Cardiac Operative Risk Evaluation (EuroSCORE) has been widely used by both patients and medical staff members in Europe to predict mortality from cardiac surgery [1,2]. The predicted mortality (in percent) is calculated by simply adding the weights (from 1 to 4) assigned to each of the 17 factors (simple additive EuroSCORE). Although the additive EuroSCORE was generally found to have an excellent predictive ability, many observers noted a trend to underestimate the operative risk in very high-risk patients [3,4]. Moreover, the predictive value of the additive EuroSCORE, which was originally made for cardiac surgery, may have to be modified [5] or improved [6], especially for thoracic aortic surgery. To reduce the underestimation in very high-risk patients and to easily perform full statistical comparisons to other systems, comprehensive information on the logistic regression equation of the score was published [3], and thus the logistic regression version of the score (the logistic EuroSCORE) could be calculated [3].

The analysis in our 327 patients undergoing surgery of the thoracic aorta over the past 30-year period shows that the logistic EuroSCORE works better than the additive EuroSCORE in predicting the operative risks in the very high-risk patients undergoing the thoracic aortic surgery, although both models are useful for risk stratification in the thoracic aortic surgery.
Table 1
Distribution of thoracic aortic aneurysms over a 30-year period

<table>
<thead>
<tr>
<th>Aneurysm Type</th>
<th>Number of patients</th>
<th>Mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute aortic dissection</td>
<td>80</td>
<td>7 (9%)</td>
</tr>
<tr>
<td>Stanford Type A</td>
<td>76</td>
<td>7 (9%)</td>
</tr>
<tr>
<td>Stanford Type B</td>
<td>4</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Chronic aortic dissection</td>
<td>37</td>
<td>6 (16%)</td>
</tr>
<tr>
<td>Stanford Type A</td>
<td>18</td>
<td>2 (11%)</td>
</tr>
<tr>
<td>Stanford Type B</td>
<td>19</td>
<td>4 (21%)</td>
</tr>
<tr>
<td>True aortic aneurysm</td>
<td>170</td>
<td>26 (15%)</td>
</tr>
<tr>
<td>Ascending aorta</td>
<td>17</td>
<td>2 (13%)</td>
</tr>
<tr>
<td>Aortic arch</td>
<td>100</td>
<td>15 (15%)</td>
</tr>
<tr>
<td>Peri-aortic arch</td>
<td>15</td>
<td>1 (7%)</td>
</tr>
<tr>
<td>Descending aorta</td>
<td>27</td>
<td>7 (26%)</td>
</tr>
<tr>
<td>Thoraco-abdominal aorta</td>
<td>11</td>
<td>1 (9%)</td>
</tr>
<tr>
<td>AAE</td>
<td>37</td>
<td>2 (5%)</td>
</tr>
<tr>
<td>Others</td>
<td>3</td>
<td>1 (33%)</td>
</tr>
<tr>
<td>Total</td>
<td>327</td>
<td>42 (13%)</td>
</tr>
</tbody>
</table>

2.2. Operative procedure

The operative procedures performed on 327 patients included: 113 patients underwent graft replacement of the thoraco-abdominal aorta, and 1 patient underwent an extra anatomic bypass from the descending aorta to the abdominal aorta.

2.3. Method

The preoperative and postoperative data were collected and the logistic and the additive EuroSCORE were calculated. The predicted mortality was compared for the observed or actual mortality for each risk category. Mortality was defined as death from any cause within 30 days of operation or within the same hospital admission. The continuous data were expressed as the mean ± standard deviation, and categorical variables were expressed as percentages. A statistical analysis was conducted using the StatView5.0 software package (SAS Institute Inc., Cary, NC, USA). Comparisons of the two groups were performed for categorical variables using either the chi-square test or Fisher’s exact test as appropriate. The receiver operating characteristic (ROC) curves were plotted for each score system and the area under the ROC curve was calculated as an index for the predictive value of the model. The parametric methods for comparing the two curves were based on the estimates of the binormal parameters associated with each curve, and their variances and covariances. Variables were judged to be significant at a level of p < 0.05.

3. Results

Table 2 shows the distribution of the operative procedures and the in-hospital mortality for each procedure. The mortality in acute type A aortic dissection (9%) is lower than in chronic dissection (16%) and true aortic aneurysm (15%). Because operations for both chronic dissection and true aortic aneurysm were carried out for 30 years but the operation for acute type A dissection was carried out during approximately recent 20 years, this difference might affect the mortality rate of these operations. Table 3 shows the impact of the EuroSCORE risk factors on in-hospital mortality as assessed by a univariate analysis. Five of 17 variables such as a female sex, chronic pulmonary disease, extracardiac arteriopathy, serum creatinine over 200 μmol/l, and a critical preoperative state significantly influenced hospital mortality. A female sex, however, which was related with a lower mortality in the original data of the EuroSCORE [2] was a definite risk factor in the present study. The overall in-hospital mortality was 13% (42 of 327 patients) and the simple hospital mortality of the first decade (1976—1985), the previous decade (1986—1995), and the present decade (1996—2005) was 21%, 16%, and 9%, respectively.

Fig. 1 shows the serial changes of the observed hospital mortality in the first, previous, and present decades. In the low-risk group, the observed mortality in the first decade (20%, p < 0.05) was significantly higher than those of other decades, and it also improved in the previous decade (6%) and has continued to maintain a low level in the present decade (3%). In the moderate-risk group, the observed mortality in the first decade (33%) and the previous decade (22%) was significantly improved in the present decade (10%; p < 0.05 vs other decades). In the high-risk group, no patient was
observed in the first decade and the observed mortality was also quite high in the previous decade (50%) and even in the present decade (33%).

As for the prediction ability of both the additive and logistic EuroSCORE, the ROC curve is shown in Fig. 2A and B. Although the present decade showed a better prediction value (the area under ROC curve) of >0.7 in both algorithms (Fig. 2A and B), the general estimation of risk throughout the 30-year period was appreciably good (0.67 and 0.68) in both algorithms (Fig. 2A and B).

According to the additive EuroSCORE, all patients were divided into three groups (Fig. 1): Group 1 (low-risk group (additive EuroSCORE 3—6), 147 patients (45%) with an actual mortality of 7%); Group 2 (moderate-risk group (additive EuroSCORE 7—11), 161 patients (49%) with an actual mortality of 16%); and Group 3 (high-risk group (additive EuroSCORE >12), 19 patients (6%) with an actual mortality of 37%).

As shown in Fig. 3, the mortality expected by the logistic EuroSCORE perfectly matched with the actual mortality in any of the three risk groups. In contrast, the mortality expected by the additive EuroSCORE tended to dissociate when the number of the risks increased (Fig. 3). In fact, the mortality expected by the additive EuroSCORE algorithm in the high-risk group was significantly (p = 0.0080) lower than that expected by the logistic EuroSCORE algorithm.

4. Discussion

Recently, Barmettler et al. [5] reported that a slight modification of the EuroSCORE successfully improved the predictive value of the EuroSCORE for thoracic aortic surgery.
In addition, Kawachi et al. [6] reported a relatively low accuracy (area under the ROC curve = 0.61) of the additive EuroSCORE for thoracic aortic surgery. From the present study, in contrast, both the additive and logistic models of EuroSCORE were generally validated in the thoracic aortic surgery for estimating the operative risk (Fig. 2A and B). In the high-risk group, however, the mortality expected by the additive EuroSCORE tended to dissociate, while the logistic EuroSCORE matched very well with the actual mortality.

The benefit of the additive EuroSCORE is its simple way to evaluate the operative risk in individual patients by themselves. They or their doctors can calculate the risk by mental arithmetic or 'on the back of an envelope'. In the logistic EuroSCORE, the predicted mortality is given by the relatively complex formula [3]. Such a risk calculation, however, can be used or downloaded easily from the EuroSCORE website (http://www.euroscore.org). Therefore, both the patients and cardiologists (or cardiac surgeons) can estimate the operative risk.

Seventeen variables have been reported to be associated with an increased mortality in the EuroSCORE [1,2]. Five out of these 17 variables, namely, a female sex, chronic pulmonary disease, extracardiac arteriopathy, a serum creatinine over 200 \( \mu \text{mol/l} \), and a critical preoperative state, significantly influenced hospital mortality in the present study. A female sex, however, was found to be a factor for a lower mortality rate in our study. Although a female gender has been reported to be a significant risk factor for hospital mortality in the original report [7], Kawachi et al. [6] reported no significant differences in the hospital mortality and morbidity between the two genders, and female octogenarians have also shown a significantly better actuarial survival than males [8].

It is well known that the operative mortality for the thoracic aneurysm has improved because of the technical modification. Thus, it might not be feasible to apply the EuroSCORE for operations carried out during 30 years. Simple crude mortality figures, however, are no longer sufficient to assess the quality control of any treatment at an institute. Especially, in order to improve quality of thoracic aortic surgery of one institute, it is necessary to verify the operative results using a simple but reliable algorism. Although the simple hospital mortality in our patients has gradually improved during the first (21%), previous (15%), and present (9%) decades, the risk stratification according to the EuroSCORE showed more information from the view point of quality control indicating that hospital mortality of the low- and moderate-risk patients has already improved during past 20 years; however, we have to still improve mortality of the high-risk patients.
One limitation of this study was the relatively small number of patients especially in the high-risk group patients and the first decade patients from a single institute. Because the number of high-risk patients may continue to increase, we should continue to collect data to verify the observations of this study. Although the logistic model has been recognized as a better model for high-risk patients group [3, 4], we demonstrated that the EuroSCORE was a valuable model for measuring the quality of surgical care in aortic surgery in Japan.

In conclusion, although both the additive and the logistic EuroSCOREs reliably predicted the operative mortality for thoracic aortic surgery, the logistic EuroSCORE was better than the additive EuroSCORE especially in predicting the operative risk in the very high-risk group. Risk stratification using the EuroSCORE helped us to improve the quality control of surgical treatment for thoracic aneurysms.

Acknowledgment

We thank Mr Brian Quinn for carefully reviewing of this manuscript.

References


Editorial comment

EuroSCORE and the Japanese aorta

In this paper, [1] Nishida and colleagues have reviewed 327 patients who have undergone surgery of the thoracic aorta in their hospital and stratified them using EuroSCORE (both logistic and additive versions). They found that both models work well in this patient population but that the logistic model is superior in the higher risk group. They also found that the logistic model remains well calibrated in this Japanese thoracic aortic surgical patient group.

The study has some weaknesses. It is from a single institution and therefore, when such studies purport to assess a risk model they may end up telling us more about the institution and its performance than about the risk model itself. Some of the data collection must have been retrospective as EuroSCORE was not available at the beginning of the thoracic aortic surgical experience as reported in the study. The authors duly recognise the limitation of the relatively small number of patients. They are perhaps a little too self-critical in that respect: the paper deals with a highly selected subset of cardiac surgical procedures and 327 patients is a very respectable number of operations on the thoracic aorta. Finally, the conclusion that the logistic model works better in high-risk patients is one that is already well known and documented in general cardiac surgery, so it is not surprising to find this fact confirmed in surgery of the thoracic aorta. Indeed, thoracic aortic surgery is in itself a relatively higher risk subgroup of cardiac surgery; so logistic models in general are better suited to such a patient population than additive models [2].

The interesting findings in this study are that, even within a small patient subgroup such as this, we have a risk model which is still capable of differentiating risk strata and that the model remains well calibrated in this cohort. This is despite several recent reports of overscoring and claims that the model may soon be out of date. That EuroSCORE remains valuable for measuring the quality of surgical care in aortic surgery in Japan and that its discriminatory power has, if anything, increased with time over the last three decades are new findings which are both interesting and satisfying. Nevertheless, it has to be recognised that the data from which EuroSCORE was originally constructed are now 10 years old. The results of cardiac surgery have changed in the last decade. There is evidence from the United Kingdom [3] and elsewhere that around the year 2002, cardiac surgical outcomes have improved significantly despite a general tendency towards operating on older and sicker patients. Undoubtedly, the time has come for a re-evaluation of the risk model and the EuroSCORE project group are in the process of developing this re-evaluation. The current tendency is towards a repeat exercise in data collection which may be worldwide rather than confined to European countries. Institutions with an interest in participating in this venture are encouraged to register their interest with the project group.

Finally, the authors conclude that risk stratification has helped them improve the quality control of thoracic aortic surgery. I agree with them and would go even further: