The novel EuroSCORE II algorithm predicts the hospital mortality of thoracic aortic surgery in 461 consecutive Japanese patients better than both the original additive and logistic EuroSCORE algorithms

Takahiro Nishida*, Hiromichi Sonoda, Yasuhsia Oishi, Yoshihsia Tanoue, Atsuhiro Nakashima, Yuichi Shiokawa and Ryuji Tominaga

Department of Cardiovascular Surgery, Kyushu University Graduate School of Medical Sciences, Fukuoka, Japan

* Corresponding author: Department of Cardiovascular Surgery, Kyushu University Graduate School of Medical Sciences, 3-1-1 Maidashi Higashi-ku, Fukuoka 812-8582, Japan. Tel: +81-92-6425557; fax: +81-92-6425566; e-mail: tnishida@heart.med.kyushu-u.ac.jp (T. Nishida).

Received 13 August 2013; received in revised form 4 November 2013; accepted 18 November 2013

Abstract

OBJECTIVES: The European System for Cardiac Operative Risk Evaluation (EuroSCORE) II was developed to improve the overestimation of surgical risk associated with the original (additive and logistic) EuroSCOREs. The purpose of this study was to evaluate the significance of the EuroSCORE II by comparing its performance with that of the original EuroSCOREs in Japanese patients undergoing surgery on the thoracic aorta.

METHODS: We have calculated the predicted mortalities according to the additive EuroSCORE, logistic EuroSCORE and EuroSCORE II algorithms in 461 patients who underwent surgery on the thoracic aorta during a period of 20 years (1993–2013).

RESULTS: The actual in-hospital mortality rates in the low- (additive EuroSCORE of 3–6), moderate- (7–11) and high-risk (≥11) groups (followed by overall mortality) were 1.3, 6.2 and 14.4% (7.2% overall), respectively. Among the three different risk groups, the expected mortality rates were 5.5 ± 0.6, 9.1 ± 0.7 and 13.5 ± 0.2% (9.5 ± 0.1% overall) by the additive EuroSCORE algorithm, 5.3 ± 0.1, 16 ± 0.4 and 42.4 ± 1.3% (19.9 ± 0.7% overall) by the logistic EuroSCORE algorithm and 1.6 ± 0.1, 5.2 ± 0.2 and 18.5 ± 1.3% (7.4 ± 0.4% overall) by the EuroSCORE II algorithm, indicating poor prediction (P < 0.0001) of the mortality in the high-risk group, especially by the logistic EuroSCORE. The areas under the receiver operating characteristic curves of the additive EuroSCORE, logistic EuroSCORE and EuroSCORE II algorithms were 0.6937, 0.7169 and 0.7697, respectively. Thus, the mortality expected by the EuroSCORE II more closely matched the actual mortality in all three risk groups. In contrast, the mortality expected by the logistic EuroSCORE overestimated the risks in the moderate- (P = 0.0002) and high-risk (P < 0.0001) patient groups.

CONCLUSIONS: Although all of the original EuroSCOREs and EuroSCORE II appreciably predicted the surgical mortality for thoracic aortic surgery in Japanese patients, the EuroSCORE II best predicted the mortalities in all risk groups.

Keywords: Risk stratification • Hospital mortality • Aortic surgery

INTRODUCTION

The original additive European System for Cardiac Operative Risk Evaluation (EuroSCORE) was established in 1999 [1, 2] to predict the hospital mortality after cardiac surgery, and has been widely used by both patients and medical staff not only for risk prediction, but also for the assessment of the quality of cardiac surgery. However, several reports [3–5] have demonstrated the overestimation of the hospital mortality by the logistic EuroSCORE. The EuroSCORE II [6] is a new algorithm used for estimating hospital mortality after cardiac surgery and was updated in 2011 to correct for the overestimation of hospital mortality by the logistic EuroSCORE.

Although the risk stratification in thoracic aortic surgery is a topic of major interest, it is not easy to develop optimal algorithms for predicting hospital mortality for thoracic aortic surgery because of the variety of aortic diseases, surgical methods and corresponding lesions of the aorta. In fact, the hospital mortality for thoracic aortic surgery cannot be calculated by the Society of Thoracic Surgery (STS) risk calculator [7]. In contrast, the original (both additive and logistic) EuroSCOREs have provided relatively good predictions of hospital mortality for thoracic aortic surgery. In 2006 [8], we reported the superiority of the logistic EuroSCORE over the additive EuroSCORE in predicting mortality, especially in the high-risk group after thoracic aortic surgery in Japanese patients. Recently, the mortality rate after thoracic aortic surgery has decreased due to improvements in surgical techniques, artificial grafts and other materials used. Therefore, the risk stratification algorithms should be modified further and improved in order to more accurately predict hospital mortality after cardiac surgery [6, 9].

We evaluated the newly developed EuroSCORE II for predicting mortality rates after thoracic aortic surgery in our Japanese patients for the past 20 years.
Patients and Methods

Patients

Between January 1993 and January 2013, 461 consecutive patients underwent surgery for thoracic aortic disease at our institution. The mean age of the patients was 63.5 ± 0.7 years and 163 (35%) were female. Table 1 shows the distribution of the various aortic diseases. One hundred and thirty-five (29.3%) patients suffered from acute aortic dissection, and 51 (11.1%) had chronic aortic dissection. True and pseudo-aortic aneurysms were observed in 221 (47.9%) patients. Other diseases including primary or metastatic aortic tumours, traumatic aortic rupture and infected aneurysms were observed in 38 (8.2%) patients. Other diseases including primary or metastatic aortic tumours, traumatic aortic rupture and infected aneurysms were observed in 3 (0.65%), 5 (1.1%) and 1 (0.22%) patients, respectively.

Surgical procedure

Table 2 shows the surgical procedures performed in the 461 patients. We performed graft replacement of the ascending aorta in 84 patients, the aortic root in 69 and the aortic arch in 220. Descending aorta replacement was carried out in 51 patients. Graft replacement of the thoraco-abdominal aorta was performed in 30 patients.

Methods

The pre- and postoperative data were collected, and the additive and logistic EuroSCOREs and EuroSCORE II were used to calculate the predicted mortality rates. The predicted mortality rates were compared with the observed actual mortality for each risk category. The actual mortality was strictly defined as follows: death in the same hospital or at another hospital but before discharge from the hospital or death within 90 days of surgery, regardless of the location [6]. We used the JMP software (SAS Institute Japan Ltd, Tokyo, Japan) for the statistical analysis. Continuous data were expressed as the means ± standard error of the mean, and categorical variables were expressed as percentages. Comparisons of the two groups were performed for categorical variables using either the $\chi^2$ 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 (AUC) was calculated as an index for the predictive value of the model. A value of $P < 0.05$ was considered to be significant.

Results

Table 2 shows the distribution of the surgical procedures and the observed hospital mortality for each procedure. Table 3 shows the impact of the EuroSCORE II risk factors on in-hospital mortality as assessed by univariate analysis with the $\chi^2$ tests. Seven variables were found to be significant risk factors, including the level of renal impairment (severe), extracardiac arteriopathy, a critical preoperative state, New York Heart Association (NYHA) Class IV, recent myocardial infarction, urgency, and weight of the intervention (two procedures). There were 33 hospital deaths, so the actual overall hospital mortality rate was 7.2%.

As shown in Fig. 1, the predicted mortality in the overall cases was significantly higher ($P < 0.0001$) in comparison with the predicted mortality calculated by either the additive EuroSCORE (9.5 ± 0.1%) or the EuroSCORE II (7.4 ± 0.4%), or with the observed actual mortality (7.3%).

According to the additive EuroSCORE, all 461 patients were divided into the following three groups (Fig. 1): the low-risk group (additive EuroSCORE 3–6) including 75 (16%) patients with an actual mortality rate of 1.3%, the moderate-risk group (additive EuroSCORE 7–11) including 289 (63%) patients with an actual mortality rate of 6.6% and the high-risk group (additive EuroSCORE >11) including 97 (21%) patients with an actual mortality rate of 14.4%. As shown in Fig. 1, the mortality rate expected by the EuroSCORE II perfectly matched the actual mortalities in all three risk groups. In contrast, the mortality expected by the logistic EuroSCORE tended to become dissociated when the grade of the risk increased. In fact, significant differences were observed between the mortality rate expected by
the logistic EuroSCORE and the actual mortality rate in both the moderate- (P = 0.0002) and high-risk groups (P < 0.0001).

With regard to the predictive value of each EuroSCORE algorithm, the ROC curves for the additive (Fig. 2A), logistic EuroSCORE (Fig. 2B) and EuroSCORE II (Fig. 2C) are shown in Fig. 2. Although no statistically significant differences were found among the three ROC curves calculated by the three algorithms, the AUC was over 0.75 only for the EuroSCORE II prediction (AUC = 0.7697; Fig. 2C), whereas it was <0.75 for both the additive EuroSCORE (AUC = 0.6937; Fig. 2A) and the logistic EuroSCORE (AUC = 0.7169; Fig. 2B) predictions. These results indicate that the EuroSCORE II could best predict the hospital mortality in Japanese patients undergoing thoracic aortic surgery at our institution.

DISCUSSION

In this report, we have clearly shown that the EuroSCORE II algorithm predicted the surgical mortality of thoracic aortic surgery in Japanese patients more suitably than did the logistic EuroSCORE algorithm. Although the logistic EuroSCORE overestimated the mortality of the high-risk patient group, the EuroSCORE II provided more accurate predicted mortality rates in all three risk groups. Moreover, only the AUC of the EuroSCORE II was >0.75.

Risk stratification of cardiac surgery is a topic of interest; however, it is not easy to establish a simple but reliable algorithm.

Table 3: Impact of the seven EuroSCORE II risk factors on hospital death in 461 patients with thoracic aortic diseases

<table>
<thead>
<tr>
<th>Risk factor</th>
<th>With risk factor</th>
<th>Without risk factor</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Death/number of patients</td>
<td>Mortality (%)</td>
<td>Death/number of patients</td>
</tr>
<tr>
<td>Renal impairment (severe)</td>
<td>15/123</td>
<td>12.2</td>
<td>17/329</td>
</tr>
<tr>
<td>Extracardiac arteriopathy</td>
<td>21/121</td>
<td>17.4</td>
<td>12/340</td>
</tr>
<tr>
<td>Critical preoperative state</td>
<td>16/26</td>
<td>61.5</td>
<td>17/435</td>
</tr>
<tr>
<td>NYHA (IV)</td>
<td>20/163</td>
<td>12.3</td>
<td>13/298</td>
</tr>
<tr>
<td>Recent myocardial infarction</td>
<td>4/10</td>
<td>40.0</td>
<td>29/451</td>
</tr>
<tr>
<td>Emergency</td>
<td>20/163</td>
<td>12.3</td>
<td>13/298</td>
</tr>
<tr>
<td>Weight of the intervention (two procedures)</td>
<td>13/78</td>
<td>16.7</td>
<td>18/377</td>
</tr>
</tbody>
</table>

Figure 1: The actual and predicted mortality rates calculated using the EuroSCORE II, additive EuroSCORE and logistic EuroSCORE in the overall cases, low-risk group (additive EuroSCORE 3–6), moderate-risk group (additive EuroSCORE 7–11) and high-risk group (additive EuroSCORE >11). Significantly (P < 0.0001) higher mortality rates were predicted by the logistic EuroSCORE in both the overall cases and the high-risk group in comparison with either the actual mortality rate or the predicted mortality rates calculated by the additive EuroSCORE or EuroSCORE II. In the moderate-risk group, only the predicted mortality rate estimated by the logistic EuroSCORE was significantly (P = 0.0002) higher than the actual mortality rate.

Figure 2: The ROC curves of the additive EuroSCORE (A), logistic EuroSCORE (B) and EuroSCORE II (C). The AUCs were 0.6937, 0.7169 and 0.7697, respectively.
to predict the surgical mortality in patients undergoing thoracic aortic surgery, because there are various aortic diseases in various locations. To establish the original (additive and logistic) EuroSCOREs [2], the majority of the patients collected for these projects were suffering from coronary artery disease, while only a few of the patients underwent ascending aortic surgery. Despite the fact that only a few aortic surgery cases were collected, the previous two original EuroSCOREs could provide good mortality prediction rates for thoracic aortic surgery due to the addition of the variable of ‘surgery on thoracic aorta’.

Some reports, including those by our group [8, 10], have already shown the superior validity of the logistic EuroSCORE for the prediction of surgical mortality in thoracic aortic surgery, although other reports [11–13] have insisted that some modifications are required to further improve the validity of the logistic EuroSCORE.

In this context, Barmettler et al. have reported that it was important to take both the aortic dissection and malperfusion prior to thoracic aortic surgery into account to improve the predictive value of the logistic EuroSCORE. On the other hand, Matsuura et al. [13] reported that the subtraction of the age contribution from the original EuroSCOREs improved the accuracy of the predicted mortality rates. Although the EuroSCORE II does not list either malperfusion or circulatory arrest time, it could satisfactorily predict the morality of patients who have undergone thoracic aortic surgery in our 461 Japanese patients. This may suggest that the well-known risk factors of aortic surgery, such as malperfusion and circulatory arrest time, are not mandatory in predicting patient mortality for aortic surgery. Preoperative patient condition due to particular malperfusion in some organs may have been reflected on other variables such as recent myocardial infarction, critical preoperative state or poor mobility in the EuroSCORE II algorithm.

It is well known that mortality during and after cardiac surgery has improved due to improvements in surgical methods, including the surgeon’s skills, in anaesthesia and in postoperative care. Thus, the risk prediction model should also be consistently improved. From this viewpoint, the additive EuroSCORE, originally established in 1999 [1, 2], was modified and improved to the logistic EuroSCORE in 2003 [14], because the additive EuroSCORE had tended to underestimate the risk in high-risk patients. As we have previously reported in 2006 [8], the logistic EuroSCORE was indeed better than the additive EuroSCORE, especially for predicting the surgical risk in the high-risk group of patients undergoing thoracic aortic surgery. Our results [8] are also compatible with the recent consensus that the additive EuroSCORE underestimates the mortality of high-risk patients in general cardiac surgery.

Recently, some reports [3, 5, 11] have shown that the second-generation algorithm of the logistic EuroSCORE conversely tends to overpredict the surgical risks in the high-risk group because of the improvements in cardiac surgery. Thus, these two original EuroSCOREs were updated to generate the EuroSCORE II in 2011 [6]. The four newly added variables were Canadian Cardiovascular Society Class 4 angina, insulin-dependent diabetes mellitus, creatinine clearance calculated by the Cockcroft-Gault formula and the weight of procedure. Owing to these modifications, the EuroSCORE II is far better calibrated than the original two models [15, 16]. In contrast to the small number of patients included in the original EuroSCOREs, a total of 1637 (7.3%) patients who underwent thoracic aortic surgery were included in the patient cohort to establish the EuroSCORE II [6]. From this sufficiently large patient cohort, all aortic operations from the aortic root to thoraco-abdominal repair, from dissection to traumatic aneurysms, are able to be handled as one group using the EuroSCORE II. Although better accuracy for predicting the risk of thoracic aortic surgery is thus expected from the EuroSCORE II, no reports proving the validity of this third-generation algorithm for thoracic aortic surgery have been published so far.

To the best of our knowledge, our present study is the first to confirm that the EuroSCORE II predicted the surgical mortality of thoracic aortic surgery more accurately than the original EuroSCOREs. The 17 variables are all mandatory in the final EuroSCORE II algorithm [6] in order to predict the risks. Seven of these 17 variables, namely, severe renal impairment, extracardiac arteriopathy, critical preoperative state, NYHA Class IV, recent myocardial infarction, emergency and two procedures, significantly influenced the hospital mortality rate in our present study. Interestingly, the highest risk of mortality was observed in the patients with ‘severely impaired renal function’, but a high level of risk was not observed in the patients on haemodialysis in the present study. Similar findings have already been reported in the study establishing the EuroSCORE II [6].

Recently, JapanSCORE has been developed by analysing the data from the Japan Adult Cardiovascular Surgery Database [17]. The average mortality calculated by the JapanSCORE in our most recent 200 cases (among the 461) of thoracic aortic disease was 8.0 ± 0.5%, whereas the average mortality calculated by the EuroSCORE II was 7.2 ± 0.6% in the same 200 patients (P = 0.2868; not significant by paired t-test). The JapanSCORE is expected to be more suitable for our Japanese patients because this algorithm is based on the Japanese database. Further study is mandatory to confirm whether the JapanSCORE is better than the other algorithms including the EuroSCORE II.

The simple crude mortality rate may not be sufficient to assess the quality controls of high-risk operations. To improve the quality of thoracic aortic surgery, for example, it is necessary to analyse and verify the results using a reliable algorithm. In our institute, the simple in-hospital mortality rate of thoracic aortic surgery gradually improved from 15% during the period from 1986–1995 to 9% in the period 1996–2005 and to 7.2% in the period 2006–2013. Based on the risk stratification according to the additive EuroSCORE, the mortality in the high-risk patients (additive EuroSCORE >11) was especially poor (50% in 1986–1995 and 33% in 1996–2005) in our institute [8] and it decreased to only 8.5% (4 of 47) in the period 2006–2013. Thus, not only the analysis of the simple mortality rate, but also the risk stratification, apparently clarify where we should focus our efforts in terms of the type of patients who should undergo thoracic aortic surgery in order to improve the overall surgical results.

In conclusion, based on our present study, the novel algorithm of the EuroSCORE II provided satisfactorily accurate mortality predictions in all risk groups of Japanese patients undergoing thoracic aortic surgery, while the logistic EuroSCORE overestimated the mortality rate, especially in the high-risk group. The use of this excellent algorithm may provide not only good risk prediction, but also a way to improve the quality of the overall surgical results.

ACKNOWLEDGEMENTS

We thank Brian Quinn for reviewing this manuscript.

Conflict of interest: none declared.

REFERENCES

surgeons should calculate the risk-adjusted mortality ratio (RAMR) by dividing the actual mortality (observed) by the expected mortality. Then, the predicted mortality calculated by EuroSCORE II is multiplied by the unit’s or the individual surgeon’s RAMR. This reflects in a more faithful manner the current mortality for a given patient in a given surgical unit. If, for example, the unit’s RAMR for thoracic aortic surgery is 2, and the predicted mortality by EuroSCORE II is 7.4%, the corrected mortality rate for thoracic aortic surgery for this hypothetical unit would be 14.8%. In consequence, this is a more effective way to adapt the EuroSCORE II to the ‘real world’.

Conflict of Interest: none declared

References


Conflict of Interest: none declared

References


© The Author 2014. Published by Oxford University Press on behalf of the European Association for Cardio-Thoracic Surgery. All rights reserved.

Nishida et al. need to be congratulated for achieving an overall mortality rate of 7.2% in 461 patients undergoing aortic surgery [1], however, their manuscript raises a number of issues. Mortality prediction models are classically assessed via receiver operating curve (ROC) and Hosmer-Lemeshow statistic assessment [2,3]. The use of a 20-year study period is probably too long, due to improvements in surgical, anaesthetic and medical care, hence the recalibration of EuroSCORE to create EuroSCORE II. With such a long study period the chronobiology of aortic dissection may be an important issue [4,5]. We agree with their interpretation that a ROC of nearly 0.8 is impressive for the EuroSCORE II risk model in aortic surgery in their patient group; however their demonstration that in high risk patients none of the models are accurate means clinical usage is limited. Unfortunately no Hosmer-Lemeshow statistic was presented. Failure to achieve an adequate ROC and Hosmer-Lemeshow statistic means adoption can not be recommended.

The operation types for this study seem skewed, as 220 patients underwent isolated arch surgery, and only 7 patients underwent root and arch surgery. In addition no mention of redo numbers were made in the manuscript.

Risk modelling for CABG or valve surgery requires 10,000 to 20,000 patients per procedure to be operated on in the modern era to avoid being underpowered. As aortic surgery can be simplicistically divided into root, arch and descending aorta surgery, risk modelling for aortic surgery is probably beyond any single institution due to number restrictions. An international collaborative project is needed.

Conflict of Interest: none declared

References


© The Author 2014. Published by Oxford University Press on behalf of the European Association for Cardio-Thoracic Surgery. All rights reserved.

Nishida et al. need to be congratulated for achieving an overall mortality rate of 7.2% in 461 patients undergoing aortic surgery [1], however, their manuscript raises a number of issues. Mortality prediction models are classically assessed via receiver operating curve (ROC) and Hosmer-Lemeshow statistic assessment [2,3]. The use of a 20-year study period is probably too long, due to improvements in surgical, anaesthetic and medical care, hence the recalibration of EuroSCORE to create EuroSCORE II. With such a long study period the chronobiology of aortic dissection may be an important issue [4,5]. We agree with their interpretation that a ROC of nearly 0.8 is impressive for the EuroSCORE II risk model in aortic surgery in their patient group; however their demonstration that in high risk patients none of the models are accurate means clinical usage is limited. Unfortunately no Hosmer-Lemeshow statistic was presented. Failure to achieve an adequate ROC and Hosmer-Lemeshow statistic means adoption can not be recommended.

The operation types for this study seem skewed, as 220 patients underwent isolated arch surgery, and only 7 patients underwent root and arch surgery. In addition no mention of redo numbers were made in the manuscript.

Risk modelling for CABG or valve surgery requires 10,000 to 20,000 patients per procedure to be operated on in the modern era to avoid being underpowered. As aortic surgery can be simplicistically divided into root, arch and descending aorta surgery, risk modelling for aortic surgery is probably beyond any single institution due to number restrictions. An international collaborative project is needed.

Conflict of Interest: none declared

References


© The Author 2014. Published by Oxford University Press on behalf of the European Association for Cardio-Thoracic Surgery. All rights reserved.

Nishida et al. need to be congratulated for achieving an overall mortality rate of 7.2% in 461 patients undergoing aortic surgery [1], however, their manuscript raises a number of issues. Mortality prediction models are classically assessed via receiver operating curve (ROC) and Hosmer-Lemeshow statistic assessment [2,3]. The use of a 20-year study period is probably too long, due to improvements in surgical, anaesthetic and medical care, hence the recalibration of EuroSCORE to create EuroSCORE II. With such a long study period the chronobiology of aortic dissection may be an important issue [4,5]. We agree with their interpretation that a ROC of nearly 0.8 is impressive for the EuroSCORE II risk model in aortic surgery in their patient group; however their demonstration that in high risk patients none of the models are accurate means clinical usage is limited. Unfortunately no Hosmer-Lemeshow statistic was presented. Failure to achieve an adequate ROC and Hosmer-Lemeshow statistic means adoption can not be recommended.

The operation types for this study seem skewed, as 220 patients underwent isolated arch surgery, and only 7 patients underwent root and arch surgery. In addition no mention of redo numbers were made in the manuscript.

Risk modelling for CABG or valve surgery requires 10,000 to 20,000 patients per procedure to be operated on in the modern era to avoid being underpowered. As aortic surgery can be simplicistically divided into root, arch and descending aorta surgery, risk modelling for aortic surgery is probably beyond any single institution due to number restrictions. An international collaborative project is needed.

Conflict of Interest: none declared

References


© The Author 2014. Published by Oxford University Press on behalf of the European Association for Cardio-Thoracic Surgery. All rights reserved.

Nishida et al. need to be congratulated for achieving an overall mortality rate of 7.2% in 461 patients undergoing aortic surgery [1], however, their manuscript raises a number of issues. Mortality prediction models are classically assessed via receiver operating curve (ROC) and Hosmer-Lemeshow statistic assessment [2,3]. The use of a 20-year study period is probably too long, due to improvements in surgical, anaesthetic and medical care, hence the recalibration of EuroSCORE to create EuroSCORE II. With such a long study period the chronobiology of aortic dissection may be an important issue [4,5]. We agree with their interpretation that a ROC of nearly 0.8 is impressive for the EuroSCORE II risk model in aortic surgery in their patient group; however their demonstration that in high risk patients none of the models are accurate means clinical usage is limited. Unfortunately no Hosmer-Lemeshow statistic was presented. Failure to achieve an adequate ROC and Hosmer-Lemeshow statistic means adoption can not be recommended.

The operation types for this study seem skewed, as 220 patients underwent isolated arch surgery, and only 7 patients underwent root and arch surgery. In addition no mention of redo numbers were made in the manuscript.

Risk modelling for CABG or valve surgery requires 10,000 to 20,000 patients per procedure to be operated on in the modern era to avoid being underpowered. As aortic surgery can be simplicistically divided into root, arch and descending aorta surgery, risk modelling for aortic surgery is probably beyond any single institution due to number restrictions. An international collaborative project is needed.

Conflict of Interest: none declared

References
