Preoperative risk prediction and intraoperative events in cardiac surgery

Serban C. Stoica\textsuperscript{a}, Linda D. Sharples\textsuperscript{b}, Ishtaq Ahmed\textsuperscript{a}, François Roques\textsuperscript{c}, Stephen R. Large\textsuperscript{a}, Samer A.M. Nashef\textsuperscript{a,}\textsuperscript{*}

\textsuperscript{a}Department of Cardiothoracic Surgery, Papworth Hospital, Papworth Everard, Cambridge CB3 8RE, UK
\textsuperscript{b}MRC Biostatistics Unit, Cambridge CB2 2SR, UK
\textsuperscript{c}CHU Fort de France, Martinique, France

Received 2 August 2001; received in revised form 14 October 2001; accepted 19 October 2001

Abstract

Objective: To examine the relationship between preoperative risk prediction and intraoperative events. Methods: A total of 3118 patients operated in 1999 and 2000 at our institution were analysed, all of whom had their EuroSCORE collected prospectively. The intraoperative variables studied were consultant or trainee operating, long bypass time, long ischaemic time, return on bypass in theatre and use of intra-aortic balloon pump at the end of the procedure. The outcomes are reported as hospital mortality, prolonged length of stay in the intensive therapy unit (pLOS-ITU, >48 h) and death or pLOS-ITU. Risk models were constructed by logistic regression for predicting these three outcomes. Results: With the exception of prolonged cross-clamp time, all variables analysed were independently predictive of a negative outcome. Trainee operating had an apparent protective effect. All risk models performed well. The area under the receiver operating characteristic (ROC) curve (95% CI) increased from 0.857 (0.81, 0.90) for EuroSCORE to 0.874 (0.83, 0.92) for the risk of death model. Similarly, the area under the ROC curve for the pLOS-ITU model increased from 0.687 (0.642, 0.732) to 0.734 (0.691, 0.777) and for the death or pLOS-ITU model from 0.717 (0.677, 0.756) to 0.757 (0.719, 0.795). Conclusions: Knowledge of adverse intraoperative events enhances preoperative risk prediction. This type of analysis could be used for identifying ‘near miss’ outcomes in adult cardiac surgery.

Keywords: EuroSCORE; Mortality; Length of intensive therapy; Near miss

1. Introduction

Risk stratification in cardiac surgery influences clinical decision-making, patient consent for procedures, audit, training and resource planning. Mortality and morbidity have been the focus of many models based on preoperative, intraoperative and postoperative patient factors or a combination of them. De Leval and colleagues introduced surgical human factors and the concept of near miss into this equation. This is a negative outcome or an equivalent of failure defined as a severe temporary or permanent complication [1]. Using the neonatal arterial switch operations as an index procedure, these authors prospectively investigated a number of perioperative human factors leading to major or minor errors. The multiplicative nature of such negative events is counteracted by the ability of the team to compensate for them and this interaction directly influences the outcome.

Extrapolation of these concepts to adult cardiac surgery is reasonable and necessary. Although it is intuitive that preoperative risk is modified by subsequent events in the operating room and in the intensive therapy unit (ITU), the magnitude of this interaction has not been directly quantified. In an extreme example, a cohort of patients with zero risk of death for coronary revascularisation, according to recognised scores, may experience a high mortality if each procedure is associated with major negative events, be those human errors (e.g. inexperienced surgical team), system errors (e.g. power failure) or a combination of them. Identification of negative events and correction of their causes may be one of the major ways further to improve cardiac surgical outcomes whilst continuing to offer good training. Furthermore, if preoperative outcome prediction is enhanced by adverse events in theatre and in ITU, than this information would be of interest to many, including individual patients, their families and hospital staff (for example by helping to plan ITU bed occupancy). In this

\textsuperscript{*} Corresponding author. Tel.: +44-1480-830-541; fax: +44-1480-364-744.
E-mail address: sam.nashef@papworth-tr.anglo.nhs.uk
(S.A.M. Nashef).

1010-7940/02/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S1010-7940(01)01077-6
paper we set out to describe the association between preoperative risk assessment and adverse intraoperative events in our patient population as a first step for identifying near miss outcomes.

2. Patients and methods

2.1. Patient population

We evaluated all patients who underwent cardiac operations at our institution between January 1999 and December 2000. EuroSCORE was collected prospectively for all cases and used as a preoperative indicator of the patients’ risk profile. The operative variables analysed were: consultant or trainee operating, cardiopulmonary bypass (CPB) time, cross-clamp (XC) time, return on CPB in theatre and requirement for mechanical support (intra-aortic balloon pump) at the end of the procedure. There were three outcome measures: hospital death, prolonged stay in the ITU (pLOS-ITU, >48 h) as a surrogate marker of morbidity, and the combined outcome of death or pLOS-ITU. A pilot analysis on 2691 consecutive cases showed that the length of anaesthetic time does not influence either of these outcomes and so this variable was not included in the current study [2]. Procedures for which EuroSCORE was thought to be an inappropriate descriptor of risk in this context were excluded: operations not involving cardiopulmonary bypass, all transplants and bridges to transplant, and elective pulmonary thromboendarterectomy. There were no cases of myocardial protection by ventricular fibrillation. Patients who needed a second operation during the study period were retained if there was a new indication (e.g. blocked graft, valvular leak) and such cases were entered in the study with their new EuroSCORE. Cases of return to theatre for immediate complications (e.g. bleeding, tracheostomy) did not represent a new indication. After the selection process, 3118 cases were retained. Twenty-nine patients who died in theatre during the initial procedure or in the first 48 h in the ITU were included in the death analyses but not in the pLOS-ITU analyses.

2.2. Statistical analysis

Prolonged cardiopulmonary bypass time (pCPB) and prolonged cross-clamp time (pXC) were defined as exceeding the 75th centile of their distribution. LOS-ITU was considered prolonged when in excess of 48 h. A risk model incorporating both EuroSCORE and the intraoperative variables was created. Patients operated on in 1999 (n = 1575) were used as a derivation data set and the patients treated in 2000 (n = 1543) represented the validation set. Multivariate analysis by stepwise forward logistic regression was performed for the three outcomes studied. The Hosmer-Lemeshow goodness of fit χ² test was used to evaluate the calibration of the models. In addition, their ability to discriminate between patients with or without the outcome is described by the area under the receiver-operating characteristic (ROC) curve.

3. Results

Fig. 1 shows the distribution of procedures and Fig. 2 illustrates the case mix by surgical team for all the patients. The overall hospital mortality was 132 (4.2%) and 319 (10.2%) patients had pLOS-ITU. Table 1 presents the frequency of the variables studied in the derivation set and in the test set. Overall 82 patients returned on CPB in theatre, 153 required mechanical support at the end of the procedure and 1235 (39.6%) were operated by a trainee (28.8% assisted by a consultant and 11.8% independently).

Table 2 presents the risk-adjusted odds ratios for death, pLOS-ITU, and death or pLOS-ITU. For every percentage increase in EuroSCORE, the odds ratios for these three outcomes were 1.22, 1.17 and 1.19, respectively (P < 0.001). Trainee operating did not influence survival but had a protective effect in terms of pLOS-ITU (P < 0.001). Similarly, long ischaemic time did not influence either of these outcomes. Prolonged bypass time, return on bypass and use of mechanical support all increase the risk of death or pLOS over and above what is evaluated.
preoperatively \((P < 0.001)\). The odds ratios in the table resulted after logistic regression, i.e. every risk factor is adjusted for all the other ones, including EuroSCORE. The results of model testing in the validation set are illustrated in Table 3, and Fig. 3 shows the ROC curves for the model compared to the predictive ability of EuroSCORE alone. The area under the ROC curve (95% CI) increased from 0.857 (0.81, 0.90) for EuroSCORE to 0.874 (0.83, 0.92) for the risk of death model (Fig. 3a). Similarly, the area under the ROC curve for the pLOS-ITU model increased from 0.687 (0.642, 0.732) to 0.734 (0.691, 0.777) and for the death or pLOS-ITU model from 0.717 (0.677, 0.756) to 0.757 (0.719, 0.795) (Fig. 3b,c).

4. Discussion

Imagination in the quality of care is a priority on a good number of health policy agendas. After realising that crude mortality rates do not reflect the complexity of surgical outcomes, a number of risk models were devised to allow risk-adjusted comparisons. A common feature for many of them is that they are based on patient factors and on other technical variables related to the planned surgical procedure. EuroSCORE is a recent model which has great accuracy in predicting mortality of patients undergoing cardiac surgery at European centres [3,4]. Of interest, EuroSCORE was recently shown to be useful in predicting mortality for a wide spectrum of risk in off-pump coronary surgery [5]. Although risk scores were generally devised to predict mortality, their use in describing other outcomes, such as length of stay in the ITU, is well recognised [6]. Our group has previously described the effect of training and the effect of return on bypass on the outcome of cardiac surgery [7,8]. We wished to take this further by examining the interaction between patient factors, as described by EuroSCORE, and several human factors related to the performance of surgery. Whether a prolonged procedure time is related to patient factors or medical team factors is impossible to determine in a retrospective analysis. There is no doubt however that long bypass time and long ischaemic time adversely affect physiology and the clinical outcome. On these grounds they should be incorporated in future refinements of the existing scores.

Our results show that knowledge of adverse intraoperative events enhances preoperative risk prediction. The derivation set and the test set were different in some respects (Table 1), but it is reassuring to see that the three models retain their predictive ability in spite of these differences. The odds ratio increase per EuroSCORE percentage increase is >1 for mortality (Table 2) (when it should be very close to 1), showing that there is a true interaction between pre- and intraoperative factors. This is confirmed by the standardised rates calculated in Table 3, all of which are >1. A higher number of outcomes than expected may have several explanations: (1) patient factors not covered by EuroSCORE; (2) intraoperative factors; and (3) postoperative variables. We shall discuss the last two groups of causes.

The fact that training does not increase the mortality risk is in keeping with previous findings [7]. The apparent protective effect of trainee operating on the pLOS-ITU

![Fig. 2. Case mix by surgical team. Expressed as median EuroSCORE with interquartile ranges. C = consultant, T = trainee. Cases operated by consultant with trainee as first assistant (C + T) had a higher EuroSCORE than the other two groups of patients \((P < 0.001, \text{ Kruskall–Wallis test})\).](image)
Risk factors, but rarely include intraoperative events or the risk scores are based on either preoperative or postoperative outcomes. The advantage of severity scores is that physiological measurements are used for coronary revascularisation, which may reduce mortality and morbidity. When hospital outcomes are used as an indicator of the quality of care, it is important to note the conclusions of Table 2. Risk-adjusted odds ratios (95% confidence interval) in the derivation sample.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Death (95% CI)</th>
<th>pLOS-ITU (95% CI)</th>
<th>Death or pLOS-ITU (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EuroSCORE*</td>
<td>1.22 (1.14, 1.30)</td>
<td>1.17 (1.11, 1.22)</td>
<td>1.19 (1.13, 1.24)</td>
</tr>
<tr>
<td>Trainee operating**</td>
<td>–</td>
<td>0.45 (0.29, 0.68)</td>
<td>0.41 (0.28, 0.60)</td>
</tr>
<tr>
<td>Prolonged CPB</td>
<td>1.96 (1.12, 3.41)</td>
<td>1.94 (1.31, 2.88)</td>
<td>1.98 (1.38, 2.85)</td>
</tr>
<tr>
<td>Prolonged XC</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Return CPB</td>
<td>9.16 (4.20, 19.96)</td>
<td>5.35 (2.22, 12.91)</td>
<td>7.58 (3.51, 16.39)</td>
</tr>
<tr>
<td>Mechanical support</td>
<td>3.62 (1.77, 7.40)</td>
<td>8.13 (4.45, 14.84)</td>
<td>8.40 (4.66, 15.16)</td>
</tr>
</tbody>
</table>

* Abbreviations are defined in the text. The odds ratios in the table resulted from multiple logistic regression. *Odds ratio increase per EuroSCORE band. **Regardless of assisted by consultant or not.

The difficulty in establishing that a risk scoring system is superior to others comes from two main causes. First of all, risk scores are based on either preoperative or postoperative risk factors, but rarely include intraoperative events or the complete time sequence. Parsonnet score outperformed three general severity systems in predicting death. A possible explanation is that the majority of cardiac surgical patients have a systemic inflammatory response syndrome (SIRS) after extracorporeal circulation. This may reduce the predictive value of acute physiology scores which, although devised for general ITU patients, have been applied with some success to cardiac surgery. Mortality models can be improved by adding observations that are closer to death, such as extreme bradycardia. Open sternum, mean arterial pressure of less than 60 mmHg, use of inotropes or pressor agents are a number of independent risk factors and this information would be readily available to most units. In summary therefore, although preoperative algorithms seem to offer an advantage over severity scores, mortality prediction is still lacking the desired power. A logical approach would be to combine pre-, intra- and postoperative variables. Such a model, both simple to apply and precise, is not available yet.

Secondly, in terms of outcomes measured, the factors affecting mortality may differ from those determining morbidity. When hospital outcomes are used as an indicator of the quality of care, it is important to note the conclusions of Table 3. Expected and observed outcomes in the validation group and standardised rates for the three risk models for death, pLOS-ITU and death or pLOS-ITU.

Table 3

<table>
<thead>
<tr>
<th>EuroSCORE band</th>
<th>Observed</th>
<th>Expected</th>
<th>SR (95% CI)</th>
<th>Observed</th>
<th>Expected</th>
<th>SR (95% CI)</th>
<th>Observed</th>
<th>Expected</th>
<th>SR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–2</td>
<td>0</td>
<td>4.1</td>
<td>0</td>
<td>21</td>
<td>15.3</td>
<td>1.37 (0.89, 2.10)</td>
<td>21</td>
<td>17.4</td>
<td>1.21 (0.79, 1.86)</td>
</tr>
<tr>
<td>3–5</td>
<td>12</td>
<td>14.3</td>
<td>0.84 (0.48, 1.48)</td>
<td>55</td>
<td>48.2</td>
<td>1.14 (0.88, 1.48)</td>
<td>64</td>
<td>56.9</td>
<td>1.12 (0.88, 1.43)</td>
</tr>
<tr>
<td>6+</td>
<td>52</td>
<td>35.7</td>
<td>1.46 (1.11, 1.92)</td>
<td>95</td>
<td>83.9</td>
<td>1.13 (0.92, 1.38)</td>
<td>126</td>
<td>105.2</td>
<td>1.20 (1.01, 1.43)</td>
</tr>
<tr>
<td>Total</td>
<td>64</td>
<td>54.1</td>
<td>1.18 (0.92, 1.51)</td>
<td>171</td>
<td>147.4</td>
<td>1.16 (1.00, 1.35)</td>
<td>211</td>
<td>179.5</td>
<td>1.18 (1.03, 1.35)</td>
</tr>
</tbody>
</table>

a The standardised rate is the ratio of observed/expected outcomes.

b Hosmer-Lemeshow goodness of fit $\chi^2$ = 2.307; degrees of freedom = 8; $P = 0.97$.

c Hosmer-Lemeshow goodness of fit $\chi^2$ = 12.3; degrees of freedom = 8; $P = 0.14$.

d Hosmer-Lemeshow goodness of fit $\chi^2$ = 12.92; degrees of freedom = 8; $P = 0.11$. 
Silber et al. [16]. They elegantly showed that rates of death and failure to rescue tend to be divergent from complication rates for individual hospitals. These outcomes possibly measure different aspects of healthcare delivery, or simply some units are better at rescuing. This brings us back to the role of human factors in the delivery of surgical treatments. A possible way in which patient and medical team risk factors overlap is shown in Fig. 4. Medical system factors are clearly more elusive, but the new practice environment is forcing us to submit them to closer scrutiny. We believe that an analysis such as this could be a first step in identifying major errors leading to negative outcomes. Instead of a blanket approach to audit, those cases could be targeted and analysed so that appropriate lessons are learned.

In conclusion, knowledge of negative intraoperative events improves risk prediction. Clearly, the performance of such a model can be further increased by adding more parameters to it. We chose not to include inotropic dosage for example because these data tend to be inaccurate unless they are prospectively and carefully recorded at the same moment in time for all patients. The study is also limited by the low frequency of some of the variables analysed and by the retrospective nature of model validation. However, we have embarked on prospective validation which will include selected variables related to the patients’ state in the first few hours in the ITU. With the information from the death/pLOS-ITU analysis we have also started an analysis of negative outcomes in low risk groups. This could potentially shed light on the issue of difference between predicted and actual mortality within a single institution [17].

Acknowledgements

Mr S.C. Stoica is supported by a research grant from the Garfield Weston Foundation. We are grateful to Sarah Powell and Frances Carey at Papworth Hospital for their support in collecting the clinical data. Cambridge Clinical Perfusion kindly provided information on mechanical support usage.

References

[2] Stoica SC, Sharples LD, Carey F, Roques F, Nashef SAM, Large SR. Can EuroSCORE and operative variables better predict outcome of...


Appendix A. Conference discussion

Dr B. Osswald (Heidelberg, Germany): The EuroSCORE which you evaluated is a preoperative score, which means you have to evaluate, out of the preoperative variables, the risk for mortality. What you included are now the operation-related factors, which most likely predict death after operation much better. What is the aim of your score?

Mr Stoica: To try and make this point I will give you an extreme example. Let’s consider a patient with a EuroSCORE of zero, a zero risk of mortality, who has a long bypass run, a long ischemic time, and also they have to return on bypass in theater and come out of the operating room with a balloon pump. It means that something has happened which changed the preoperative risk prediction, and this knowledge helps us to communicate with the patient, if they are awake, or with their family, and also to plan resources in the ITU; it affects the ITU turnover. Separately from that, we can target audit those sort of events because we need to know what is happening with low risk cases who have a bad run in theater, and perhaps we can learn more about these cases. These are the issues. So these are the two main objectives of our exercise.

As I showed in the presentation, risk modelling can be taken a step forward by introducing acute physiology scores or other variables at the time of patient arrival in the ITU. In this way prediction and discriminatory power can be further improved.

Dr S. Hagl (Heidelberg, Germany): I would like to come back to your statement before. You said that you are concentrating, if I understood well, on the patients entering the operation with a very low risk and where you have problems intraoperatively. That is possible, everybody knows. But I really doubt that you have enough numbers that you really can do a risk stratification in those patients. The normal patient, the low risk patient, having trouble intraoperatively, I think we are all happy about that, is the rarity.

Mr Stoica: This is a rare occurrence, of course, and I used this as an extreme example to make the point to the previous question. Our study encompassed the whole EuroSCORE range, and in fact we analyzed the calibration of the model across three EuroSCORE bands, low risk, medium risk and high risk, and it performed well for all of them.