Preoperative prediction of intensive care unit stay following cardiac surgery☆

Jeroen De Cocker a, Nouredin Messaoudi a, Bernard A. Stockman a, Leo L. Bossaert b, Inez E.R. Rodrigus a,∗

a Department of Cardiac Surgery, Antwerp University Hospital, Edegem, Belgium
b Department of Intensive Care, Antwerp University Hospital, Edegem, Belgium

Received 18 November 2009; received in revised form 13 April 2010; accepted 17 April 2010

Abstract

Objective: Following cardiac surgery, a great variety in intensive care unit (ICU) stay is observed, making it often difficult to adequately predict ICU stay preoperatively. Therefore, a study was conducted to investigate, which preoperative variables are independent risk factors for a prolonged ICU stay and whether a patient’s risk of experiencing an extended ICU stay can be estimated from these predictors. Methods: The records of 1566 consecutive adult patients who underwent cardiac surgery at our institution were analysed retrospectively over a 2-year period. Procedures included in the analyses were coronary artery bypass grafting, valve replacement or repair, ascending and aortic arch surgery, ventricular rupture and aneurysm repair, septal myectomy and cardiac tumour surgery. For this patient group, ICU stay was registered and 57 preoperative variables were collected for analysis. Descriptives and log-rank tests were calculated and Kaplan–Meier curves drawn for all variables. Significant predictors in the univariate analyses were included in a Cox proportional hazards model. The definitive model was validated on an independent sample of 395 consecutive adult patients who underwent cardiac surgery at our institution over an additional 6-month period. In this patient group, the accuracy and discriminative abilities of the model were evaluated. Results: Twelve independent preoperative predictors of prolonged ICU stay were identified: age at surgery > 75 years, female gender, dyspnoea status > New York Heart Association class II (NYHA II), unstable symptoms, impaired kidney function (estimated glomerular filtration rate (eGFR) < 60 ml min⁻¹), extracardiac arterial disease, presence of arrhythmias, mitral insufficiency > colour flow mapping (CFM) grade II, inotropic support, intra-aortic balloon pumping (IABP), non-elective procedures and aortic surgery. The individual effect of every predictor on ICU stay was quantified and inserted into a mathematical algorithm (called the Morbidity Defining Cardiosurgical (MDC) index), making it possible to calculate a patient’s risk of having an extended ICU stay. The model showed very good calibration and very good to excellent discriminative ability in predicting ICU stay >2, >5 and >7 days (C-statistic of 0.78; 0.82 and 0.85, respectively). Conclusions: Twelve independent preoperative risk factors for a prolonged ICU stay following cardiac surgery were identified and constructed into a proportional hazards model. Using this risk model, one can predict whether a patient will have a prolonged ICU stay or not.

Keywords: Cardiac surgery; Risk stratification; Intensive care unit stay; Predictive score

1. Introduction

Over the last few decades, statistical techniques for risk modelling have seen a rise in implementation in medical research in an attempt to quantify some of the remaining intangibles of modern-day medicine. One of the most researched of these topics in the field of cardiac surgery is preoperative mortality risk, which has led to the development of many risk-scoring systems including the European system for cardiac operative risk evaluation (EuroSCORE) [1,2]. However, due to improved surgical techniques and postoperative care, mortality rates have steadily declined over the years despite increased operative risk [3,4], whereas postoperative morbidity and use of intensive care unit (ICU) resources have not [5,6]. To be able to predict the length of ICU stay after cardiac surgery, several risk-scoring algorithms have been developed over the past few years [7–11]. Unfortunately, most of these algorithms lack either the power or a good end-point definition to accurately predict ICU length of stay.

In this article, we will assess the construction of a predictive index that should help physicians stratify between the need of an extended ICU stay or not, its advantages and its limitations.
2. Materials and methods

2.1. Patient population

In this study, clinical data were gathered consecutively for all adult patients who underwent cardiac surgery at the Antwerp University Hospital for a period of 2 years from 1 January 2005 until 31 December 2006. For this patient group (n = 1566), 57 preoperative variables were assessed and collected for analysis. In the data collection and encoding process, all variables were registered conforming to the United Kingdom and Ireland Society of Cardiothoracic Surgeons Registry definitions (http://www.scts.org/documents/PDF/5thBlueBook2003.pdf, pages 335—340) published in 2003. Patients undergoing coronary artery bypass grafting (CABG), valve replacement or repair, ascending or aortic arch surgery, cardiac tumour surgery, ventricular rupture or aneurysm repair or septal myectomy were included in the analysis. Length of stay (LOS) in the ICU was registered in hours by the ICU staff and recalculated in days for the purpose of the study, using standard round-off rules. Postoperative and total hospital stays were derived from the hospital’s medical records. Mortality was defined as death prior to discharge from the hospital. Quality control of the database was obtained by double-checking each data entry and by systematically searching for errors after completion of the data collection.

2.2. Predictors of ICU stay and risk modelling

Kaplan—Meier curves and log-rank tests were drawn and calculated respectively for each potential predictive factor. To get an as much as accurate view of LOS, patients who died before they were discharged from the ICU were censored in the process. Variables with adequate influence on LOS in the univariate log-rank test (defined as $P < 0.10$), as well as every variable with major clinical suspicion to have an influence on LOS (disregarding the log-rank statistic) were included in a Cox proportional hazards model. Using this statistical modelling technique, the independent effect of each and every of the included potential predictors on the outcome variable was determined. Several model-fitting techniques such as forward and backward stepwise procedures were used to determine predictors having an influence significant enough to remain in the final model. The proportional hazards assumption was met for each of the remaining predictors. No significant interaction terms could be added to the model. Furthermore, no relevant correlations between predictors in the model could be found. Finally, the presence of under- or overfitting was assessed by means of the ratio of the number of outcome events relative to the number of included potential predictors.

2.3. Validation

To externally validate the obtained mathematical risk model, the same set of variables as used in the original study population was gathered prospectively from 395 consecutive adult patients who underwent cardiac surgery at our centre from 1 January 2008 until 30 June 2008. The same statistical techniques as used in the predictor identification process were applied, making a comparison between patient characteristics and regression coefficients between the two patient populations possible. Furthermore, the calibration of the model was assessed through comparison of the mean LOSs for every subgroup of the predictive index. Finally, receiver operating characteristic (ROC) curves and their corresponding areas under the curve (AUC) were drawn and calculated, respectively, to evaluate the discriminative ability of the model to predict an ICU stay beyond 2, 5 and 7 days. Cross-validation with both the additive and logistic EuroSCORE was subsequently performed using the Hanley—McNeil test. All statistical analyses were performed using the Statistical Package for the Social Sciences (SPSS) 16.0 software (SPSS Inc., Chicago, United States of America).

3. Results

3.1. Patient characteristics

Descriptives were obtained for all variables and are summarised in Tables 1 and 2. Mean age at surgery was 67 years, with a male predominance of 72.6%. A total of 63% of all patients underwent isolated CABG, whereas 15% had isolated valve surgery and another 15% underwent a combination of both types of surgery. Of all the 1566 operations performed, 1435 (91.6%) were elective, 1494 (95.4%) were performed on cardiopulmonary bypass (CPB) and 91 (5.8%) were redo cardiac operations. Seventy patients (4.5%) died before they were discharged from the hospital; all of them passed away while admitted to the ICU. Patients had a mean ICU stay of 5 days, whereas 57% of patients had an ICU stay of 2 days or less. Furthermore, patients had mean postoperative and total hospital stays of 13 and 16 days, respectively. The distribution of ICU and postoperative length of stay for all patients alive at the time of ICU discharge is shown in Fig. 1.

3.2. Predictors of extended ICU stay

Of the 57 investigated preoperative variables, univariate survival analysis identified 40 potential predictors of a prolonged ICU stay. Using a multivariable Cox proportional hazards model for regression and model fitting, these potential predictors were researched for their independent effect on ICU stay. Eventually, 12 such predictive variables were identified: increased age at surgery (>75 years), female gender, dyspnoea status New York Heart Association (NYHA) class III/IV, unstable symptoms, impaired kidney function (eGFR MDRD < 60 ml min$^{-1}$), extracardiac arterial disease, presence of arrhythmias, mitral insufficiency CFM grade III/IV, inotropic support, intra-aortic balloon pumping (IABP), non-elective procedures and aortic surgery. The individual effect of every predictor on ICU stay was quantified (Table 3). Three predictors that had $P$-values $< 0.10$, but $> 0.05$ were excluded because they had little influence on the outcome. Age at surgery and kidney function was categorised on the basis of a combination of observed differences in means between different ICU stay subgroups and clinical experience. Both the continuous and dichotomous forms of these variables were subsequently entered in the model, with the
dichotomous forms yielding the better fit in each case. Due to a high number of outcome events (n = 1496) compared with a relative low number of researched independent variables (n = 40), problems of over- or underfitting were not encountered.

### 3.3. Risk model

Using these predictors, a mathematical risk model was constructed to assess the risk of an individual patient to have an extended ICU stay. To obtain a risk prediction independent
of time, the baseline hazard $h_0(t)$ was omitted from the applied Cox model, leaving the following equation:

$$h_i = e^{b_1x_1 + \ldots + b_{12}x_{12}},$$

where $h_i$ denotes the resultant hazard for the given values $x$ of the 12 predictors with coefficients $b$. The obtained hazard signifies the probability of an individual patient to experience a 'non-extended' stay in the ICU. For the sake of clinical usability, this hazard can be transformed into a more intuitive risk function, making use of the formula $S_i = e^{-h_i}$. The obtained risk index ranges from 36.79 (e^-1, no known risk factors present) to 98.88 (all known risk factors present) and is termed the 'MDC-index' (Morbidity Defining Cardiosurgical index). Meticulous analysis of the MDC-index in function of ICU stay shows a trimodal distribution of risk (Table 4), which allows us to categorise the ICU score in three distinct risk groups: those at low (MDC-index 36.79—64.99, median ICU stay 2 days), moderate (MDC-index 65.00—79.99, median ICU stay 4 days) and high (MDC-index 80.00—98.99, median ICU stay 8 days) risk of an extended ICU stay.

### 3.4. Validation

Comparisons between the descriptives and model coefficients derived from the validation population and the originally studied population are displayed in Table 5. Although small variations occur, no dramatic differences are noticed between the 12 studied predictors. In general, it can be stated that predictors with large hazard ratio’s (HR) generate greater differences compared with their counterparts in the validation dataset than other predictors (which is due to the relatively rare occurrence of these important risk factors, leading to an under- or overestimation of their effect.

---

**Table 2**

Descriptives of continuous variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Range</th>
<th>Mean ($\pm$SD)</th>
<th>Laboratory parameters (cont’d)</th>
<th>Range</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographic data</td>
<td>Laboratorv parameters</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at surgery (years)</td>
<td>20—90</td>
<td>67 ($\pm$10)</td>
<td>Serum creatinine (mg dl$^{-1}$)</td>
<td>0.50—9.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>138.00—196.00</td>
<td>169.72 ($\pm$8.92)</td>
<td>eGFR MDRD (ml min$^{-1}$)</td>
<td>5.05—199.97</td>
<td>73.34</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>36.00—148.00</td>
<td>77.57 ($\pm$14.09)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body Surface Area (BSA)</td>
<td>1.22—2.77</td>
<td>1.91 ($\pm$0.20)</td>
<td>Preoperative stay (days)</td>
<td>0—109</td>
<td>1</td>
</tr>
<tr>
<td>Body Mass Index (BMI)</td>
<td>15.79—48.22</td>
<td>26.89 ($\pm$4.28)</td>
<td>ICU length of stay (days)</td>
<td>0—147</td>
<td>2</td>
</tr>
<tr>
<td>Laboratory parameters</td>
<td></td>
<td></td>
<td>Postoperative stay (days)</td>
<td>0—262</td>
<td>9</td>
</tr>
<tr>
<td>Haemoglobin (g dl$^{-1}$)</td>
<td>8.0—17.2</td>
<td>13.3 ($\pm$1.7)</td>
<td>Total hospital stay (days)</td>
<td>0—263</td>
<td>12</td>
</tr>
<tr>
<td>Haematocrit (%)</td>
<td>23.4—50.6</td>
<td>38.8 ($\pm$4.8)</td>
<td>Additive EuroSCORE</td>
<td>0—20</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Logistic EuroSCORE (%)</td>
<td>0.88—87.07</td>
<td>3.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>MDC-index</td>
<td>36.79—98.88</td>
<td>52.79</td>
</tr>
</tbody>
</table>

SD: standard deviation; eGFR: estimated glomerular filtration rate; MDRD: modified diet in renal disease; EuroSCORE: European system for cardiac operative risk evaluation; and ICU: intensive care unit.

**Table 3**

Independent predictors of ICU stay.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Hazard ratio$^a$ (with 95% CI)</th>
<th>Increased risk$^b$ (%)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increased age at surgery (&gt;75 years)</td>
<td>1.175 (1.021—1.351)</td>
<td>17.5</td>
<td>0.024</td>
</tr>
<tr>
<td>Female gender</td>
<td>1.087 (1.035—1.225)</td>
<td>8.7</td>
<td>0.008</td>
</tr>
<tr>
<td>Dyspnoea status (NYHA III/IV)</td>
<td>1.166 (1.045—1.301)</td>
<td>16.6</td>
<td>0.006</td>
</tr>
<tr>
<td>Unstable symptoms</td>
<td>1.142 (1.025—1.272)</td>
<td>14.2</td>
<td>0.016</td>
</tr>
<tr>
<td>Impaired kidney function (CKD stage III)</td>
<td>1.243 (1.102—1.401)</td>
<td>25.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Impaired kidney function (CKD stage IV/V)</td>
<td>1.870 (1.312—2.663)</td>
<td>87.0</td>
<td>0.001</td>
</tr>
<tr>
<td>Extracardiac vascular disease</td>
<td>1.193 (1.039—1.370)</td>
<td>19.3</td>
<td>0.012</td>
</tr>
<tr>
<td>Preoperative arrhythmias</td>
<td>1.321 (1.132—1.542)</td>
<td>32.1</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Mitral insufficiency (CFM grade III/IV)</td>
<td>1.333 (1.156—1.536)</td>
<td>33.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Inotropic support</td>
<td>3.033 (1.832—5.021)</td>
<td>203.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Intra-aortic balloon pumping (IABP)</td>
<td>1.849 (1.305—2.619)</td>
<td>84.9</td>
<td>0.001</td>
</tr>
<tr>
<td>Non-elective procedure</td>
<td>1.219 (1.026—1.510)</td>
<td>21.9</td>
<td>0.031</td>
</tr>
<tr>
<td>Aortic surgery</td>
<td>1.937 (1.438—2.608)</td>
<td>93.7</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

ICU: intensive care unit; CI: confidence interval; NYHA: New York Heart Association classification; CKD: chronic kidney disease; and CFM: colour flow mapping.

$^a$ The hazard ratio (or relative risk) is the quotient of the hazard when the predictor is present, through the hazard when the predictor is absent.

$^b$ Risk increased compared to baseline ICU stay.
when evaluated in a relatively small population such as the validation data set). Furthermore, the ability of the MDC-index to discriminate an ICU stay of more than 2, 5 and 7 days was assessed using ROC curves to calculate the C-statistic (also known as AUC) at each cut-off point. C-statistics of 0.78 (0.72–0.83; P < 0.001); 0.82 (0.74–0.90; P < 0.001) and 0.85 (0.77–0.94; P < 0.001) were obtained for ICU stays of >2, >5 and >7 days, respectively (Fig. 2(A)). Using the Hanley–McNeil test, the additive and logistic EuroSCOREs (Fig. 2(B) and (C)) generated significantly lower C-statistics of 0.64, 0.74 and 0.79; logistic EuroSCORE: 0.64, 0.75 and 0.80; P < 0.001) when compared with the MDC-index. Therefore, we can conclude that the MDC-index possesses very good (C-statistic > 0.75) to excellent (C-statistic > 0.80) discriminative abilities.

### 4. Discussion

As pointed out in a recently published literature review [12], preoperative prediction of ICU stay is an important issue in contemporary cardiac surgery. Whereas, in some cases, clinical judgement may help experienced surgeons to stratify a patient’s risk of having an extended stay, many other cases remain where clinical judgement alone tends to be inaccurate, especially in high-risk patients [13]. For these patients, mathematical risk-scoring algorithms allow stratification between patients with and without extended ICU stays. In the past, many such risk-scoring algorithms have been developed [7–11]. In addition, risk models originally constructed to predict other outcomes such as mortality have been validated to predict prolonged ICU stay [14–16]. However, because of a lack of uniform predictor and endpoint definitions, a great variation in identified predictors exists among these risk scores. Moreover, these predictive indices all neglect the presence of patients who died before they were discharged from the ICU, while most of them would probably have had extended ICU stays if they stayed alive. Finally, by making use of dichotomous end points, these risk models are only able to predict whether a patient will remain admitted to the ICU beyond a predefined time, for example, 3 days.

In this study, we addressed these issues by making use of a Cox proportional hazards model [17,18]. A statistical method was used mainly in medical research for survival analysis, it possesses the unique properties necessary to accurately predict an extended ICU stay following cardiac surgery. First, the censoring function makes it possible to take patients into account when they were discharged from the ICU. Therefore, we can conclude that the MDC-index possesses very good (C-statistic > 0.75) to excellent (C-statistic > 0.80) discriminative abilities.
account who died while they were admitted to the ICU. Furthermore, instead of categorising the outcome variable, ICU stay is handled as a continuous variable. This way, predictors that repeatedly predispose to longer ICU stays than other risk factors have a larger effect on the predictive index (by means of a larger coefficient) when compared to conventional logistic regression methods. When all the coefficients of the risk factors that are present are added to each other, one can accurately predict an individual patient’s risk to experience an extended ICU stay.

When compared with previously published data, most of the predictors identified in this study have been reported before [12]: increased age at surgery (ten studies), female gender (three studies), dyspnoea status > NYHA II (one study), impaired kidney function (four studies), extracardiac arterial disease (two studies), presence of arrhythmias (three studies), mitral insufficiency > CFM grade II (two studies), inotropic support (one study) and non-elective procedures (six studies). The only two newly identified predictors were preoperative IABP and aortic surgery. However, both of them seem to be important omissions in previous studies because of the magnitude of their influence on ICU stay (risk increase of 84.9% and 93.7%, respectively). On the contrary, a few relevant previously reported risk factors were not included in our risk model: left ventricular ejection fraction (six studies) correlated significantly with preoperative inotropic support; recent myocardial infarction (four studies) was of little influence and borderline significance ($P = 0.038$); chronic pulmonary disease (four studies; $P = 0.146$) and on-pump surgery (four studies; $P = 0.164$) were non-significant. Anyhow, the inclusion of additional predictors with no more than a minor effect only has a tendency to lead to overfitting and subsequent model inaccuracy [19]. Furthermore, it is inevitable that some intangibles will remain despite the advances cardiac surgery is bound to make over the next few decades.

However sparse, validation techniques, specifically for Cox proportional hazard models, have been described before [19]. In validating any type of mathematical model, three distinct entities need to be verified to assess model accuracy: (1) model fit, (2) model calibration and (3) model discrimination. Model fit is the only one of the three that can be checked internally (on the same data set from which the model was derived) and relates to the amount and combination of predictors needed to obtain maximal model accuracy, meaning that adding or subtracting any predictor will only worsen model accuracy. When sufficient outcome events are present, the right model fit can be obtained by making use of several consecutive semi-automated stepwise procedures to evaluate the significance of each possible predictor to the model. In our study, 37 outcome events were observed for each potential predictor entered in the model, whereas only 10 are required to avoid problems of under- and overfitting [20]. Model calibration and discrimination need to be assessed on an external dataset to obtain a non-biased evaluation of the risk model. Model calibration (or reliability) relates to the extent to which the risk model can accurately predict the observed ICU stay, whereas model discrimination (or resolution) refers to the ability of the model to separate patients with different responses (e.g., staying beyond 3 days or not). Our model showed excellent calibration and discrimination, as demonstrated in Table 3 and Fig. 2(A). Finally, external cross-validation with other predictive indices allows comparison and ranking of the most performant predictive index. Because it was recently demonstrated that the EuroSCORE has a good ability to predict an ICU stay longer than 2, 5 or 7 days [14,15] and because it is one of the

![Fig. 2. Receiver operating characteristic (ROC) curve predicting ICU stay >2, >5 and >7 days in the validation dataset for (A) the MDC-index; (B) the additive EuroSCORE and (C) the logistic EuroSCORE.](image-url)
most widely used predictive indices in cardiac surgery, both its additive and logistic forms were calculated systematically for every patient in the validation population and compared to our predictive index, making use of ROC curves. As shown in Fig. 2, the MDC score is significantly more performant than either EuroSCORE form, hence consolidating its status as a highly accurate and well-validated predictive tool to predict extended ICU stays following cardiac surgery.

The shortcomings of the study should be addressed as well. First, because of the retrospective nature of the model construction part of this single-centre study, mistakes in data registration might have occurred. Second, variables such as dyspnoea status and unstable symptoms may seem prone to personal interpretation and can therefore lead to predictive inaccuracy. However, because the variables in question only exercise a relatively small effect on ICU stay, this imprecision — if occurred — will only have affected ICU stay marginally. Nevertheless, when used in clinical practice, it will be of utmost importance to clearly define any of the identified predictors, thus rendering any possible inaccuracy into futility. Further, because of the limitations inherent to Cox modelling (the baseline hazard is not accurately estimated), we were unable to add a time-component into our risk model and consequently, predict ICU stay as a number of days. Although we were able to validate a model that included the Cox-derived baseline hazard, we feel that parametric Weibull modelling is mandatory to estimate ICU stay in a number of days. In addition, patients who died while admitted to the ICU were censored in the Cox regression. Therefore, several predictors in our model are potentially underestimated since several predictors for mortality are similar to those of an extended ICU stay (e.g., several risk factors are present in both the EuroSCORE and the MDC-index). Although not implemented in this article, multi-state models may offer a more accurate estimation of the predictor coefficients [21]. Furthermore, since we only considered preoperative risk factors, our model loses its predictive qualities from the moment the first incision is made. As Wong and colleagues have demonstrated, the effects of preoperative risk factors almost disappear when pre- and postoperative factors (such as pump time, sternal wound dehiscence and prolonged ventilation) are taken into account [10]. However, most of the reasons why one would want to know a patient’s probable ICU stay (such as contemplating the advantages and disadvantages of cardiac surgery, timing of surgery, patient communication and allocation of resources) tend to relate to the preoperative setting. Nevertheless, we can only encourage the construction of a dynamic predictive index in which all potential risk factors are taken into account. This way, intensivists can continually update patients and their families about their expected stay in the ICU. Finally, risk modelling is heavily dependent on time and place [22]. This is the main reason why some risk scores cannot be validated in certain populations, no matter how large the sample size [23,24]. However, both time and place are problems that can be accounted for through longitudinal risk factor updating [25] (e.g., every 5–10 years) and model construction in various corners of the world. In a unified easy-access risk model (e.g., on the World Wide Web), the relevant risk factors and coefficients could then be selected automatically by entering the area or country where the centre in question is located.

In conclusion, using a multivariable proportional hazards analysis, we were able to construct and validate a preoperative predictive index for ICU stay following cardiac surgery consisting of 12 different risk factors. This predictive index allows an accurate estimation of the probability of an extended ICU stay, which is a clinically useful parameter. However, given the nature of this single-centre study, its results can only be applied regionally. Therefore, we strongly urge other European centres to join us in the creation of a user-friendly multi-institutional European predictive index for ICU stay, following the methodology outlined in this article. This way, prediction of ICU stay can become an integral part of standard patient communication and decision making in cardiac surgery.

References


