View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Apollo

Prediction of patient length of stay on the Intensive Care Unit following cardiac surgery: a logistic regression analysis based on the cardiac operative mortality risk calculator, EuroSCORE

Katherine Meadows,[1] Richard Gibbens,[2] Caroline Gerrard,[3] Alain Vuylsteke[3]

[1]School of Clinical Medicine, University of Cambridge, Cambridge, UK

[2]Computer Laboratory, University of Cambridge, Cambridge, UK

[3]Papworth Hospital NHS Foundation Trust, Cambridge, UK

Corresponding author:

Dr Katherine Meadows

School of Clinical Medicine, University of Cambridge

Addenbrooke's Hospital, Hills Road | Cambridge | CB2 0SP | UK |

Telephone: +44 (0)1223 336700

Email: k.e.meadows@cantab.net

Funding Statement

This work was supported by Papworth Hospital NHS Foundation Trust

Conflict of Interest Statement

Conflict of interest: none declared.

*Structured Abstract/Summary

Abstract

Objectives: The aim of this study was to develop a statistical model based on patient parameters in order to predict the length of stay (LOS) on the Intensive Care Unit (ICU) following cardiac surgery in a single centre.

Design: Data was collected from patients admitted to the ICU following cardiac surgery over a 10-year period (2006-2016). Both the additive and logistic EuroSCORE were calculated and logistic regression analysis was carried out to formulate a model relating the predicted LOS to the EuroSCORE. This model was used to stratify patients into short stay (less than 48 hours) or long stay (more than 48 hours).

Setting: Intensive care unit (ICU) at Papworth Hospital, Cambridgeshire

Participants: 18,377 consecutive patients who had been in ICU following cardiac surgery (coronary graft bypass surgery, valve surgery or a combination of both)

Interventions: This was an observational study

Measurements and main results: We have shown that both the additive and logistic EuroSCORE can be used to stratify cardiac surgical patients in various predicted length of stay in ICU. Further adjustments can be made to increase the number of patients correctly identified as either short stay or long stay. Comparison of the model predictions to the data demonstrated a high overall accuracy of 79.77% and ROC curve analysis showed the AUC to be 0.7296.

Conclusions: This analysis of an extensive data set shows that patient LOS in ICU after cardiac surgery in a single centre can be predicted accurately using the simple cardiac operative risk scoring tool EuroSCORE. Using such predictions has the potential to improve ICU resource management.

Keywords: cardiac surgery, EuroSCORE, intensive care unit, length of stay

Introduction

Bed availability in intensive care units (ICUs) is a critical factor in determining patient flow in hospitals providing cardiac surgery. An accurate tool to predict patient length of stay (LOS) would facilitate efficient patient scheduling and maximise available capacity. Patients requiring a prolonged stay make up a significant proportion of total patient days on the ICU (1) and there is a much greater financial cost for their care in addition to the high-cost cardiac surgery. (2)

A variety of approaches have been used in order to predict LOS on the ICU for patients. (3-6) Scoring tools that provide an indication of the severity and type of illness, such as APACHE (Acute Physiology and Chronic Health Evaluation) have been used. (7-9) These predictions are not always accurate for individual patients, (5) are based on physiological factors obtained at the point of admission to ICU and often require the use of complex statistics or modelling. (10)

We sought to develop a simple model that would allow straightforward pre-operative stratification of patients according to a predicted ICU LOS. We took the view that such a model should be based on variables routinely available to the clinical team before the operation, such as the ones used to calculate the EuroSCORE, a tool for calculating cardiac operative risk. (11,12)

Methods

Patient population and data collection.

Following institutional approval, data on consecutive patients who underwent cardiac surgery from 2006 to 2016 were obtained from the perioperative surgical and intensive care electronic record at a large tertiary referral and teaching hospital dedicated to adult cardiac surgery.

Papworth hospital is a large specialist adult cardiothoracic hospital located in the United Kingdom and treating around 24,400 inpatients and day cases and 73,600 outpatient per year.

It is a tertiary centre accepting referrals from the whole country, and operates within the English national Health Service.

The study was conducted in accordance with our local review board and the need for informed consent was waived as no identifiable data were used.

Patient management

All patients underwent cardiac surgery at a single institution with a standard, but not rigid, perioperative protocol. All patients were admitted postoperatively in a level 3 unit (ICU) and progressed rapidly to level 1. The unit is led by consultant intensivists and all patients are reviewed at least once a day by a multidisciplinary team that decides if they are fit for discharge to the ward.

Data

In addition to age and length of stay on the unit, the following parameters were extracted: gender, serum creatinine, left ventricular ejection fraction, systolic pulmonary pressure, the presence of extracardiac arteriopathy, pulmonary disease, neurological dysfunction, previous cardiac surgery, recent myocardial infarct, active endocarditis, unstable angina or critical perioperative state, and the type of surgery (emergency operation, ventricular septal rupture, thoracic aortic surgery, other than isolated coronary surgery).

Statistical methods

All data analysis was performed using the statistical software package R. (13) The logistic model was fitted by the glm() function, a standard operation in R. The primary objective of this model was to predict the LOS in ICU of individual patients. The data was classified into two distinct sets based on LOS, with short stay arbitrarily defined as less than 48 hours and long stay more than 48 hours. The EuroSCORE, a cardiac mortality risk scoring tool was used as the basis of the logistic regression model. (11) Initial analysis was carried out with the

EuroSCORE (additive), which gives an integer score for each patient between 0 and 24. The EuroSCORE (logistic) was also used, which gives a probability of early mortality following cardiac surgery, expressed as a percentage. The logistic EuroSCORE provides a more accurate estimate of mortality in higher risk patients than the additive EuroSCORE and hence both scoring systems were analysed to compare the overall accuracy. (11)

Fitting the logistic regression model to the data yielded estimates for the parameters used, along with the standard errors and *p*-values, indicating which were statistically significant. Using these parameter estimates, predicted outcomes were generated and a contingency table of observed and predicted outcomes was produced. This was then analysed to determine the true and false positive rates for the model and the overall accuracy.

Testing the model accuracy in this way has the potential to overestimate accuracy since the data the model is being compared against is the same data used to generate the model. In order to overcome this issue *k*-fold cross-validation was used.(14) This method divides the data into *k* groups (in this case 10) and assesses the model accuracy on each of the groups. An average is then taken as the mean test accuracy. This value was not found to show a statistically significant difference from the initial accuracy result, validating this method of analysis.

Results

A summary of patient characteristics is shown in table 1. The maximum LOS on the ICU was 123 days, with a mean stay of 2.4 days and a median of 1.0 days. Of the patients admitted to the unit, 13,637 were admitted as an elective case, 3,822 were urgent and 918 were emergency admissions.

Initially, the LOS data was plotted in relation to the additive EuroSCORE to evaluate its distribution and other characteristics. There was a wide range of LOS with varying EuroSCORE, with no clear initial structure. To elucidate an underlying distribution, the data was separated into three categories of length of stay: short (less than 48 hours), intermediate (48 to 120 hours) and prolonged (over 120 hours). A cutoff of 48 hours for short stay patients was used as this is a commonly used value in previous LOS prediction models. (15,16) The distribution of data when separated into these categories showed a clear difference in the short stay category, but a large degree of overlap in the intermediate and prolonged stay categories (Figure 1a). Box-plots revealed similar medians and overlapping interquartile ranges for the intermediate and prolonged stay categories (Figure 1b). Due to this degree of overlap, further analysis was carried out with data separated into just two categories: short stay (less than 48 hours) and long stay (over 48 hours), which demonstrated a clear difference between categories (Figure 1c) and is confirmed by the box-plot analysis which shows only slightly overlapping distributions (Figure 1d). The above analysis, carried out with the additive EuroSCORE, was repeated with the logistic EuroSCORE and revealed a similar pattern of distribution.

In order to develop a predictive scoring tool based on the EuroSCORE, a logistic regression model was used to model the probability of a patient requiring a long stay conditional on the associated EuroSCORE (additive).

$$P(\text{stay} = \text{long} \mid \text{EuroSCORE} = X) = \frac{e^{\beta_0+\beta_1 X}}{1 + e^{\beta_0+\beta_1 X}}$$

Fitting this model to the patient data yielded the following parameter estimates: $\beta_0 = -0.0323 \pm 0.0475$ and $\beta_1 = 0.2666 \pm 0.0060$. To use the logistic regression model to predict patient length of stay a classifier, η was chosen with $0 < \eta < 1$, so that a patient is predicted to have a long stay if

$$P(\text{stay} = \text{long} \mid \text{EuroSCORE} = X) \geq \eta$$

for the EuroSCORE value X calculated for that patient. Initially η was chosen to be 0.5, however, the importance of this choice and implications on the predictive accuracy of the model is discussed later. As an example, with a EuroSCORE value of 15, the conditional probability of a long stay is calculated to be 0.724. Assuming that $\eta = 0.5$, then any patients with a EuroSCORE of 15 would be assigned to the predicted long stay category. Repeating this calculation with the EuroSCORE values from the patient data set and comparing the predicted LOS category with the actual outcome is shown in table 2. Formulating a model which assigns patients to one of two predicted categories enables the stratification of patients into either a short stay or long stay pathway, with the aim of optimising patient flow.

Evaluation of this table gives the accuracy of the model to be 79.77% with a true positive rate of 0.1946 and a false positive rate of 0.0262 for the identification of long stay patients. This analysis suggests that whilst the overall accuracy of the model is good, it is particularly effective at identifying short stay patients, but less reliable for long stay patients. This simple analysis using the EuroSCORE as a single predictive variable allows a cut-off value to be set, above which a patient is classified as long stay. Plotting the EuroSCORE (additive) against the threshold probability, shows that for a classifier value of $\eta = 0.5$, the classification threshold value is 11.35 (Figure 2a), i.e. any patients with a EuroSCORE above this value will be predicted to be long stay. Plotting this value on the EuroSCORE distribution curves discussed earlier illustrates that whilst the model has a good specificity for the identification of long stay patients, there are still a significant number of patients in this category that are incorrectly assigned (Figure 2b).

To assess the effect that the classifier threshold value of η has on the true and false positivity rate, a receiver operating characteristic (ROC) curve was plotted. For an ideal predictive model with a true positive rate of 100% and a false positive rate of 0%, the ROC curve would have an area under the curve (AUC) of 1. Conversely, a random selector which arbitrarily assigns a patient to either the short or long stay category would have an AUC of 0.5, with the true positive rate varying linearly with the choice of η. Figure 3 shows that for the logistic regression model used herein, the AUC is 0.7267 and the black dot indicates the true and false positive rates for the chosen classifier value of $\eta = 0.5$. The curve clearly demonstrates the impact that the choice of classifier value has on the rates of true and false positives and the implicit trade-off between the two.

To explore this further, additional analysis was carried out with different values of η. As would be expected, decreasing η allows correct identification of more long stay patients, i.e. the true positive rate increased. However, there is also a corresponding increase in the number of short stay patients incorrectly predicted to be long stay, i.e. the false positive rate also increases. Table 3 shows a summary of statistical measures for different η values, which

demonstrates that the increase in true positive rate with decreasing η comes at the expense of a decrease in overall accuracy. For the purpose of resource management, it may be preferable to accept a decrease in accuracy in order to correctly identify more long stay patients. Failure to identify a significant proportion of long stay patients could result in an unexpected demand for bed space, nursing requirements and higher unanticipated costs.

The detailed analysis described above was repeated with the EuroSCORE (logistic) data which yielded similar results. Evaluation of the resulting contingency table gave an overall accuracy of 79.73%, almost identical to that of the EuroSCORE (additive) of 79.77%. The false positive rate was 0.0278, true positive rate 0.1982 and cut-off value 26.09% with a classifier value of $\eta = 0.5$. ROC curve analysis showed the AUC to be 0.7296.

Since analysis using the EuroSCORE (additive) and EuroSCORE (logistic) yielded extremely similar results, further investigation was carried out to determine whether the model accuracy could be improved. Both forms of the EuroSCORE are calculated from a weighted model based on 18 underlying parameters measured peri-operatively. The logistic regression model was adapted to treat each parameter as an individual variable with different weightings using the following conditional probability

$$P(\text{stay} = \text{long} \mid X_1, X_1, \ldots, X_{18}) = \frac{e^{\beta_0+\beta_1 X_1+\ldots+\beta_{18}X_{18}}}{1 + e^{\beta_0+\beta_1 X_1+\ldots+\beta_{18}X_{18}}}$$

where $X_1, X_2, \ldots, X_{18}$ are the values of the 18 parameters.

Fitting the patient data to the logistic regression model provided estimates for $\beta_0, \beta_1, \ldots, \beta_{18}$ and the p value of each variable in the model. Out of the 18 parameters, three were found not to be statistically significant: gender, diagnosis of active endocarditis and presence of ventricular septal rupture. We hypothesise that due to the extremely small number of patients with active endocarditis or ventricular septal rupture (81 and 58 respectively), there are simply not enough data points for these variables to influence the model.

As before, the contingency table for the model was calculated and analysed with $\eta = 0.5$, and revealed an overall accuracy of 80.54%. This is a modest increase compared to the initial analysis with EuroSCORE as the only variable. However, with each parameter modelled individually, the true positive rate increased to 0.2504, a 29% rise compared to the single-variable model. Additionally, the false positive rate remained extremely low at 0.0326. Repeat calculations with the three non-significant variables removed resulted in very similar results with accuracy 80.56%, true positive rate 0.2506 and false positive rate 0.0323.

This simple logistic regression model does not currently take into account patient deaths which has the potential to influence LOS. In this data set there were 376 deaths out of a total of 18,377 admissions (2.05%). The mean and median EuroSCORE (additive) values for these patients were both 10, with mean and median LOS of 249 and 119 hours respectively. This data demonstrates that the patients that died were more likely to have higher EuroSCORE values and require a longer LOS and including these patients in the analysis is unlikely to affect the results.

Discussion

We have used the simple cardiac risk scoring tool, EuroSCORE to predict LOS on the ICU. Statistical analysis using a logistic regression model has demonstrated no significant difference in results using either the additive or logistic EuroSCORE. Using this model, it is possible to determine a threshold EuroSCORE above which, a patient would be predicted to have a long stay in ICU. Analysis has demonstrated that altering this threshold affects the overall predictive accuracy, in addition to the positive and negative predictive values of the model, as would be expected. This illustrates the balance that must be chosen between maximising the overall accuracy of the model and correctly identifying all patients who will require a long stay.

We have shown that higher overall accuracy can be achieved by analysing the variables used to calculate EuroSCORE individually. Modelling each parameter as an independent variable demonstrated that some of these parameters had no statistical influence on the accuracy of the model. In particular, gender was not found to be a statistically significant variable in terms of accurately predicting LOS. The presence of either active endocarditis or ventricular septal rupture was also found not to influence accuracy, most likely due to the small number of patients with these conditions in the data set. By removing these variables from the model and treating the other parameters as individual variables, the model accuracy was improved compared to using EuroSCORE as a single variable.

Implementation of this model in clinical practice could influence the patient cohort selected for admission to the ICU. For example, by setting the EuroSCORE threshold to a low value, it would be possible to select patients who are highly likely to only require a short stay. Conversely, admitting patients with a EuroSCORE above a high threshold value would increase the likelihood of this cohort requiring a long stay. This strategy faces one challenge as patients with a high EuroSCORE have a higher mortality rate, and therefore average LOS could be lower than predicted if patients do not recover.

Whilst predictive models of this type have good accuracy for the patient population as a whole, they are generally poorer at predicting the outcome of an individual patient. (3,7,17) As such, the primary use of these models is in the assessment of efficiency and patient flow through ICUs, and comparison between units. There is also the potential to identify institutional factors that may influence ICU stay, which would have implications for costs of care and resource use. (7,18-20) The value of this model for resource planning is to enable cohorts of patients to be chosen so that the relative proportions of likely short stay and long stay patients remains constant. In practice, this could mean ensuring that only a certain number of high EuroSCORE patients are scheduled for elective procedures within a particular time frame. (4) The optimal proportions of short and long stay patients would need to be evaluated to ensure uninterrupted flow of patients through the unit and allowing flexibility for emergency cases.

The advantage of this method of predicting LOS over previous studies is multifactorial. Firstly, EuroSCORE is a cardiac specific scoring tool, which is now widely used in specialist centres and therefore predictions based on this will be straightforward to calculate. EuroSCORE also has the advantage over other scoring systems of taking into account peri-operative factors which can have a significant influence on LOS in previously well patients.(16) Our predictive model compares favourably with previous studies with an AUC equivalent to or greater than that calculated by other methods.(15, 21)

Unlike, previous studies, we have not only calculated the accuracy of predictions based on the EuroSCORE but have demonstrated the flexibility of the model to enable more accurate predictions in specific circumstances. For example, our model can be varied to ensure a higher proportion of long stay patients are identified to avoid unexpectedly high bed occupancy.

The strength of this model lies in its use in resource allocation and patient flow management. By selecting an acceptable proportion of short and long stay patients that are treated on the ICU at any particular time, the model can be used to select patients for elective surgery and to calculate the number of contingency beds required for emergency cases. This allows resources to be used as efficiently as possible by maintaining a relatively constant bed occupancy, whilst aiming to avoid bed shortages.

The main limitation of this study is the use of data from a single centre, and hence the results may not all be transferrable to all other intensive care units. However, the large data set means that the model is likely to be valid to similar units admitting patients following cardiac surgery but this needs to be tested. Furthermore, it would be straightforward to generate parameters specific to different units if similar data sets were available.

There are clearly other variables which will influence length of stay including many operative factors, however, for the purposes of this study we have chosen to focus on pre-operative scoring systems which will be simpler to evaluate for patient flow management.

Similarly, we have elected not to evaluate any temporal changes in length of stay over the period of data collection. This enables the use of a much larger data set on which to develop the model, however, integrating trends in length of stay over time would provide another area of interest for further research.

Finally, the model could be used to select an appropriate ratio of short stay to long stay patients described above, which would be transferable to other units that may have alternative ICU LOS prediction tools.

A number of options could be explored to further improve the accuracy of the model. The more recently published EuroSCORE II system has been shown to be a better predictor for cardiac mortality than either the additive or logistic EuroSCORE due to the changes in cardiac surgical mortality over time. (22) Whilst the logistic regression model is the standard initial choice for a problem of this type, there are alternatives that may improve accuracy, but require more rigorous statistical assumptions to be made.

Conclusions

Being able to predict LOS on the ICU has a number of advantages for both individual patients whose management can be planned accordingly, and for organisations, where resource utilisation can be optimised. Using a simple logistic regression model based on the cardiac operative risk scoring tool EuroSCORE, we have demonstrated that the LOS can be accurately predicted. With EuroSCORE data alone, the overall accuracy of the model is 79.77%. The model can be adapted to focus on the correct identification of long stay (or short stay) patients alone, which may be of more practical use for improved resource allocation.

Funding Statement

This work was supported by Papworth Hospital NHS Foundation Trust

Conflict of Interest Statement

Conflict of interest: none declared.

References

1. Wagener G, Minhaz M, Wang S, et al: The Surgical Procedure Assessment (SPA) score predicts intensive care unit length of stay after cardiac surgery. J Thorac Cardiovasc Surg 2011;142:443-50.

2. Andersen LW, Holmberg MJ, Doherty M, et al: Postoperative Lactate Levels and Hospital Length of Stay after Cardiac Surgery. J Cardiothorac Vasc Anesth 2015;29:1454-60.

3. Verburg IW, Atashi A, Eslami S, et al: Which Models Can I Use to Predict Adult ICU Length of Stay? A Systematic Review. Crit Care Med 2017;45:e222-31.

4. Ghotkar SV, Grayson AD, Fabri BM, et al: Preoperative calculation of risk for prolonged intensive care unit stay following coronary artery bypass grafting. J Cardiothorac Surg 2006;1:14-21.

5. Hein OV, Birnbaum J, Wernecke K, et al: Prolonged intensive care unit stay in cardiac surgery: risk factors and long-term-survival. Ann Thora Surg 2006;81:880-5.

6. Janssen DP, Noyez L, Wouters C et al: Preoperative prediction of prolonged stay in the intensive care unit for coronary bypass surgery. Eur J Cardiothorac Surg 2004;25:203-7.

7. Zimmerman JE, Kramer AA, McNair DS, et al: Intensive care unit length of stay: Benchmarking based on Acute Physiology and Chronic Health Evaluation (APACHE) IV. Crit Care Med 2006;34:2517-29.

8. Bukan RI, Moller AM, Henning MA, et al: Preadmission quality of life can predict mortality in intensive care unit--a prospective cohort study. J Crit Care, 2014;29:942-7.

9. Higgins TL, McGee WT, Steingrub JS, et al: Early indicators of prolonged intensive care unit stay: impact of illness severity, physician staffing, and pre-intensive care unit length of stay. Crit Care Med 2003;31(1):45-51.

10. Render ML, Kim HM, Deddens J, et al: Variation in outcomes in Veterans Affairs intensive care units with a computerized severity measure. Crit Care Med 2005;33:930-9.

11. Nashef SA, Roques F, Michel P, et al: European system for cardiac operative risk evaluation (EuroSCORE). Eur J Cardiothorac Surg 1999;16:9-13.

12. Roques F, Michel P, Goldstone AR, et al: The logistic EuroSCORE. Eur Heart J 2003;24:881-2.

13. R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Veinna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

14. Kohavi, R: A study of cross-validation and bootstrap for accuracy estimation and model selection. Proc 14th IJCAI 1995;2:1137-1143

15. Ettema RGA, Peelen LM, Schuurmans MJ, et al: Prediction models for prolonged intensive care unit stay after cardiac surgery. Circulation 2010;122:682-9.

16. Nilsson J, Algotsson L, Hoglund P, et al: EuroSCORE predicts intensive care unit stay and costs of open heart surgery. Ann Thorac Surg 2004;78:1528-35.

17. Marik PE, Hedman L. What's in a day? Determining intensive care unit length of stay. Crit Care Med 2000;28:2090-3.

18. Woods AW, MacKirdy FN, Livingston BM, et al: Evaluation of predicted and actual length of stay in 22 Scottish intensive care units using the APACHE III system. Acute Physiology and Chronic Health Evaluation. Anaesth 2000;55:1058-65.

19. Vasilevskis EE, Kuzniewicz MW, Cason BA, et al: Mortality probability model III and simplified acute physiology score II: assessing their value in predicting length of stay and comparison to APACHE IV. Chest 2009;136:89-101.

20. Verburg IW, de Keizer NF, de Jonge E, et al: Comparison of regression methods for modelling intensive care length of stay. PLoS One 2014;9:e109684.

21. Messaoudi N, De Cocker J, Stockman BA, et al: Is EuroSCORE useful in the prediction of extended intensive care unit stay after cardiac surgery? Eur J Cardiothorac Surg 2009;36:35-9.

22. Nashef SA, Roques F, Sharples LD, et al: EuroSCORE II. Eur J Cardiothorac Surg 2012;41:734-5.

Figure Captions

Figure 1: Density plots (a,c) and box plots (b,d) of EuroSCORE (additive) for different length of stay categories: short, intermediate and prolonged (a,b) or short and long (c,d).

Figure 2: Graph of threshold probability with EuroSCORE (additive) showing effective classification threshold of 11.35 (a) and density plot of EuroSCORE (additive) values for short and long stay categories with classification threshold indicated (b).

Figure 3: ROC curve for the classifier η based on EuroSCORE (additive). Black dot indicates $\eta = 0.5$. The dashed line corresponds to a random classifier.

Table 1:

Summary of patient demographics from the ICU LOS data set

Gender	Male 13,169	Female 5,802	Total 18,377	
Age	Minimum 16	Maximum 96	Mean 68	Median 70
LOS (days)	Minimum 0	Maximum 123	Mean 2.4	Median 1.0
Type of surgery	Elective 13,637	Urgent 3,822	Emergency 918	

Table 2:

Contingency table with classifier $\eta = 0.5$ for logistic regression model using EuroSCORE (additive)

	Predicted		
Observed	short	long	Total
short	13849 (75.3%)	372 (2.0%)	14221 (77.3%)
long	3345 (18.2%)	808 (4.4%)	4153 (22.6%)
Total	17194 (93.6%)	1180 (6.4%)	18374 (100%)

Table 3:

Overall accuracy, false positive rate and true positive rate for varying values of the classifier η

η	Accuracy	False Positive Rate	True Positive Rate
0.50	79.77%	0.0262	0.1946
0.25	72.53%	0.2302	0.5728
0.10	42.79%	0.7104	0.9015
0.05	25.16%	0.9643	0.9909

Figure 1

Click here to download high resolution image

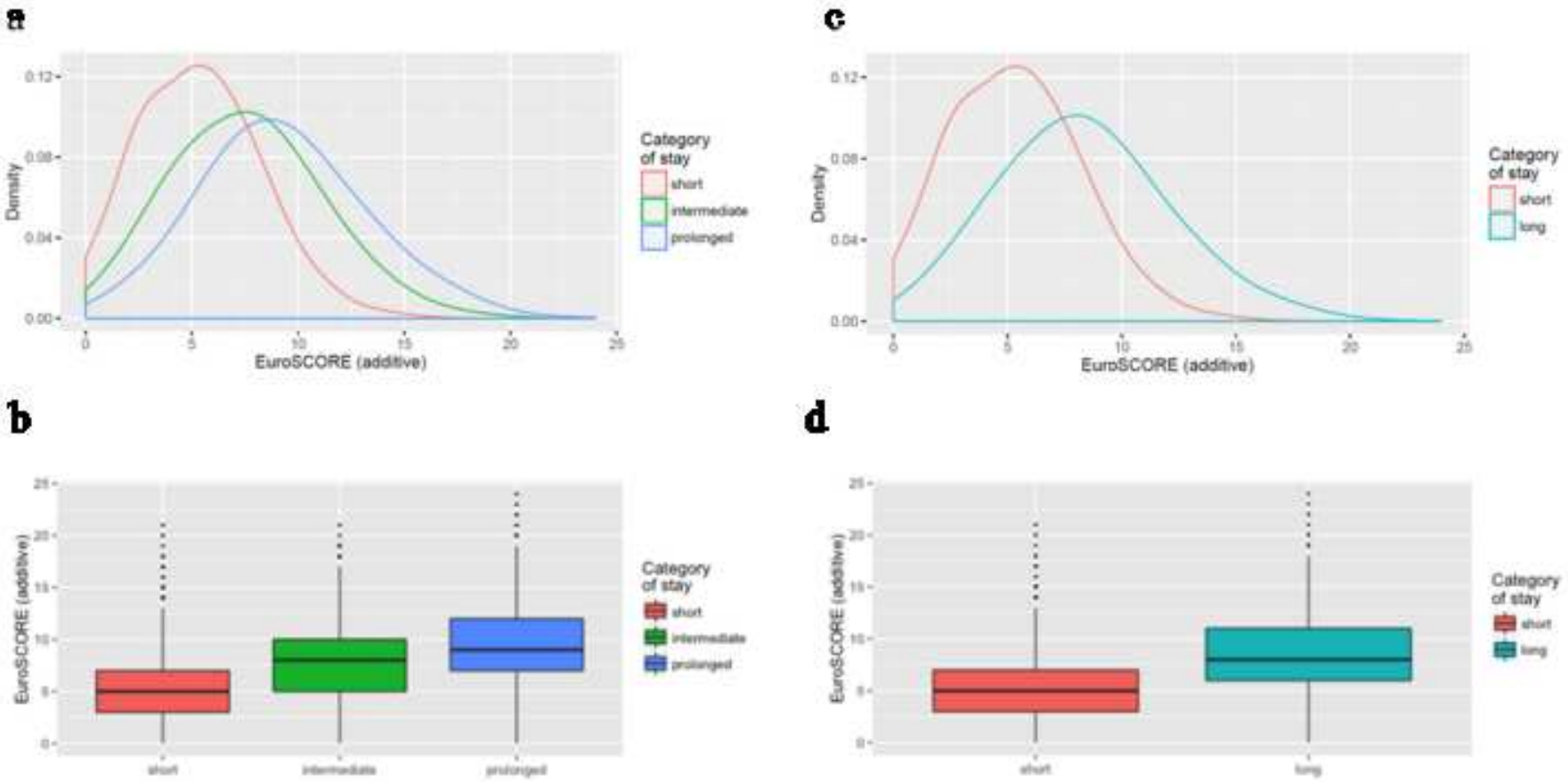

Figure 2
Click here to download high resolution image

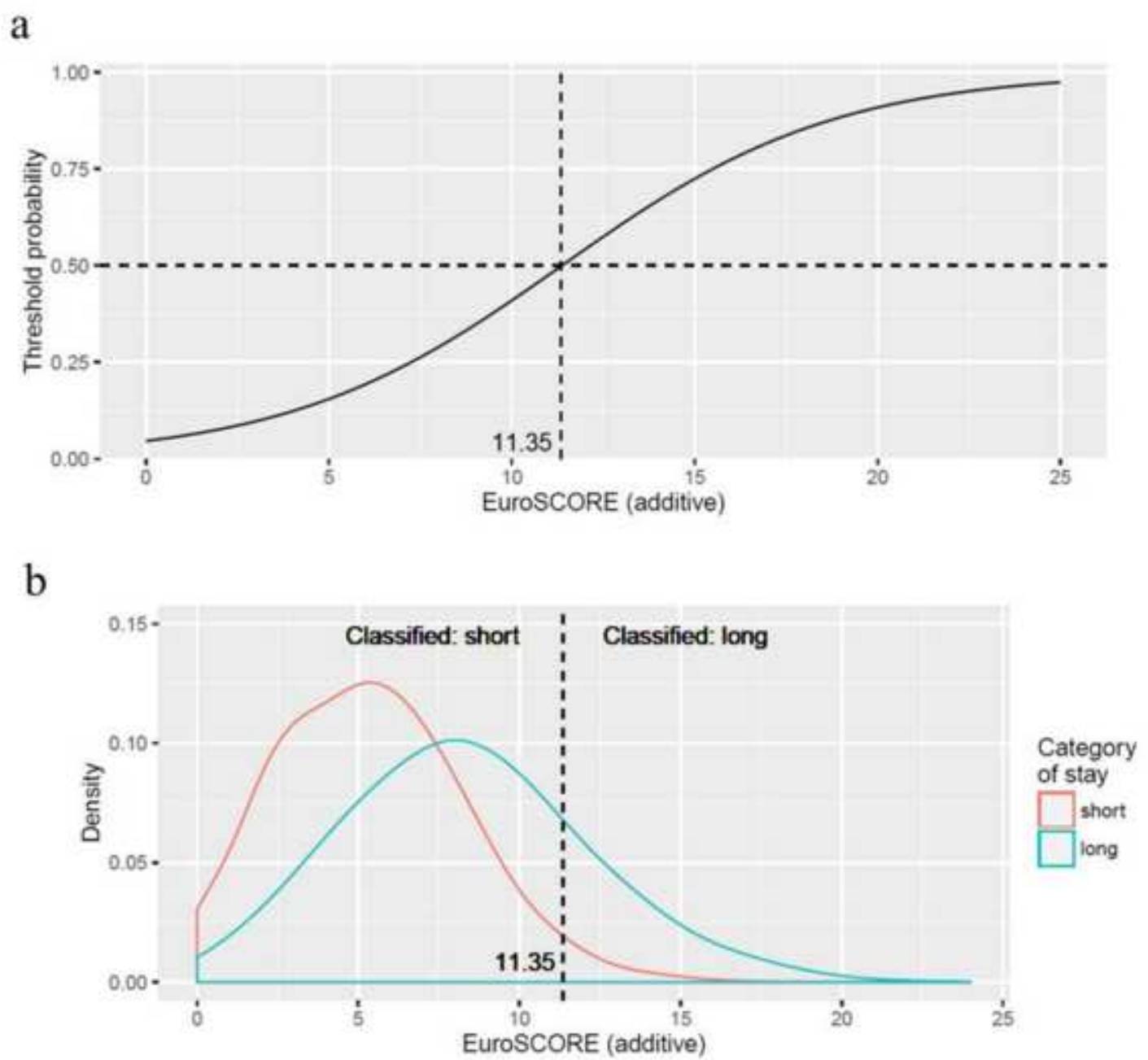
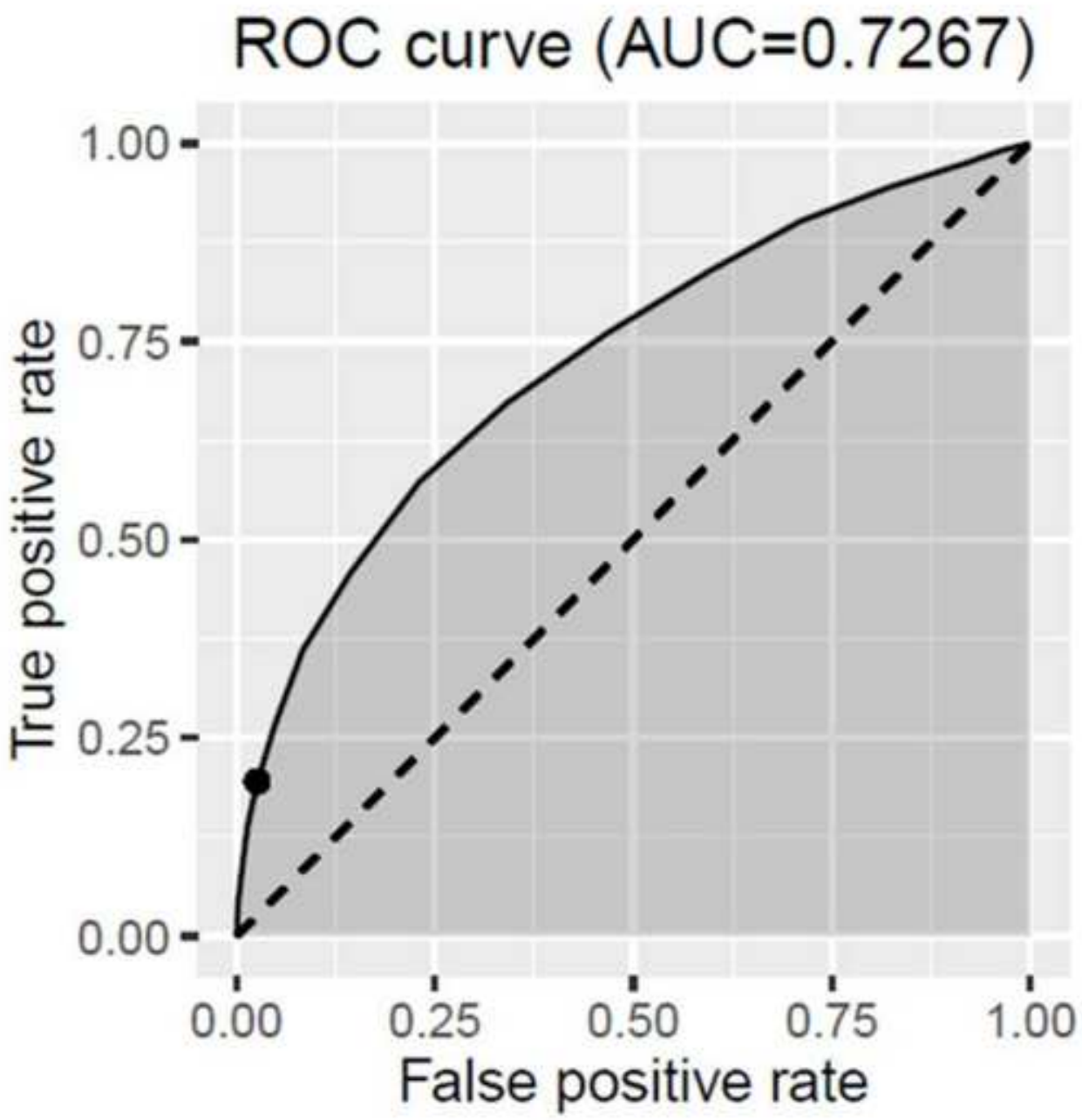
Figure 3
Click here to download high resolution image