Comparison or different statistical models for Intensive Care length of stay

I.W.M. Verburg MSc, N.F. de Keizer PhD, N. Peek PhD

Department of Medical Informatics, Academic Medical Center, University of Amsterdam, the Netherlands

Introduction Intensive Care Units (ICUs) are increasingly interested in assessing and improving their performance. Since ICU patients receive complex care, costs are high and hospitals face continuous pressure to both improve quality and reduce costs. Until now benchmarking has been used in the process of continuous quality improvement in ICUs, often using hospital mortality as an indicator for ICU quality. Benchmarking is less frequently used to compare efficiency among ICUs. Since Length of Stay of ICU patients (LoS) is related to ICU costs it can serve as a marker for efficiency. To compare LoS between different hospitals, one should adjust for differences in case mix. However, only few models exist to predict case mix adjusted LoS and no consensus exists which method should be used to model LoS for benchmarking purposes [1-3]. The aim of this study was to compare the performance of different statistical regression methods to predict LoS for unplanned ICU admissions.

Methods The Dutch National intensive Care evaluation (NICE) quality registry records data from intensive care patients in the Netherlands, including severity-of-illness data from the first 24h of ICU stay, and outcomes such as mortality and LoS. This study used data from admissions between January 1st, 2009, and December 31st, 2011 from 84 ICUs. All patients satisfying the APACHE IV inclusion criteria were included, excluding patients that underwent elective surgery. As the distribution of LoS and the association between patient variables and LoS differs between survivors and non-survivors, separate models were developed for these two groups of patients.

The distribution of LoS is highly skewed. Models which are available in literature often make use of Ordinary Least Squares (OLS) regression of LoS or OLS regression of log-transformed LoS. Theoretically, these are suboptimal choices. In this study we compared OLS regression of LoS and of log-transformed LoS with Generalised Linear Model (GLM) regression with Gaussian, Gamma, Poisson and negative binomial distribution families and a logarithmic link function, and with Cox proportional hazards (PH) regression. LoS was defined as the number of fractional days spent at ICU. Each model started with a set of variables known to be associated with LoS [1;4]. Subsequently, the models were simplified by stepwise backwards elimination of variables. Several performance measures were used to evaluate each model's ability to predict LoS, being the squared Pearson correlation coefficient (R^2), the root mean squared prediction error (RMSPE), the relative mean absolute prediction error (relative MAPE), and the relative bias (mean difference between predicted and observed LoS divided by mean observed LoS). In general, good predictions yield low values for the RMSPE, the relative MAPE and the relative bias, and high values for the R^2. All performance measures were calculated on the same dataset which was used to develop the models, and afterwards corrected for optimistic bias that was estimated from 100 bootstrap samples [5].

Results From January 1st, 2009, until December 31st, 2011, 222,529 ICU admissions were recorded in the NICE database. After applying all exclusion criteria, 94,251 (42.4%) admissions remained. Of these, 81,190 (86.1%) survived ICU stay and 13,061 (13.9%) died.

Performance statistics for each of the models are shown in the table below. GLM models showed the best performance, while the Cox PH model had poorest performance. The model based on OLS regression of log-transformed LoS had a large relative bias. All models showed better results for survivors than for non-survivors, especially for R^2 and the relative bias.

	ICU survivors				ICU non-survivors			
	R^2	RMSPE	relative BIAS	relative MAPE	R^2	RMSPE	Relative BIAS	relative MAPE
OLS regression LoS	0.174	7.448	0.008	0.812	0.107	9.618	0.005	0.891
OLS regression log LoS	0.183	7.714	-0.400	0.674	0.107	10.213	-0.510	0.762
GLM-Gaussian	0.197	7.335	0.001	0.771	0.134	9.462	-0.009	0.868
GLM-Poisson	0.194	7.349	0.000	0.769	0.128	9.504	0.000	0.872
GLM-Negative binomial	0.186	7.388	0.005	0.773	0.120	9.545	-0.001	0.872
GLM-Gamma	0.184	7.407	0.005	0.773	0.112	9.602	-0.001	0.877
Cox PH regression*	0.097	9.002	-0.693	0.938	0.075	11.388	-0.808	0.906

*Performance estimates were not corrected for optimistic bias.

Discussion From our results we conclude that GLM-Gaussian and GLM-Poisson are the best choices to model LoS of unplanned ICU admissions, closely followed by plain OLS regression of untransformed LoS. Cox PH regression and OLS regression of log-transformed LoS are better avoided. A limitation of this study is that differences in performance were not statistically tested. Our study confirms the results by Austin et al. that was performed in a cohort of patients undergoing (elective) CABG surgery [6]. Future research will focus on developing a model which can be used for benchmarking purposes.

Reference List

[1] Vasilevskis EE, Kuzniewicz MW, Cason BA, Lane RK, Dean ML, Clay T, 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 Jul;136(1):89-101.

[2] Woods AW, MacKirdy FN, Livingston BM, Norrie J, Howie JC. 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. Anaesthesia 2000 Nov;55(11):1058-65.

[3] Zimmerman JE, Kramer AA, McNair DS, Malila FM, Shaffer VL. Intensive care unit length of stay: Benchmarking based on Acute Physiology and Chronic Health Evaluation (APACHE) IV. Crit Care Med 2006 Oct;34(10):2517-29.

[4] Kramer AA, Zimmerman JE. A predictive model for the early identification of patients at risk for a prolonged intensive care unit length of stay. BMC Med Inform Decis Mak 2010;10:27.

[5] Steyerberg EW. Overfitting and Optimism in Prediction Models. Clinical Prediction models: A practical Approach to Development, Validation, and updating. Springer; 2009. p. 83-99.

[6] Austin PC, Rothwell DM, Tu JV. A Comparison of Statistical Modeling Strategies for Analyzing Length of Stay after CABG Surgery. 2002 Jun 1;(2):107-33.