

A comparison of two alternative methodologies for estimation of length of hospital stay

A.E. Smith¹, V. Stankovski² & D.M. Marsh¹

¹*Queen's University, Belfast, Northern Ireland*

²*University Medical Centre, Ljubljana, Slovenia*

Abstract

This study utilised a large audit database from a trauma unit, to model the length of stay of the patient in the unit, i.e. bed occupancy, for all those with upper and lower limb fractures. For this purpose, we applied two methodologies: a decision tree approach utilising Quinlan's *See5* machine learning environment was compared with a modification of Cox's regression, a statistical approach and the two systems evaluated for classification accuracy. Data from 747 lower limb and 487 upper limb cases were analysed, with the model being extracted from key parameters, which were retrieved from the original database. We also used information from the medical personnel involved and indicators of significance from preliminary studies. Exactly the same model and parameters were utilised in both approaches to ensure comparability. Results are given for the classification performance for both approaches, with respect to the length of stay.

1 Introduction

Estimating the length of stay of patients in hospitals is important for patient treatment management, bed occupancy management and cost effectiveness measures. Mathematical models have been produced as decision aids [1] and many trauma scores give an indication of patient survival probability [2, 3] and severity of trauma indices [4]. All of these were considered and incorporated into the model, where appropriate.

The data set for this study has arisen from a very large audit database that has been set up to record all details of patients passing through a large regional Trauma Unit in Belfast, Northern Ireland.

2 Methodology

2.1 Data preparation and model formation

The original database is a relational database in ACCESS, which means that, for the purposes of research the database had to be addressed in SQL, extracting all the variables / attributes of interest. The result has many duplicates because of the multiplicity of relational tables and so the removal and cleaning up of the model in a flat form has been carried out.

We utilised two models - (1) All upper limb fractures. (2) All lower limb fractures. The primary approach is concerned with the factors that influence the stay of the patient in the unit, or in other words, bed occupancy, for all those with lower limb fractures. We extracted attributes considered possibly relevant, by the Trauma Research Group, along with demographic details. The aim was to classify outcomes in term of length of stay and discover the risk factors for this.

Our model for prediction consisted of the following attributes. (a) A deprivation index computed from the locality codes in the area and social coding for these (*Depriv*). (b) The age of the patient at the time of their accident (*Ageatacc*). (c) The sex of the patient (d) The Accident Severity Index, a nationally recognised score (*Aisscore*). (e) The Physical (*PCS*) and (f) Mental Condition (*MCS*) scores on entry to the unit. (g) The level of training of the most senior clinician involved (*Consult*) (Senior House Officer = 1, Senior Registrar = 2, Consultant = 3). (h) Whether the patient had an operation or was conservatively treated (*OPCS*). (i) The level of cigarette consumption (*Cigs*) and (j) the level of alcohol consumption (*Alco*). The benchmark was actual length of stay in the unit.

A randomly chosen five fold leave-n-out train and test approach was employed to indicate consistency of results. All these attributes were used in each model and each fold of the model to ensure that direct comparability was maintained between all of the methodologies used. Some attributes e.g., injuryICD10 scores proved to be impracticable to be included in the model, since there are dozens of these categories, which would result in dozens of tiny data sets, which would not achieve any power in prediction. The trauma research team considered that the aisscore reflected these sufficiently and gave some degree of the severity of the injuries.

Several outliers in the data were found, that the Trauma team identified as being specifically unusual, e.g. a 94 year old lady who was in for 60 days was considered to be skewing the data. Therefore it was decided to take a 95% percentile, i.e. 95% of the patients, this gave maximum stay for lower limbs of <25 days and upper limbs of <20 days (spine and others). Only 5% of patients remained after this time and they were considered to be untypical.

2.2 Statistical approaches

An adaptation of Cox's regression to give a time estimate for each patient in a data set was utilised [5]. This gave a formula (1) which, when written into syntax produced an automated output, in order to make this directly comparable with the machine learning model.

Here B_i s are the partial regression co-efficients and X_i 's the values of the covariates for each patient. m is the slope of the fitted baseline derived from the data, and t is the time attribute being estimated. The natural log of the probability ($\ln p$) is taken at 0.5, the half life of the time curve. By this method time estimates in days or part thereof were produced.

$$\hat{t} = [\exp(\exp(-\sum_{i=1}^n B_i X_i) * \ln p) - 1] / m \quad (1)$$

2.3 Machine learning techniques

In contrast to the traditional medical methodology, the modified methodology for medical research [5] exploits the flexibility of machine learning and retains the kind of statistical tests that are generally accepted in the medical field for the confirmation of hypotheses. First, the medical problem is defined and data for an observed population are collected; then a machine learning tool is used to generate hypotheses regarding the problem; finally, statistical methods are used to determine the validity of the generated hypotheses.

Decision trees and their creation has been defined. "The traditional approach to constructing a decision tree from a training set of cases described in terms of a collection of attributes is based on successive refinement. Tests on the attributes are constructed to partition the training set into smaller subsets until each subset contains cases belonging to a single class. These tests form the interior nodes of the decision tree and each subset is associated with one of its leaves. An unseen case is classified by tracing a path from the root of the tree to the appropriate leaf and asserting that the case belongs to the same class as the set of training cases associated with that leaf" [6]. Each interior node in the tree represents a statistical hypothesis as described in [5].

The decision trees for the purpose of this research were generated by using the *See5* environment [7], which is a PC application that uses the C5.0 algorithm. There has been much work done on decision trees since the 70s, both by Quinlan [6-8] and others in the machine learning communities [5, 10].

3 Results

3.1. Initial results of statistical multivariate model

The most significant attributes for length of stay in hospital are *Ageatacc*, *AISscore*, *Consult*, *MCS*. However, all the attributes in the model were retained for each methodology and to enable direct comparisons to be made. The results for the Cox's regression adaptation for the two complete datasets are given in Table 1.

With *Cox's regression* prediction of risk factors for length of stay, care in interpretation has to be taken, because negative results for hazards actually indicate a shorter time - to - event.

Table 1. Cox's regression

(Lower limb)

attribute	Regr. Coeff.	Sig.	Odds Ratio	(95% C.I.)
Depriv	-.0022	0.5436	0.9978	(0.9908,1.0049)
Ageatacc	-.0092	0.0010	0.9908	(0.9854,0.9963)
Sex	.1334	0.1163	1.1428	(0.9674,1.3499)
Aisscore	-.2551	<0.0001	0.7748	(0.6850,0.8764)
Pcs	-5.19E-04	0.9179	0.9995	(0.9897,1.0094)
Mcs	.0094	0.0730	1.0095	(0.9991,1.0200)
Consult	-.2493	0.0016	0.7794	(0.6678,0.9096)
Cigs	.0038	0.5348	1.0038	(0.9919,1.0158)
Alco	-6.02E-04	0.9222	0.9994	(0.9874,1.1115)
Opcs	-.0871	0.2566	0.9166	(0.7886,1.0654)

(Upper limb)

attribute	Regr. Coeff.	Sig.	Odds Ratio	(95% C.I.)
Depriv	-4.36E-04	0.9216	0.9996	(0.9909,1.0083)
Ageatacc	-.0050	0.1484	0.9950	(0.9883,1.0018)
Sex	-.1058	0.3745	0.8996	(0.7122,1.1362)
Aisscore	-.3219	<0.0001	0.7313	(0.6355,0.8415)
Pcs	.0029	0.6820	1.0029	(0.9889,1.0171)
Mcs	.0110	0.1077	1.0111	(0.9976,1.0248)
Consult	-.2705	0.0062	0.7630	(0.6287,0.9260)
Cigs	-.0025	0.7325	0.9975	(0.8931,1.0121)
Alco	.0032	0.6688	0.9975	(0.9886,1.0181)
Opcs	-.1663	0.1002	1.1809	(0.9685,1.4399)

3.2 Initial results of machine learning models

A decision tree model for the *Orthosta* class (length of stay in hospital) for the lower limb injuries is presented (for demonstration purpose) in Figure 1. The tree model for upper limbs is similar.

The induced model can be interpreted as follows. If the level of training of the doctor involved is lower (*Consult* [0-2]) and the patient is younger (*Agetacc* ≤ 55 years), then the bed is occupied for up to 5 days. Conversely, if the patients are older than 55 years then they would occupy the bed from 6 to 10 days. In this subtree, the bed occupancy is related to the age of the patient and the level of the training of the doctor involved. On the other hand, if the level of training of the doctor is higher (*Consult* = 3), the second most important attribute is the Accident Severity Index (*Aisscore*). Further interpretation of the induced decision tree shown in Figure 1 is left to the reader.

So, the modified methodology for medical research [5] can be suitably employed in the clinical decision making process, since it is naturally understood by medical professionals.

Figure 1: A decision tree model for length of stay in hospital for lower limb injuries

```
See5 [Release 1.15]

Decision tree:

Consult in [0-2]:
...Agetacc <= 55: 5 (63/31)
: Agetacc > 55: 10 (16/6)
Consult = 3:
...Aisscore in [3-4]:
...OPCS = 0: 10 (24/10)
: OPCS = 1:
: ...MCS <= 58.01: 10 (31/22)
: MCS > 58.01: 15 (17/9)
Aisscore in [1-2]:
...Agetacc <= 34:
...Alco > 3: 10 (53/23)
: Alco <= 3:
: ...Cigs <= 0: 10 (30/17)
: Cigs > 0: 5 (24/12)
Agetacc > 34:
...Alco > 15: 15 (15/8)
: Alco <= 15:
: ..Depriv> 6.06: 5 (15/10)
: Depriv <= 6.06:
: ..PCS<=55.06:10 (58/27)
: PCS> 55.06: 5 (28/19)
```

In order to improve the acceptance of the results in the medical field, in addition to the study of the prediction error of the induced model, the validity of the generated hypotheses should also be tested by using appropriate statistical tests [5].

Our results show that the key parameters influencing the length of stay in hospital are: the level of training of the doctor involved (*Consult*), the age of the patient at the time of the accident (*Agetacc*), and also the Accident Severity Index (*AISscore*).

3.3 Comparison of the two alternative methodologies

Decision tree methodology gives results that fall into correct classification or not, so the Cox's results were adjusted to be classified in the same manner, that is within 5 days if the true length of stay was classified as being correct and incorrect if otherwise. Results for each of the five folds in the data that we employed, is presented in Table 2. The overall classification result is 71.5% for the Cox's adaptation and 50.7% for the decision tree approach.

Table 2. The correct classification results

Fold	# 1	# 2	# 3	# 4	# 5
Cox adaptation (LL)	58.5	68.1	68.7	61.7	52.8
" " (UL)	87.4	77.0	83.5	79.6	78.1
Decision Tree (LL)	31.0	39.1	41.5	42.6	33.3
" " (UL)	63.2	62.1	73.6	61.3	59.4

Upper limb = UL Lower limb = LL

Apart from some differences in the classification, both techniques defined *AISScore*, *Consult* and *Agetacc* as being the key attributes that determine the length of stay in the Trauma unit.

4. Conclusions and further work

Testing for the differences in the distribution between the estimates produced by the decision tree approach and the statistical approach, compared with the actual time of stay in the trauma unit was limited due to the nature of the results so only classification percentages are given. Exact statistical testing is therefore not possible. We plan, however, to use regression tree approaches to give direct comparability between methodologies that can be utilised to give time-to-event estimations for patients' bed occupancy/length of stay in hospital units. We hope to gain an optimisation of the estimates rather than a five-day classification window.

Cox's regression techniques may be less instinctively obvious to medical personnel, since it is an iterative process and therefore lacks clarity. The advantage of utilising decision trees is in combining both an approach that is easily understood by clinicians, i.e. machine learning techniques and some statistical testing stringency. It is hoped in this way to produce hybrid techniques that will make such data mining more acceptable to the medical field.

References

- [1] McClean SI. Modelling and simulation for health applications. In: Modelling hospital resource use. RSM Press, London 1994; pp21-28
- [2] Champion HR, Sacco WI, Carnazzo AJ, et al. Trauma score. Crit Care Med. 1981;9; 672-
- [3] Boyd CR, Tolson MA, Cope WS. Evaluating trauma care. The TRISS method. J Trauma 1987; 27: 370-392
- [4] Champion HR, Copes WS, Sacco WI. A new characterisation of injury severity. J Trauma 1990; 30; 539-545
- [5] Stankovski V, Bratko I, Demšar J, Smrke D. Induction of hypothesis concerning hip arthroplasty: A modified methodology for medical research, Methods Inf. Med. 2001; 40; 392-6
- [6] Quinlan JR. Induction of decision trees, Mach. Learn. 1996; 1; 81-106
- [7] Quinlan JR. Data mining tools C5.0 and See5. 1996: Rulequest Research, Web Ref: <http://www.rulequest.com/see5-info.html>
- [8] Quinlan JR. Programs for machine learning. Morgan Kaufmann, 1993
- [9] Smith AE, Anand SS. Patient survival estimation with multiple variables: adaptation of Cox's regression to give an individual's point prediction. Proc. ECAI, IDAMAP workshop. Berlin. Aug. 2000. p 51-54
- [10] Brieman L, Friedman JH, Olshen RA et al. Classification and regression trees. Wadsworth, Belmont, California, 1984