Patients flow: a mixed-effects modelling approach to predicting discharge probabilities.

Shola Adeyemi
Thierry Chaussalet
Haifeng Xie
Peter Millard

School of Informatics

Copyright © [2007] IEEE. Reprinted from the proceedings of the 20th IEEE International Symposium on Computer-Based Medical Systems. IEEE CBMS 2007, Maribor, Slovenia, 20-22 June 2007. IEEE, Los Alamitos, USA, pp. 725-730. ISBN 0769529054.

Patients Flow: A Mixed-Effects Modelling Approach to Predicting Discharge Probabilities

Shola Adeyemi, Thierry Chaussalet, Haifeng Xie and Peter Millard
Health and Social Care Modelling Group, School of Informatics
University of Westminster, London, UK
s.adeyemi3@westminster.ac.uk, (chausst, xieh)@wmin.ac.uk ,phmillard@tiscali.co.uk

Abstract

A mixed effects approach is hereby introduced to patients flow and length of stay modelling. In, particular, a class of generalized linear mixed models has been used to demonstrate the usefulness of this approach. This modelling technique is used to capture individual patients experience during the process of care as represented by their pathways through the system. The approach could predict the probability of discharge from the system, as well as detect where the system may be going wrong.

1. Introduction

Average length of stay (LOS) has been shown to be a good proxy for hospital resource consumption and disease severity [1]. In the literature, LOS in health and social care systems has been described using patient flow models. Modelling patients flow in health care systems is considered to be vital in understanding the operational and clinical functions of the system and has proved to be useful in improving the functionality of the health care system. These papers have used different techniques to model patient flow in health care systems with several different assumptions imposed on the system. Markov models have been used extensively to capture probabilistic laws that govern the dynamics of patients between states in the healthcare system. A two-stage continuous-time Markov model that describes the movement of patients through two compartments in geriatric hospitals was developed in [2]. Such an approach takes into account, conceptually, different types of patients and their corresponding length of stay. The (Markov) model [2] was extended to describe the behaviour of patients moving through three stages in a geriatric department [3]. A Markov model in continuous time for the LOS of elderly people moving within and between residential home care and nursing home care has also been developed using the framework of aggregated Markov processes [4]. Phase-type distributions have also been employed to represent the variable nature of LOS. This class of distributions describes the time to absorption of a finite Markov chain in continuous time, where there is a single exit (absorbing state) and the stochastic process starts in a transient state, [5]. Figure 1, below, describes the Coxian phase-type distribution where the absorbing state is discharge by death.

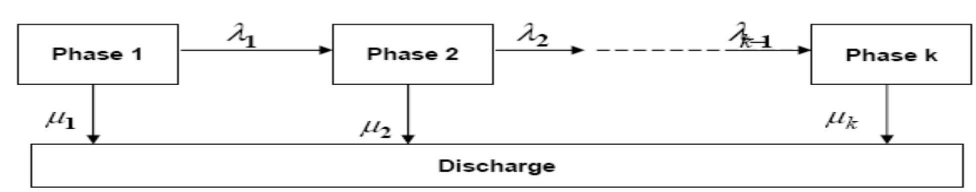

Figure 1: An Illustration of the Coxian phase-type distribution

Twentieth IEEE International Symposium on Computer-Based Medical Systems (CBMS'07)
0-7695-2905-4/07 $20.00 © 2007 IEEE

These models describe duration until an event occurs in terms of a process consisting of a sequence of latent phases- the states of a latent Markov model. However, the generality of phase-type distributions makes it difficult to estimate all the parameters of the model. To overcome this problem Coxian Phase-type distributions were introduced. Coxian phase-type distributions describe duration until an event in terms of a process which consists of a sequence of latent phases. Compartmental and simulation modelling has been an approach for modelling LOS [6], The compartmental framework has also been incorporated into queuing theoretical approach however, due to the complexity, discrete event simulation (DES) has been employed to solve the system equations [7]. It should be stressed that the compartments were virtual therefore the analyses could only be considered a hypothetical system rather than a real one.

In almost all studies of patient flow and LOS modelling what has been overlooked is what an individual patient experiences during the delivery of care [8]. In order to have an insight into what an individual patient experience during the course of care, we seek to develop a mixed modelling framework for patient flow paths through the healthcare system. This approach could predict the probability of discharge from the system, as well as detect where there may be problems with flow.

2. Modelling Approach

In the healthcare system, either the improvement of patients depends on activities (or interventions) in each state visited in the care process or the recovery of patients progresses in stages (states) until discharge. According to [9], this represents patient journeys (flow paths) that a particular clinical process may entail which make patients pass through the health system and attain different health-states. These can, for example, also reflect the physical progression of patients between different hospital locations. Since our aim is to model the experience of individual patients in the healthcare system, these experiences are the different states (locations or health states) visited captured by individual patient's flow paths through the system. In a multi-state system, the flow paths represent clustered or repeated observations within an individual patient. When repeated measurements are taken on patients, classical regression assumptions are violated; therefore mixed models are developed to model the outcome in a view to capture the correlation structure induced and the patient specific random effects. There are two distinct approaches to the analysis. First, the heterogeneity can be explicitly modelled; we will refer to this as the 'patient-specific' approach. These patient specific effects are assumed to follow a parametric distribution across population, usually normal. Second, the population-averaged response can be modelled as a function of covariates without explicitly accounting for patient to patient heterogeneity. A mixed effect model satisfies, [10]

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b} + \boldsymbol{\varepsilon} \tag{2.1}$$

where

$\mathbf{Y}$ is the vector of different combinations of path ways

$\mathbf{X}$ is the design matrix for the fixed (patients' population) effects

$\boldsymbol{\beta}$ is the vector of fixed (average patients' population) effect parameters

$\mathbf{Z}$ is the design matrix for the random (patient specific) effects

$\mathbf{b}$ is the vector of random (patient specific) effects

$\boldsymbol{\varepsilon}$ is the vector of experimental (measurement) errors

Twentieth IEEE International Symposium on Computer-Based Medical Systems (CBMS'07)
0-7695-2905-4/07 $20.00 © 2007 IEEE

The regression coefficients (fixed effect parameters) have interpretations for the population rather than for any individual. If Y is Gaussian, the model is referred to as a linear mixed model while a generalized linear mixed model whenever Y is non-Gaussian. The linear mixed model and the generalized linear mixed models are linear both in the predictor and the parameter however when Y is non-Gaussian and the mean is modelled as a nonlinear function of the parameters and the predictor, the model (2.1) is referred to as a nonlinear mixed model. In this paper, we demonstrate a class of the generalized linear mixed models.

3. Illustration of the Approach

The figure below depicts activities within an artificial system. This system assumes that patients entering into the system will undergo some interventions and then progress from doing primary, secondary and discharge planning activities before discharge. Discharge can be by death, discharge to patients' home or transfer to another healthcare institution.
For simplicity we have forced each state of the system to have three exits, this is not a restriction for using this approach however, at least two exit paths are necessary.
The coding of the paths are ordered categories which gives the highest level of improvement by patients for state *i* as $\gamma_{i2} = 2$, and $\gamma_{i1} = 1$ while $\gamma_{i0} = 0$ is the lowest level of improvement.

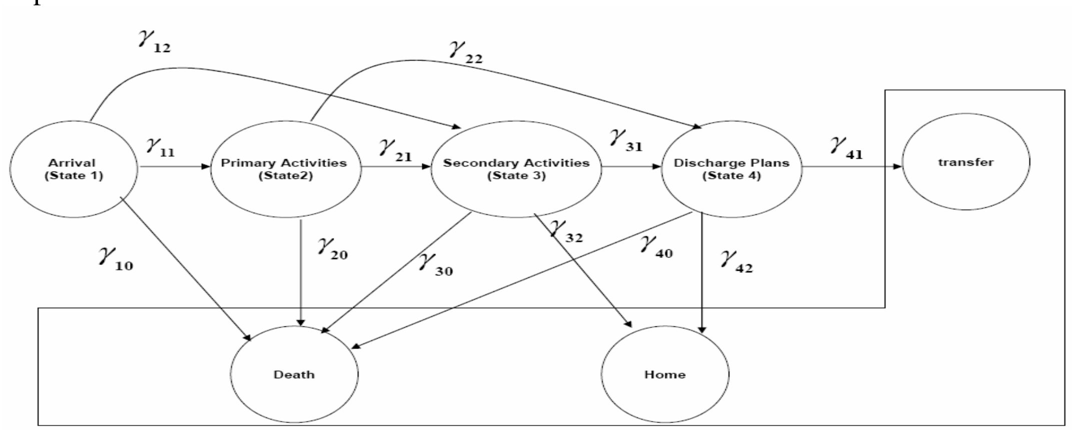

Figure 2: Disease progression through an artificial system- Adapted from Cote 2000

The probability distribution of potential exit (or discharge) from the states through any of the paths (γ_{ij}, *i*=1, . .,4 and *j*=0, 1, 2) in Figure 1, is modelled by

$$\pi_{pyi} = \frac{\exp\sum_{u=0}^{y}\left(\theta_p - \beta_{iu}\right)}{\sum_{k=0}^{m_i-1}\exp\sum_{u=0}^{k}\left(\theta_p - \beta_{iu}\right)}; \quad y = 0,1,........,m_i - 1. \qquad (1)$$

The probability of patient *p* passing through *y* (i.e. through any γ_{ij}) on the m_i -path state *i* is a function of the person's position on θ (person specific random effect) and the difficulties of the m_i paths in state *i*. The observation *y* can be seen as a count of surpassed paths, and only the difficulties of these *y* surpassed paths appear in the numerator of the model. The score, number of surpassed paths in the system, for each patient is a sufficient statistics for $\boldsymbol{\theta_p}$, [11] distributed normally with mean 0 and variance $\sigma_\theta^{\ 2}$. The β_i s show the relative ease (or difficulty) of passing through each state by different patients having person specific random effect (e.g. health status). These random effects scaled between -3 and +3.

The Table below gives the parameter estimates for the model proposed in (1) based on maximum likelihood estimation using non-adaptive Gaussian quadrature with 20 nodes and implemented in the SAS V9.

Table 1: Parameter estimates for the model presented in (1)

Parameter	Estimate	Standard Error	Degree of Freedon	T Value	p-values
β_1	-1.7047	0.4671	59	-3.65	0.0006
β_2	0.6931	0.2582	59	2.68	0.0094
β_3	2.3026	0.3464	59	6.65	< 0.0001
β_4	1.7918	0.2981	59	6.01	<0.0001

Presented in the figures (Figures 3 – 6) below are the discharge curves for each state. These curves are model based i.e. obtained from the estimated parameters. On the x-axis are the patient specific random effects theta and some threshold parameters. These figures give the probability of discharge with respect to the location of patients on the health status (theta) scale. Because each person can only pass through one path and the three options are mutually exclusive, the sum of the three probabilities at a particular value of theta is 1. The probability of discharge, from the system, with respect to the location of patients on the health status (theta) scale can be calculated. Suppose θ_{pi} is the value of the random effect for patient *p* in state *i* and γ_{ij} is the observable (ordinal variate) paths *j* through *i*. Under the model, there are thresholds values, $\beta_{i\mathbf{1}}, and \beta_{i\mathbf{2}}$ such that

$$\gamma_{ij} = \begin{cases} 0 & \text{if } \theta_\text{p} < \beta_{i1} \\ 1 & \text{if } \beta_{\text{i}1} \le \theta_p < \beta_{i2} \\ 2 & \text{if } \theta_\text{p} \ge \beta_{i2} \end{cases} \qquad i = 1,....,4; j = 1,2,3$$

For instance, observing state 1 (Figure 3), path 0 and path 1 intersect at $\beta_{i1} = -1.5$ and path 1 and path 2 intersect at $\beta_{i2} = 0\,.1$.

These thresholds show that the most likely path is path 0 below β_{i1} and path1 is the most likely between β_{i1} and β_{i2} while path2 is the most likely above β_{i2}. Note that if $\beta_{i2} < \beta_{i1}$, the interpretation changes with path 1 not being most likely at any point. This might be interpreted as indicating a problem with path 1, but this is not always true. State2 (Figure 4), shows the threshold have changed. β_{i1} has changed from -1.5 in state 1 (Figure 3) to -0.6

Twentieth IEEE International Symposium on Computer-Based Medical Systems (CBMS'07)
0-7695-2905-4/07 $20.00 © 2007 IEEE

and β_{i2} remains fixed at 0.1. These thresholds serve as indicators of possible paths that can be taken by a patient given the person specific random effect and give a way of predicting the success of interventions in the states for any new patient.

Observing state 3 (Figure 5), $\beta_{i2} < \beta_{i1}$ which may be interpreted that path 1 may not be needed in the system since patients who are not dead and have patient specific random effect above β_{i1} could be discharged from state 3 to recuperate in their own home. This is an indication that this modelling technique can detect where the system may be going wrong.

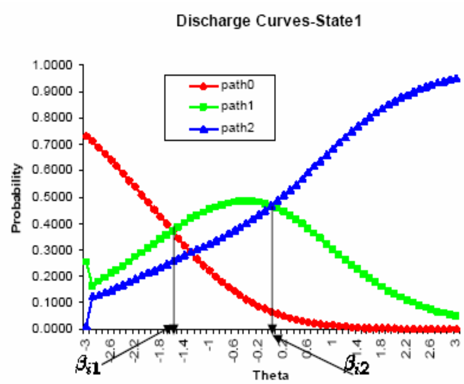

Figure 3: Discharge curves through the three paths showing thresholds for State 1

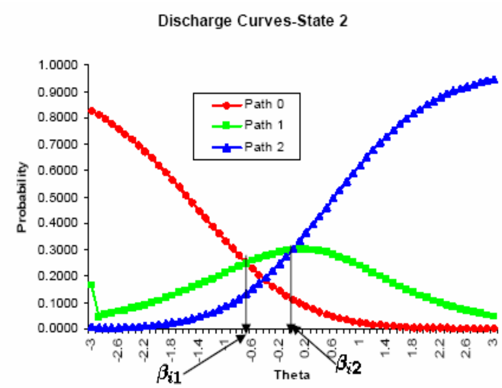

Figure 4: Discharge curves through the three paths showing thresholds for State 1

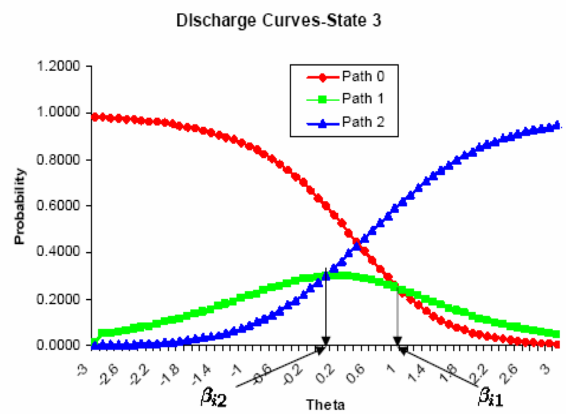

Figure 5: Discharge curves through the three paths showing thresholds for State 1

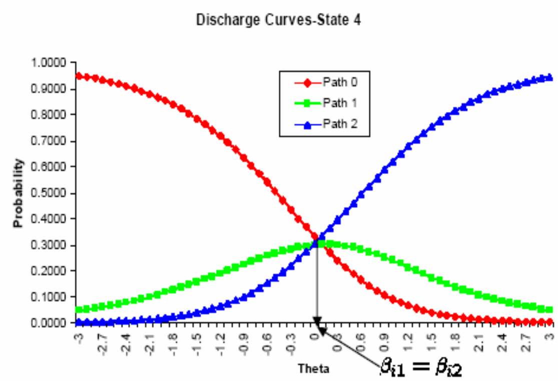

Figure 6: Discharge curves through the three paths showing thresholds for State 1

For state 4 (Figure 6) $\beta_{i1} = \beta_{i2}$ =0, this is a point at which it is equally likely for any patient to pass through any of the paths. Points less than $\beta_{i1} = \beta_{i2}$ =0 are more likely to pass through path0 and path2 is more likely above the threshold. Another interpretation is that patients below the population average are more likely to pass through path 0 while patients whose health status is greater than the population average are more likely to pass though the path 2. Discharge probabilities from the system through any combination of paths can easily be calculated.

4. Conclusion

Previous work have independently model LOS without capturing the flow paths of patients through the system. However, in this paper we have demonstrated the fact that mixed models are useful in health and social care modelling since patient experience in the care process is highly important for improvement of the process of care and better planning. We have presented a model for the prediction of discharge probabilities based on individual patient's flow paths (trajectories) and patient-specific random effects. However, an elegant way will be to jointly model LOS and patients flow paths. Therefore, an extension of this work to the joint modelling of LOS and individual patient's flow paths is currently being investigated.

5. References

[1] E. Kulinskaya, D. Kornbrot, and H. Gao, "Length of stay as a performance indicator: robust statistical methodology," *IMA J Management Math*, vol. 16, pp. 369-381, 2005.

[2] V. Irvine, S. Mcclean, and P. Millard, "Stochastic-Models for Geriatric Inpatient Behavior," *IMA Journal of Mathematics Applied in Medicine and Biology*, vol. 11, pp. 207-216, 1994.

[3] S. Mcclean and P. Millard, "Patterns of Length of Stay after Admission in Geriatric-Medicine - an Event History Approach," *Statistician*, vol. 42, pp. 263-274, 1993.

[4] H. Xie, T. J. Chaussalet, and R. H. Millard, "A continuous time Markov model for the length of stay of elderly people in institutional long-term care," *Journal of the Royal Statistical Society Series a-Statistics in Society*, vol. 168, pp. 51-61, 2005.

[5] M. J. Faddy and S. I. McClean, "Analysing data on lengths of stay of hospital patients using phase-type distributions," *Applied Stochastic Models in Business and Industry*, vol. 15, pp. 311-317, 1999.

[6] G. W. Harrison and P. H. Millard, "Balancing Acute and Long-Term Care - the Mathematics of Throughput in Departments of Geriatric-Medicine," *Methods of Information in Medicine*, vol. 30, pp. 221-228, 1991.

[7] E. El-Darzi, C. Vasilakis, T. Chaussalet, and P. H. Millard, "A simulation modelling approach to evaluating length of stay, occupancy, emptiness and bed blocking in a hospital geriatric department," *Health Care Management Science*, vol. 1, pp. 143-149, 1998.

[8] M. J. Cote and W. E. Stein, "A stochastic model for a visit to the doctor's office," *Mathematical and Computer Modelling*, vol. 45, pp. 309-323, 2007.

[9] M. J. Cote and W. E. Stein, "An Erlang-based stochastic model for patient flow," *Omega-International Journal of Management Science*, vol. 28, pp. 347-359, 2000.

[10] N. M. Laird and J. H. Ware, "Random-Effects Models for Longitudinal Data," *Biometrics*, vol. 38, pp. 963-974, 1982.

[11] G. N. Masters, "A Rasch Model for Partial Credit Scoring," *Psychometrika*, vol. 47, pp. 149-174, 1982.

0-7695-2905-4/07 $20.00