AUT UNIVERSITY
Faculty of Business and Law

Department of Economics

Working Paper Series

A Model for Predicting Readmission Risk in New Zealand

Rhema Vaithianathan, Nan Jiang and Toni Ashton

2012/02

A Model for Predicting Readmission Risk in New Zealand

Rhema Vaithianathan, Nan Jiang and Toni Ashton

Abstract

Predictive Risk Models which utilize routinely collected data to develop algorithms are used in England to stratify patients according to their hospital admission risk. An individual's risk score can be used as a basis to select patients for hospital avoidance programmes. This paper presents a brief empirical analysis of New Zealand hospital data to create a prediction algorithm and illustrates how a hospital avoidance business case can be developed using the model. A sample of 134,262 patients was analyzed in a Multivariate logistic regression, various socioeconomic factors and indictors of previous admissions were used to predict the probability that a patient is readmitted to hospital within the 12 months following discharge. The key factors for readmission prediction were age, sex, diagnosis of last admission, length of stay and cost-weight of previous admission. The prognostic strength of the algorithm was good, with a randomly selected patient with a future re-admission being 71.2% more likely to receive a higher risk score than one who will not have a future admission.

Keywords: Hospital readmission; Risk prediction; Prognostic strength

JEL Classification: I10; C13; O22

Corresponding authors: Rhema Vaithianathan: Department of Economics, The University of Auckland, New Zealand. Email: r.vaithianathan@auckland.ac.nz**;** Nan Jiang: Department of Economics, Auckland University of Technology, New Zealand. Email: nan.jiang@aut.ac.nz; Toni Ashton: School of Population Health, The University of Auckland, New Zealand. Email: t.ashton@auckland.ac.nz

1. Introduction

In most health care systems, a small fraction of the population accounts for the bulk of health care usage costs (Reuben, *et al.* 2002; Hughes, *et al.* 2004; Department of Health, 2004; Pannatoni, *et al.* 2011). With respect to New Zealand, in the Counties Manukau District Health Board (CMDHB), less than one-fifth of the population account for almost half of all adult medical bed days. The highly skewed nature of utilisation and health care expenditure implies that the cost-effectiveness of "upstream interventions", such as hospital readmission avoidance programmes, will be improved if they target patients who are at sufficiently high risk of hospitalisation and identify these patients with sufficient lead time for the intervention to have an impact (Roseman, 2003).

In response to the challenge of identifying these high risk patients, Billings *et al.* (2006) developed the 'patient at risk for rehospitalisation' (PARR) predictive risk modelling (PRM) tool. PARR was primarily designed for use by primary care trusts or community services to identify patients at high risk of readmissions to hospital in the next 12 month. Data from hospital episode statistics in National Health Service (NHS) trusts in England was used. A second tool, the Combined Predictive Risk Model, was subsequently developed (Pannatoni, *et al.* 2011), combining a comprehensive dataset of patient information, including inpatient, outpatient, and accident & emergency data from secondary care sources as well as general practice electronic medical records. These tools are reported to be used by 72% of National Health Service (NHS) agencies in England responsible for managing chronic care in the community (Ham, 2009).

Pannatoni *et al.* (2011) argue that a similar tool is feasible for New Zealand given that details of all public hospital admissions are routinely collected in the National Minimum Data Set (NMDS). In this paper we develop a case finding algorithm for the Waitemata District Health Board (WDHB) to enable it to risk stratify its patient

population, using hospital episode statistics for the three years from 01 July 2006 to 30 June 2009.

The WDHB has the largest catchment of the 20 DHBs in New Zealand with more than 525,000 residents. It is located in Auckland, the largest city in New Zealand, and therefore principally serves an urban population. The pressures placed on health systems in both Australia and New Zealand has lead to increased interest in hospital avoidance programmes. Predictive Risk Models are a practical way to increase the efficacy of such programmes.

The rest of this short paper is organized as follows: Section 2 explains the data and methodology, Section 3 discusses the results, Section 4 illustrates how to develop a hospital avoidance business case using the model and Section 5 concludes the main findings.

2. Method

The methodology used in this study follows that developed by Billings *et al.* (2006) for the PARR case finding algorithm in England. Multivariate logistic regression analysis is conducted in which the predicted variable was the probability that a patient is readmitted to hospital in the 12 months following the date of discharge. The model was estimated on a random 50% sample and validated on the remaining 50% of the population. The potential net costs of different levels of spending on hospital readmission avoidance programmes were estimated for given levels of efficacy and risk thresholds.

We used a subset of WDHB hospital episode data of all adult acute admissions (i.e. unplanned) over 3 years from 01 July 2006 to 30 June 2009. Only patients who were admitted between 01 January 2008 and 30 June 2008 were selected. For each patient,

a "triggering" admission date (TAD) was identified. If patients had more than one admission during the period, the earliest admission was considered as his/her TAD.

Patients who died in the 12 months following the TAD or who were less than 17 years old at the TAD were excluded from the analysis. Patients who had more than one TAD on the same day were dropped due to our inability to ascertain the reason for this multiplicity. For each patient in the sample, we identified whether the patient had a readmission during 01 July 2008 and 30 June 2009 (see Figure 1).

For each patient, data on the patient's previous acute hospital admissions back to 01 July 2006 were coded to determine the number of acute admissions in the previous 90, 180 and 365 days, the total number of previous acute admissions and whether or not this patient was previously admitted for a reference condition as defined by Billings *et al.* (2006), such as congestive heart disease, chronic obstructive pulmonary disease, diabetes, sickle cell disease etc. for which improved management may help to prevent future admissions. The final sample size was 134,262 patients.

<Insert Figure 1 about here>

We constructed disease categories using Diagnosis Related Group (DRG) codes. DRG codes are diagnostic fields in computerized hospital admission. DRG is a system to classify hospital cases into specific group expected to have similar hospital resource use. Individual characteristics of sex, age and ethnicity are recorded within the NMDS. A step-wise multivariate statistical analysis was undertaken in order to develop an algorithm to predict patients at high risk of readmission in the 12 months following the TAD. This algorithm was developed on half the sample (N=67,131). The coefficients for the 41 most powerful variables were chosen on the basis of maximizing the log-likelihood ratio. We then applied the model to the remaining 50% of the population (N=67,131) to validate the findings of the algorithm from the first sample. All analyses were conducted using the logit command and Stata 11 (64 bits). The basic model can be expressed as the following:

$$Y_i^* = \boldsymbol{\beta} \cdot \mathbf{X}_i + u_i \qquad (1)$$

Where Y_i^* is a latent variable associated with patient i, and Y_i is a dummy variable equals to 1 if patient i was actually readmitted in the 12 month following discharge and 0 otherwise. We cannot observe Y_i^*, what we observe is the dummy variable Y_i defined by:

$$Y_i = 1 \quad if \quad Y_i^* > 0$$
$$Y_i = 0 \quad if \quad Y_i^* \le 0$$

$\mathbf{X}_i$ is a vector of patient characteristics, such as age, sex, ethnicity, current hospital admission diagnosis, severity, and previous admissions status, etc. $\boldsymbol{\beta}$ is the corresponding vector of parameters to be estimated, and u_i is the random errors associated with this patient which is assumed to have a logistic distribution. The logistic distribution is a continuous probability distribution, its cumulative distribution function is the logistic function:

$$F(\boldsymbol{\beta} \cdot \mathbf{X}_i) = \frac{e^{\boldsymbol{\beta} \cdot \mathbf{X}_i}}{1 + e^{\boldsymbol{\beta} \cdot \mathbf{X}_i}} = \frac{1}{1 + e^{-\boldsymbol{\beta} \cdot \mathbf{X}_i}} \qquad (2)$$

Denote by P_i the probability that $Y_i^* > 0$:

$$P_i = prob(Y_i = 1) = prob(Y_i^* > 0) = prob(\boldsymbol{\beta} \cdot \mathbf{X}_i + u_i > 0) = prob(u_i > -(\boldsymbol{\beta} \cdot \mathbf{X}_i)) \qquad (3)$$

With the logistic function, we have the following logit model for P_i:

$$P_i = \frac{1}{1+\exp(-Y_i^*)} = \frac{1}{1+\exp(-(\boldsymbol{\beta}\cdot\mathbf{X}_i + u_i))} \quad (4)$$

$$1-P_i = \frac{\exp(-(\boldsymbol{\beta}\cdot\mathbf{X}_i + u_i))}{1+\exp(-(\boldsymbol{\beta}\cdot\mathbf{X}_i + u_i))} \quad (5)$$

From which it follows that:

$$\ln(\frac{P_i}{1-P_i}) = \boldsymbol{\beta}\cdot\mathbf{X}_i + u_i \quad (6)$$

Unlike Billings *et al.* (2006), we estimate the risk *for all readmissions* rather than only those in a subset of "reference conditions" to establish the prognostic strength for an entire range of hospital readmissions. Whether or not these hospital readmissions can be prevented through improved management or effective early intervention is another empirical question to be addressed.

The predictor variables were chosen as follows. Variables which were always insignificant no matter how many other variables were included in the logistic regression were dropped. The estimated beta weights were applied to the development sample to derive the prognostic characteristics of the algorithm.[1] The variables that were included in the model were functions of sex, age, ethnicity, DRG of admission, length of stay, number of admissions in past 6 months and cost-weights of previous admissions.

3. Results

Table 1 reports on the ability of the algorithm to act as a basis for selecting patients for an intervention for WDHB. The Positive Predictive Value (PPV) indicates the

[1] The beta weights for the model can be obtained from the author upon request.

percentage of patients who are flagged (i.e. exceed the risk-score threshold) and are subsequently readmitted. The one minus PPV (1-PPV) value, on the other hand, is critical in assessing the potential for the tool to increase the cost effectiveness of hospital readmission avoidance programmes. If this value is large, then the algorithm results in patients who are incorrectly identified and incorrectly recruited to the hospital avoidance programme. Therefore, the total savings from the initiative will be lower. This is because the potential savings derived from reducing subsequent admissions are unavailable for these patients who are incorrectly identified.

<Insert Table 1 about here>

Use of a variety of risk thresholds ascertains how sensitive the results are, in terms of finding patients who are potentially in need of intervention (i.e. have a subsequent readmission). At a risk score threshold of 70, the algorithm identifies 2,403 patients, of whom 73.37% would have been correctly identified. This is closely comparable with 77% in the UK Parr Tool (Billings, *et al.* 2006). The remainder (26.63%) would have been flagged incorrectly. A risk score threshold of 90 identified only 526 patients, of whom 83.46% were actually readmiited in the next 12 month following discharge. There is a clear trade-off between achieving high targeting accuracy rate and selecting the number of patients to be targeted. The higher the risk score threshold, the fewer number of patients are selected, but a higher percentage of them are those who will actually be readmitted.

The receiver operating characteristic (ROC) curve graphically illustrates the trade-offs between sensitivity and 1 minus specificity (See Figure 2). The area under the curve indicates a 71.2% probability that a randomly selected patient with a future re-admission will receive a higher risk score than a randomly selected patient who will not have a future admission. This compares favourably with the ROC curve of the UK Parr Tool which had a lower area under the ROC curve (68.5%) which suggests that the latter has a slightly worse prognostic strength – although we cannot test whether this difference is statistically significant (Delong, *et al.* 1988).

<Insert Figure 2 about here>

4. Development of a Business Case

To judge the usefulness of the model, we use it to build a business case for hospital readmission avoidance programmes (See Table 2). We assume three levels of intervention costs as in Billings *et al.* (2006): $500, $750 and $1,000 per patient. We also assume that the intervention is effective in reducing readmission rates by 10% amongst those patients who did have a readmission. We calculate the savings from avoiding admissions on the basis of the average cost-weight per admission in our sample of discharges (using the average cost-weight multiplied by the national reference price of $4,410).

The business case estimates the net costs or savings from an intervention with a target level of efficacy and risk threshold. For example, an intervention which costs $500 per patient and reduces hospital admission by 10% yields an expected net saving of $444,189 if targeted at the 1% highest risk group, and $1,691,969 if targeted at the 20% highest risk group. Those figures will be reduced to $396,189 and $1,057,969 respectively if the intervention cost doubled to $1000 per patient. Higher risk score threshold therefore does not necessarily imply more savings from early intervention. As mentioned before, it will identify fewer patients to be treated. The design of an effective intervention program cannot be isolated from budget consideration, the prognostic characteristics of the predictive risk model is also an important factor.

<Insert Table 2 about here>

5. Discussion and Conclusion

This paper presents a model for predicting readmission risk within 12 months following an admission to hospital and investigates whether such a model can be used to help establish a business case for hospital readmission avoidance interventions.

The PRM we estimated achieved reasonable prognostic strength using routinely collected data for hospital admissions. Greater accuracy may be achieved if a more comprehensive data set was to be used: for example, by including outpatient data, general practice consultations and pharmacy data. However the attraction of the model presented here is that it does not require the linkage of different data sets and can therefore be run fairly readily by analysts.

Our business model suggests that, when linked with an effective intervention, the use of the PRM has the potential to make substantial savings through avoided readmissions. Further research is required to compare the prognostic strength of a PRM approach with standard clinical judgment, and to examine the impact that using a PRM to identify at-risk patients can have on the cost-effectiveness of interventions to prevent readmissions.

An alternative approach to using a PRM such as this is to utilize a threshold model wherein patients are recruited into a hospital avoidance programme on the basis of meeting a certain number of fixed criteria. The advantage of threshold models is that they are simple to use. A disadvantage is that threshold models do not allow providers to calibrate the number of patients who are flagged (Pannatoni, *et al.* 2011). In contrast, with a PRM, one may readily identify the top 1%, 0.1% or 0.01% of risk groups. Threshold models are also thought to suffer more starkly from the problems of regression to the mean in the sense that they are more inclined to identify patients that have been at high risk of readmission in the past rather than being at high risk in the future.

To explore the differences between the PRM and a threshold model, we compared this to the case finding tool used by the Frequently Admitted Medical Admissions (FAMA) programme, an intensive case management programme designed to reduce hospitalization in Counties Manukau District Health Board (Roseman, 2003). FAMA applied criteria similar to other chronic care management programmes (such as Evercare in the UK) and flagged patients on the basis of 2 or more previous admissions in last 12 months for a total of 5 or more bed days. Patients who are no more than 15 years old and/or dead during their triggering admission are excluded. When applied to our sample, the threshold flagged 15,629 patients (11.64% of the 134,262 individuals) in our sample. Within the flagged group, 9,071 (58%) were actually readmitted in the future period.

To directly compare this method with a PRM approach we calculated the risk score that would be required to flag 15,629 patients. This yields a score of 0.49595. Of this alternative 15,629 patients flagged by PRM, 9,792 (63%) of them were readmitted in the 12 months after the triggering period - which is 721 more than were flagged by FAMA.

It is also important to ask whether the PRM is a better at case-finding than clinical judgment. The argument that doctors might be better able to judge the risk of admission of a patient into hospital than a statistical algorithm requires further analysis.

Our business case was not based on any particular hospital avoidance programme. In our example, there are net savings at all risk thresholds, even though we assumed conservatively that the intervention would prevent only 10% of readmissions. Further research is required to understand what factors contribute to readmission by the particular set of patients that are flagged by the PRM, what types of interventions are effective in reducing their readmissions, and how much these interventions cost.

Our business case was developed assuming that the avoided hospital admission would cost the same as an average admission. It may be argued that the average cost of a hospital admission is higher than the marginal cost (Roberts *et al.* 1999; Taheri *et al.* 2000). Therefore, hospital avoidance programmes which use the average cost of avoided hospital stay over-estimate the savings in the sense that, unless a hospital is able to shut a ward and reduce staff numbers, it is unlikely that avoidance programmes will result in real savings to the health system. This is a somewhat specious argument as it fails to recognize that at some scale of hospital avoidance, the marginal cost and the average cost are the same. Indeed, it could be argued that it is this type of flawed reasoning that continues to see extremely high levels of hospital admissions which could be avoided. We therefore would argue that pricing hospital stays at their average cost is a valid approach. The average cost of hospital admissions avoided was estimated to be around $4568 in NZ by Love and Gullery (2011), very close to the $4756 we used to develop the business case. Future research on the cost distribution of hospital admissions may shed more light on this issue.

References

Billings, J., Dixon, J., Mijanovich, T. and Wennberg, D. (2006) "Case finding for patients at risk of readmission to hospital: development of algorithm to identify high risk patients*," British Medical Journal* **333**(7563), 327-333.

DeLong, E.R., DeLong, D.M., and Clarke-Pearson, D.L. (1988) "Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach," *Biometrics* **44**(3), 837-845.

Department of Health (2004) "*Improving Chronic Disease Management*," available from: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4075214

Fox-Rushby, J. and Cairns, J. (2005) "Understanding Public Health," *Economic Evaluation*. London: Open University Press.

Ham, C. (2009) "Chronic Care in the English National Health Service: Progress and Challenge," *Health Affairs* **28**(1), 190-201.

Hughes, J.S., Averill, R.F., Eisenhandler, J., Goldfield, N.I., Muldoon, J., Neff, J.M. and Gay, J.C. (2004) "Clinical Risk Groups (CRGs): A Classification System for Risk-Adjusted Capitation-Based Payment and Health Care Management," *Medical Care* **42**(1), 81-90.

Love, T. and Gullery, C. (2011). "*Pharmacist medicine management review services*," Pharmaceutical Society of New Zealand Inc., Available from: http://psnz.org.nz/members/members_home/documents/GulleryLove.pharmacistmedicinesmanagementreviewservices.Mar2011.pdf

Pannatoni, L.E., Vaithianathan, R., Ashton, T. and Lewis, G.H. (2011) "Predictive Risk Modelling in Health: Options for New Zealand and Australia," *Australian Health Review* **35**(1), 45-51.

Reuben, D.B., Keeler, E., Seeman, T.E., Sewall, A., Hirsch, S.H. and Guralnik, J.M. (2002) “Development of a Method to Identify Seniors at High Risk for High Hospital Utilization,” *Medical Care* **40**(9), 782-793.

Roberts, R.R., *et al.* (1999) “Distribution of variable vs fixed costs of hospital care,” *The Journal of the American Medical Association* **281**(7), 644-649.

Roseman, P. (2003) "*Frequent Adult Medical Admissions Final Report*," Counties Manukau Acute Demand Management Group, Available from: http://www.sah.co.nz/Funded-Services/CCM/docs/reports/FAMAreport.pdf

Taheri, P.A., Butz, D.A. and Greenfield, L.J. (2000) “Length of stay has minimal impact on the cost of hospital admission,” *Journal of the American College of Surgeons* **191**(2), 123-130.

Figure 1: Timing of the events

TABLE 1: PROGNOSTIC STRENGTH AT DIFFERENT RISK SCORE THRESHOLDS

	Risk score threshold			
	70	**80**	**90**	**99**
No. of patients flagged	2,403	1,268	526	96
Share of those flagged who are re-admitted (PPV (%))	73.37	78.08	83.46	91.67
Share of those flagged who are *not* re-admitted (1-PPV, %)	26.63	21.92	16.54	8.33
Share of re-admitted patients correctly flagged (Sensitivity (%))	8.75	4.91	2.18	0.44
Specificity (%)	98.64	99.41	99.81	99.98
Average number of re-admissions for correctly flagged patients	4.15	4.94	6.66	11.76

FIGURE 2: ROC CURVE FOR MODEL

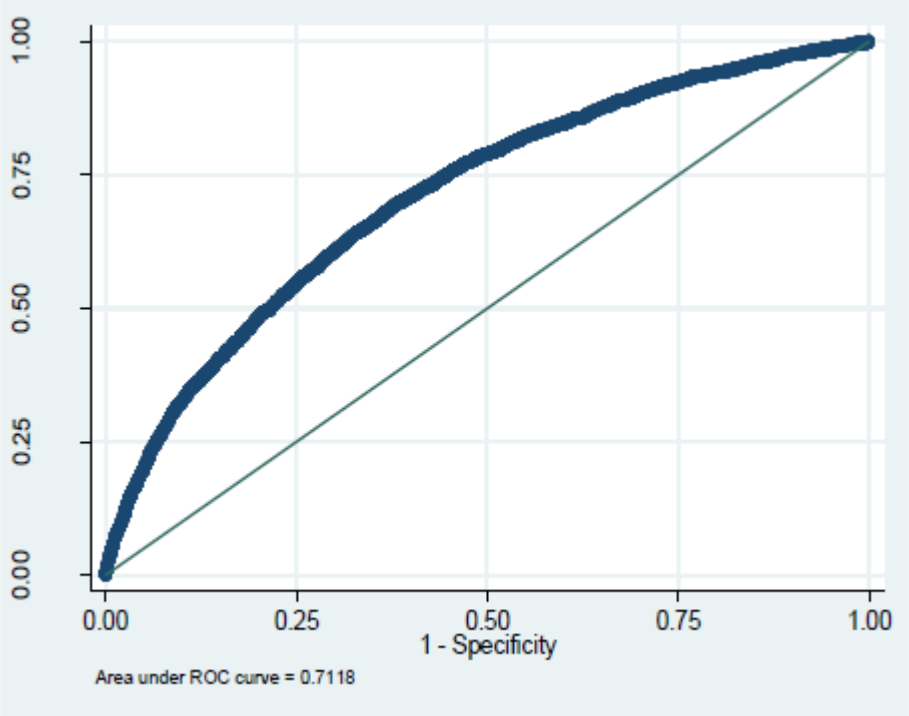

TABLE 2: BUSINESS CASE FOR HOSPITAL AVOIDANCE PROGRAMME WHICH IS 10% EFFECTIVE

Risk score threshold	No. admitted patients identified	No. patients flagged incorrectly (not admitted)	Total cost of intervention	Admissions within 12 months for correctly flagged patients	Intervention saving ($4756 per admission)	Net savings
Intervention cost of $500 per patient						
80	990	278	$634,000	4.94	$2,325,969	$1,691,969
90	439	87	$263,000	6.66	$1,390,531	$1,127,531
99	88	8	$48,000	11.76	$492,189	$444,189
Intervention cost of $750 per patient						
80	990	278	$951,000	4.94	$2,325,969	$1,374,969
90	439	87	$394,500	6.66	$1,390,531	$996,031
99	88	8	$72,000	11.76	$492,189	$420,189
Intervention cost of $1,000 per patient						
80	990	278	$1,268,000	4.94	$2,325,969	$1,057,969
90	439	87	$526,000	6.66	$1,390,531	$864,531
99	88	8	$96,000	11.76	$492,189	$396,189