Analysis of Patient Presentation and Admission Data for Prediction Modelling

J Boyle[1], J Lind[2], D Green[2], J Crilly[2], P Miller[3], M Wallis[2,4], M Jessup[4], G Fitzgerald[5]

[1]Australian E-Health Research Centre, CSIRO ICT Centre,
[2]Gold Coast Hospital, Queensland Health, [3]Toowoomba Hospital, Queensland Health
[4]Griffith University,
[5]Queensland University of Technology

Abstract

We describe data analysis undertaken as part of a patient admissions prediction project underway through Emergency Departments (EDs) at two hospitals in the Southern Area Health Service within Queensland Health. The project involves the modelling and software development of forecasting tools that predict ED admission for time and day of the year. These tools may assist with the allocation of inpatient beds to aide in alleviating two major problems of most EDs: overcrowding and access block.[1]

The modelling data consists of five years of ED presentations and admissions (1/7/02 – 30/6/07) from the Gold Coast and Toowoomba hospitals which were chosen for their different demographic characteristics. Toowoomba reflects an entire population (~90,000) served by one ED with a fairly stable population, unlike the Gold Coast, which has one of the busiest EDs in the state, a large itinerant population (~500,000) and numerous other EDs serving the area.

Many useful characteristics which can help shape health management practices have been identified from the data. For example, the date and time when admitted patients leave the ED, indicating the times of highest demand on hospital beds; patient arrival time in the ED, which represents a staffing impact with workload; and the days of the week which represent higher ED workloads and hospital bed demand. The data also enables the analysis of 'frequent-flyers' – those patients who presented multiple times during the analysis period.

From the analysis of this data, we have been generating forecast estimates and associated confidence intervals based on several forecasting approaches and validating the forecasts against actual data. The project also includes packaging the most accurate technique into a standalone software application.

Topic area and paper objectives:

Forecasting is an important aid in effective and efficient planning. The ability to forecast patient admissions can assist many areas of health management, including elective surgery scheduling, bed management, and staff resourcing. The analysis of patient presentation and admission data is an important step in determining candidate models for forecasting expected values of this data. The work forms part of a 12-month project assessing quantitative forecasting techniques applied to admissions data.

Background:

There are many examples of forecasting in health applications[2-4], and our work aims to address gaps in forecasting ED admissions. For example, recent related work by researchers from La Trobe University and Bendigo Health[3] compared the number of ED presentations for five months of 2006 with predictions from 2000-2005 data. The authors report that forecasts made with an autoregressive integrated moving average method "compared well" with observed attendance data. The study raised additional questions which our work addresses:

- Are time series for different triage categories or International Statistical Classification of Diseases (ICD) groups similar?

We are including triage and ICD groupings in our forecasting.

- Can admissions be predicted as distinct from presentations?
We are providing forecasts for both admissions and presentations.
- Can ED presentations be examined on a finer scale (daily or hourly)?
We are performing analysis in hour time blocks.
- Would accurate prediction facilitate health service and staff planning?

We are assessing impact of a forecasting model on ED staff and ED work practices.

Methods:

Following ethics and State Government Privacy legislation approvals, we obtained five years of ED presentation and admission data (1/7/02 – 30/6/07) for the Gold Coast and Toowoomba hospitals. Two databases were provided:
1. EDIS (Emergency Department Information System) - data is specific to the ED and does not capture what happens after the patients leave the ED;
2. HBCIS (Hospital Based Clinical Information System) - HBCIS captures data for patients that are admitted only (diagnosis, procedures performed, length of stay etc) but does not indicate presentations and hence the total workload seen by the ED.

One of the major challenges of the project to date has been matching these two separate information systems. Both databases have useful information and the results that we can obtain from this project are increased by the ability to integrate both datasets.

Presentations numbers for both hospitals across the five years are similar (278,000 - Gold Coast, 218,000 - Toowoomba) although the Gold Coast hospital has a higher rate of admissions (33%) than Toowoomba hospital (20%) – see Figure 1. Most patients presenting at the Gold Coast were Australasian Triage Scale (ATS) category 3 (53%), whereas most presentations at Toowoomba were ATS category 4 (59%).
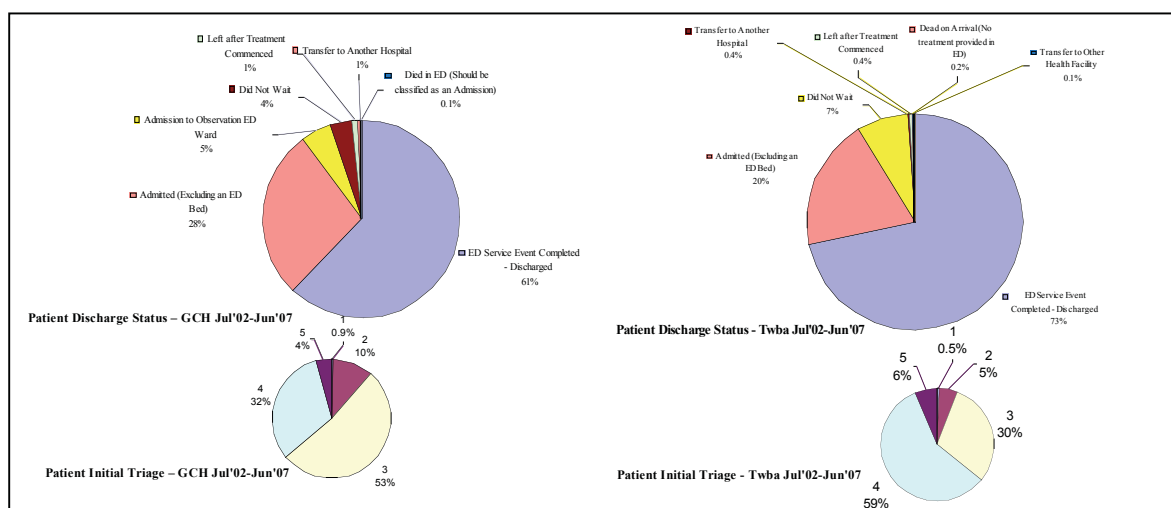

Figure 1. *Discharge Status and Triage of Presenting Patients; Left: Gold Coast Hospital, Right: Toowoomba Hospital*

The data also enables the analysis of the so-called 'frequent-flyers' – those patients who presented multiple times during the analysis period. Useful insights into health management practices can be obtained by analysis of these cases. This group is diverse, and includes mental health patients, people over 70, acute patients, and those with chronic disease.

For example, the 278,000 presentations made to the Gold Coast hospital ED comprise 156,000 individuals. Most people are seen once only, but there are many cases of multiple presentations, with the maximum presentations by a single individual being 256 times over the five years. Interestingly, the mean triage is constant across 'frequent flyers' for presentations (Figure 2). Two-thirds of all presentations are first timers, and the mean triage score is the same for the group of first-timers as it is for those on their tenth visit. As could be expected, the mean triage for admissions is less (indicating higher acuity) than that for presentations.
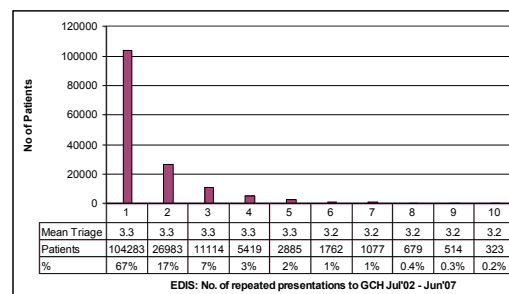

	1	2	3	4	5	6	7	8	9	10
Mean Triage	3.3	3.3	3.3	3.3	3.3	3.2	3.2	3.2	3.2	3.2
Patients	104283	26983	11114	5419	2885	1762	1077	679	514	323
%	67%	17%	7%	3%	2%	1%	1%	0.4%	0.3%	0.2%

EDIS: No. of repeated presentations to GCH Jul'02 - Jun'07

Figure 2. *Gold Coast Hospital Frequent Flyers*
(Note: above shows horizontal axis up to 10 repeated presentations – data extends to 256)

Data Analysis:

The data includes date and time of admissions which provides useful information on peak admission times experienced within the EDs. Figure 3 indicates the times of highest demand on hospital beds (admitted patients leaving the ED), indicated by a brighter colour. The vertical columns of the plots indicate the hour
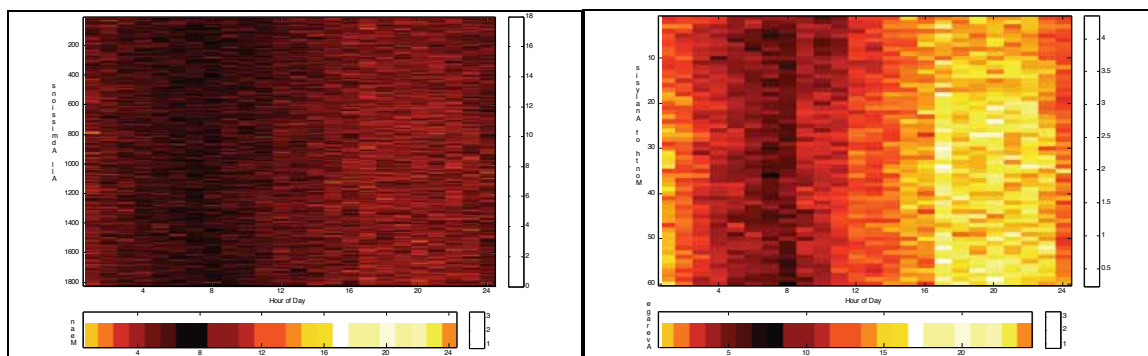

Figure 3. *Peak admission times for patients leaving the ED;*
Left: Admissions across every day in study period; Right: Monthly averages across study period.

of day, and admission numbers are indicated by the colour bar. It is apparent that the highest demand for hospital beds occurs in the afternoon and into the evening. Every row on the left-hand plot represents a day from 1/7/02 – 30/6/07, while the rows on the right-hand plot indicate monthly averages throughout the study period. Similar assessment has been done for discharge times of all presentations (not just those admitted) and also for arrival times within the ED. The skew of the data to the end of the day is apparent.

Another point of interest is the time of arrival in the ED, as this represents a staffing impact with workload. Figure 4 shows the ED discharge time for the admissions that are shown in Figure 3. It also indicates the arrival
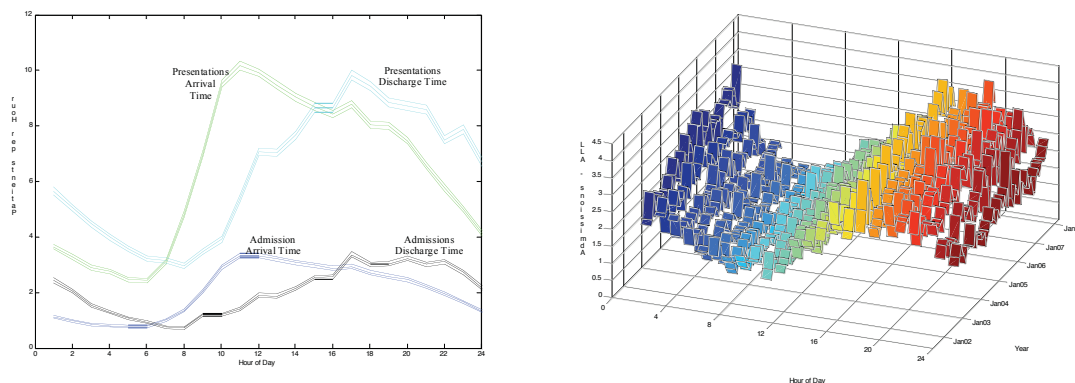
Figure 4. *Patient Arrival/Discharge in the ED*

time for this group, which peaks around 11:00hrs. However, admitted patients make up only a small subset of all the patients seen in the ED, and the two curves in the upper portion of the plot represent all presentations. Again we can see that the peak discharge time lags behind the peak arrival time by around 8 hours. The plot
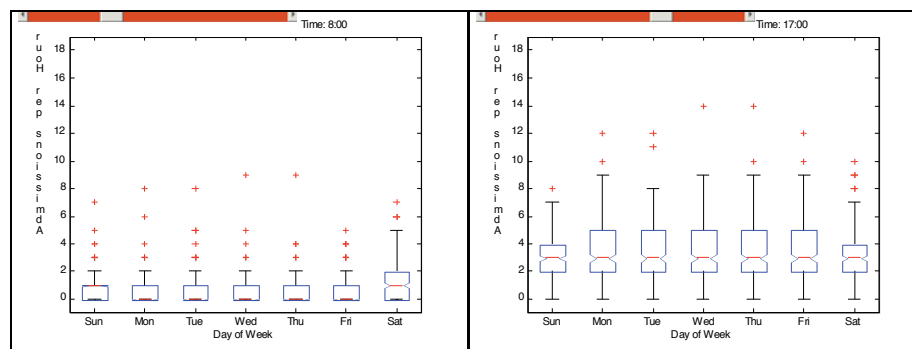

Figure 5. *Box-plots of hourly admissions; Left: 08:00hrs, Right: 17:00hrs*

shows mean patient numbers per hour along with the 95% Confidence Interval band around the mean. It can be seen that the Confidence Interval is very tight, indicating this trend is common across the dataset.

The hourly fluctuations of the data has also been studied using box-plots as shown in Figure 5, which show, for example, the quietest time (8am) and the busiest time (5pm) for admissions. Median, upper and lower quartiles and outliers are represented in the plots.

It is also of interest to determine the days of the week that represent higher ED workloads and hospital bed demands. For example, Figure 6 shows the mean and 95% Confidence Interval band for the daily and monthly trends in the arrival time of all presentations (Left) and for admitted patients (Right) at the Gold Coast
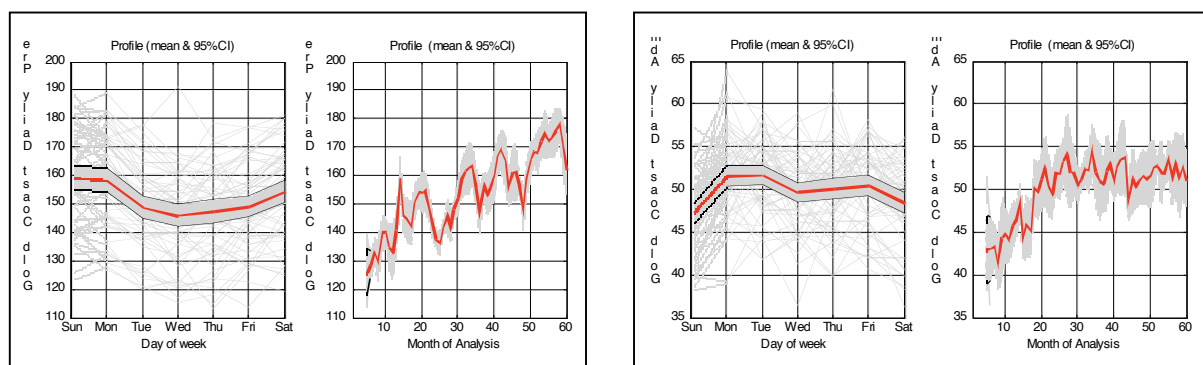

Figure 6. *Daily and Monthly Trends in Presentations (Left) and Admissions (Right) at the Gold Coast Hospital ED – July 2002 – June 2007.*

paper peer review E1 DEST 2008   4

Hospital ED. The busiest days for presentations are over the weekend and Mondays. Considering the arrival time for just those patients that are admitted, it can be seen that Mondays and Tuesdays are the busiest days. There has also been an overall increase (approximately 40%) in the number of patients presenting over the five years. Interestingly the trend over all the months-of-analysis for admitted patients shows a plateau effect, which could be attributed to bed capacity being reached, or the adoption of hospital avoidance strategies.
Implications:
From the analysis of this data, we have been generating forecast estimates and associated confidence intervals based on several forecasting techniques. This modelling is ongoing and includes probabilistic modelling, stepwise multiple regression, exponential smoothing and Box-Jenkins Autoregressive Integrated Moving Average (ARIMA) models. The current approach allows a user to generate a forecast for either a Day-of-week and Month-of-interest (for example, estimates for Mondays in January at 15:00hrs) or a specific Day-of-
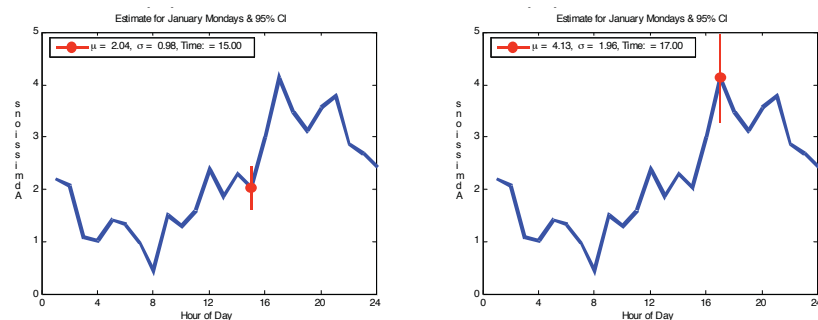

Figure 7. *Admission forecast for Mondays in January; Left: 15:00hrs estimate; Right: 17:00 hrs estimate*

interest (eg. 31 January, 10:00hrs). For example, Figure 7 shows probability estimates for Mondays in January based on the data. The red dot and bar depicts the mean and 95% Confidence Interval for the desired time of day. Two estimates are shown: one for 15:00hrs (left) and 17:00hrs (right). We can see that at 17:00hrs we
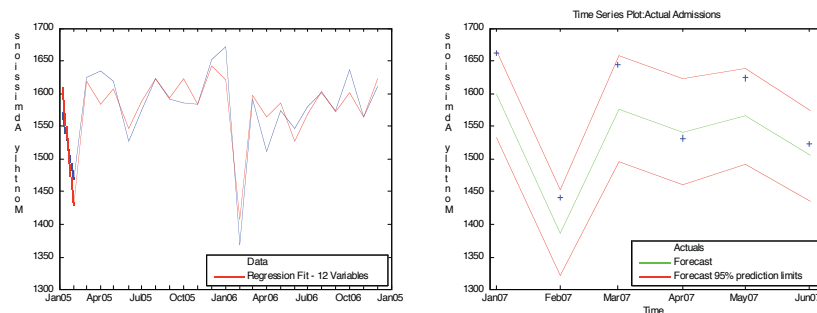

Figure 8. *Left: Regression fit to monthly admissions data Jan'05'Dec'06;
Right: Comparison of forecast predictions using model Jan'07-Jun'07.*

can expect a higher rate of admissions but less certainty in the prediction (higher standard deviation). The mean (μ) and standard deviation (σ) are overlaid on each plot.
Other current work includes validating the predictions using statistical comparison with actual data (Figure 8), and qualitatively assessing the impact of a prediction tool via semi-structured interviews with ED and bed management staff.[5]

Once this additional work is complete it is anticipated that a tool that can predict ED admissions and thus allow appropriate allocation of in-patient beds and operating theatres can be produced. With regular feed of site specific retrospective data, this tool should have considerable utility for acute facility bed management and health service planning.

References:

1. P Cameron, Hospital overcrowding: a threat to patient safety? Med J Aust 184(5):203-4
2. R Abdel-Aal, A Mangoud, Modeling and forecasting monthly patient volume at a primary health care clinic using univariate time-series analysis, Computer Methods and Programs in Biomedicine 1998; 56(3): 235-247
3. R Champion, L Kinsman, G Lee, et al., Forecasting emergency department presentations, Aust Health Rev 2007; 31(1): 83-90
4. A Pereira, Performance of time-series methods in forecasting the demand for red blood cell transfusion, Transfusion 2004; 44(5):739-46
5. N Harrop, T Wood Harper, A Gillies, Neglected user perspectives in the design of an online hospital bed-state system: implications for the National Programme for IT in the NHS, Health Informatics Journal 2006,12(4):293-303