ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/24041307

Short-Term Forecasting of Emergency Inpatient Flow

Article *in* IEEE transactions on information technology in biomedicine: a publication of the IEEE Engineering in Medicine and Biology Society · March 2009

DOI: 10.1109/TITB.2009.2014565 · Source: PubMed

CITATIONS
94

READS
1,111

3 authors, including:

Graham Byrnes
International Agency for Research on Cancer
254 PUBLICATIONS **19,036** CITATIONS

SEE PROFILE

Christopher Bain
Monash University (Australia)
51 PUBLICATIONS **486** CITATIONS

SEE PROFILE

All content following this page was uploaded by Christopher Bain on 02 May 2014.

The user has requested enhancement of the downloaded file.

Short-Term Forecasting of Emergency Inpatient Flow

Gad Abraham, Graham B. Byrnes, and Christopher A. Bain

***Abstract*—Hospital managers have to manage resources effectively, while maintaining a high quality of care. For hospitals where admissions from the emergency department to the wards represent a large proportion of admissions, the ability to forecast these admissions and the resultant ward occupancy is especially useful for resource-planning purposes. Since emergency admissions often compete with planned elective admissions, modelling emergency demand may result in improved elective planning as well. We compare several models for forecasting daily emergency inpatient admissions and occupancy. The models are applied to three years of daily data. By measuring their mean square-error in a cross-validation framework, we find that emergency admissions are largely random and hence non-predictable, whereas emergency occupancy can be forecasted using a model combining regression and ARIMA, or a Seasonal ARIMA model, for up to one week ahead. Faced with variable admissions and occupancy, hospitals must prepare a reserve capacity of beds and staff. Our approach allows estimation of the required reserve capacity.**

***Index Terms*—Patient flow, emergency occupancy, health care management, time series analysis, forecasting, ARIMA, SARIMA**

I. Introduction

INCREASING attention has been directed lately to optimal management of health care resources, i.e., providing better and more efficient patient care while minimising resource use [1], [2]. Public hospitals in Australia operate under severe and conflicting constraints. They are medically and legally obliged to accept any patient arriving at the emergency department (ED), and to provide them the best possible care. The point at which this ability is exhausted, due to lack of ED beds or of other resources, is called "hospital bypass" or "ambulance diversion." Hospital performance is scrutinised both by the government [3], and by the public, through the media. Hospitals also operate under strict budgetary limits, hence they need to allocate resources in a cost effective way. Our investigations suggest that ED resources are typically exhausted due to back-propagation of congestion within hospital wards, i.e., insufficient ward resources to admit patients from the ED (*access block* [4]). This situation arises from the manner in which the emergency and elective patient streams interact: elective patients are booked days or weeks in advance, whereas emergency patients arrive in an unplanned fashion, beyond the control of the hospital, and driven by factors such as illness patterns, time of day, and sociodemographic effects. If electives are booked assuming a certain level of emergency demand, and it turns out to be an underestimate, then the two streams conflict. In such a case, some demand is reduced by cancelling planned elective admissions (even on the day of surgery). However, previously admitted elective patients will consume beds for their entire stay, thereby reducing the hospital's ability to admit emergency patients to the overnight wards for several more days.

Decisions regarding acceptable rates of elective admissions are made with lead times ranging from several days to several weeks ahead, due to the need to schedule patients and facilities such as surgical suites. Tactical decisions regarding other resources require lead times varying from a few days for staffing rosters to much longer term for construction of new facilities or training new staff. In the experience of the third author (CB), decisions are currently based on personal experience of staff, and the planning process is done manually. This makes it inherently difficult to evaluate the long-term consequences of decisions taken, and to compare the effects of several possible plans.

These constraints highlight the need for a modelling approach that can give decision makers a reliable estimate of future patient levels, based on current known levels. However, there are fundamental limitations to what such models can achieve: purely random processes cannot be predicted. In this work, we sought to determine where these limits might lie by evaluating several time-series analysis models.

Specifically, this work examines the issue of forecasting the number of daily admissions from a hospital's emergency department (ED) to its wards, and the resultant daily ward occupancy levels. This work has several goals. First, we explore whether emergency admissions and occupancy can actually be forecast. If the time series is random (*white noise*), then it cannot be forecast. On the other hand, if we can detect some temporal patterns, such as autocorrelations or seasonality, then we may be able to reduce the level of uncertainty regarding

Manuscript received August 24, 2007; revised August 13, 2008; accepted January 21, 2009. This work was supported by the Australian Research Council through Linkage Grant LP0349153. G. B. Byrnes is supported by an NHMRC Capacity Building Grant 251533.

G. Abraham was with the Department of Mathematics and Statistics, University of Melbourne, Parkville 3010, Victoria, Australia, and with Melbourne Health, Royal Melbourne Hospital, Parkville 3052, Victoria, Australia. He is now with the Department of Computer Science and Software Engineering, The University of Melbourne, Parkville 3010, Victoria, and with the NICTA Victoria Research Lab. (e-mail: gabraham@csse.unimelb.edu.au)

G. B. Byrnes was with the Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, University of Melbourne, Parkville 3010, Victoria, Australia. He is now with the International Agency for Research on Cancer, 150 Cours Albert Thomas, 69372 Lyon CEDEX 08, France. (e-mail: byrnesg@iarc.fr)

C. A. Bain was with the Redesigning Care and Patient Access Unit at the Royal Melbourne Hospital, Parkville 3052, Victoria, Australia. He is now with the Western & Central Melbourne Integrated Cancer Service, East Melbourne 3001, Victoria, Australia. (e-mail: christopher.bain@petermac.org)

Copyright (c) 2009 IEEE. Personal use of this material is permitted.

future hospital admissions and occupancy. Therefore, we evaluate several forecasting models of admissions and occupancy and compare their forecast accuracy. We use a cross-validation methodology to estimate how well these models generalise to time periods outside the ones used to train them. Finally, we evaluate the best model's performance in terms of its precision (95% prediction intervals). From a practical point of view, this is the most important measure since it quantifies our confidence in future forecasts, thereby indicating the level of reserve capacity required for accommodating unpredicted demand. We conclude by outlining the practical implications of these findings for hospital decision-making.

Other approaches, such as Markov models [5], [6] and queueing models [7], have also been applied to patient flow. We considered both a structured model in which we attempted to predict occupancy from arrival and discharge processes, and a black-box approach where we modelled occupancy directly.

Time series analysis has been applied before to health care and particularly to patient flow. Lin [8] used ARIMA, Vector Autoregressive Moving Average (VARMA), and Holt-Winters exponential-smoothing models to forecast the monthly discharges and differences in occupancy, across several hospitals. He found that in most cases, ARIMA performed as well as or better than either VARMA or the Holt-Winters method. Jones *et al.* [9] described several ARIMA models of daily hospital occupancy resulting from emergency admissions. They found that whereas external covariates such as weather were significantly correlated with the number of admissions, they added little to the model's forecasting ability. Champion *et al.* [10] fitted ARIMA and seasonal exponential-smoothing models to monthly emergency presentations (patient arrivals to the emergency department), and found that the latter performed slightly better. Earnest *et al.* [11] used ARIMA models to forecast ward occupancy due to SARS infections, for up to three days ahead. Channouf *et al.* [12] applied ARIMA models to forecasting calls to emergency medical services. They report that for forecast horizons of up to about two weeks, the best performing model was a mixed-effects regression model with an autoregressive model of the residuals. Beyond that time, no model had any forecasting benefit.

II. Definitions of Admissions and Occupancy

Patients arriving at the ED are seen first by a triage nurse who ranks them by urgency, then later by a doctor who assesses and initiates treatment of their complaint. Some will then be discharged or transferred directly from the ED, a few seriously ill patients will die in the ED, and the rest will be *admitted* to a ward. In this work the focus was on *multiday* patients which spend at least one night in a ward, as opposed to *sameday* patients which arrive and leave on the same day. The number of whole days an admitted patient spends in a ward is their *length of stay* (LOS). Leaving the ward (separation) occurs when the patient is *discharged* or upon death.

III. Methods

A. Dataset

The Royal Melbourne Hospital (RMH) is a tertiary teaching hospital in Melbourne, Australia, with a capacity of about 350 acute overnight beds. At the RMH, admitted emergency patients represent about two-thirds of the total hospital overnight acute occupancy. Most of the remaining occupancy is taken by elective patients. This study used de-identified data from the elective surgery waiting list at the RMH, for the period 10/4/2004 to 25/3/2007 . The data was gathered in accordance with NHMRC guidelines for Quality Assurance [13], in consultation with Human Research Ethics Committee for RMH.

B. Forecasting Models

In this section we describe several models for forecasting daily levels of admissions and occupancy. The first two models serve as base cases against which the more sophisticated models are compared.

1) Moving-Average Forecasting: In the k-step moving-average model, denoted MA(k), the one day ahead forecast F_{t+1} is the average of the last k observations in a sequence Y consisting of T observations

$$F_{t+1} = \frac{1}{k} \sum_{i=t-k+1}^{t} Y_i \qquad 1 \leq k < T, \tag{1}$$

where k is the MA parameter. Setting $k = 1$ is equivalent to forecasting the same value for tomorrow as is observed today. Conversely, setting $k = t - 1$ is using the mean of most of the observations. Edge effects caused when forecasting at the beginning or at the end of the time series are overcome by temporarily using smaller values of k. Subsequent forecasts are identical to those for the one day ahead $F_{t+1} = F_{t+2} = F_{t+3} = \cdots$. The optimal MA model is found by enumerating values of k from 1 to 365 days.

Note that this model should not be confused with *moving average smoothers* or with MA models from the Box-Jenkins ARIMA methodology. In this work, MA(k) refers to the moving average forecast, whereas the Box-Jenkins MA process is denoted by ARIMA$(0, 0, q)$.

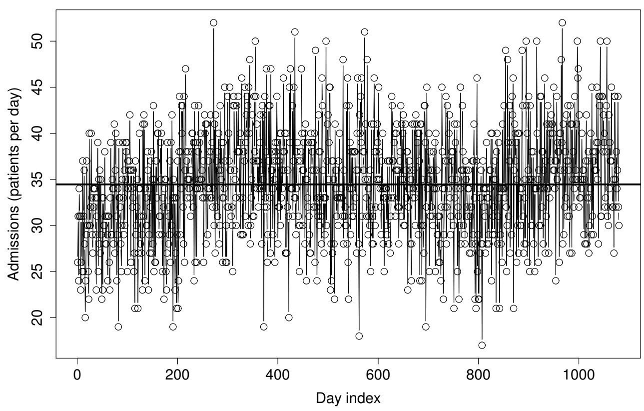
2) Single Exponential Smoothing: Single exponential smoothing (SES) is related to the moving average method, however, recent observations are weighted more highly than earlier ones. SES is expressed as

$$F_{t+1} = \begin{cases} Y_{t+1} & t = 0, \\ F_t + \alpha(Y_t - F_t) & t \geq 1. \end{cases} \tag{2}$$

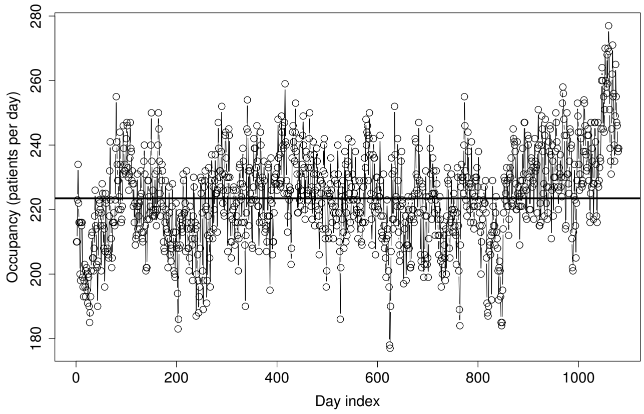
The optimal value of α was found using an optimisation method, using the criteria of minimum sum-of-squares.

Note that SES is not representative of all exponential-smoothing models. For example, the Holt-Winters method models seasonality explicitly (see, e.g., [14]). Preliminary results showed that the Holt-Winters method produced similar forecast accuracy to the ARIMA models discussed below, hence they were not pursued further.

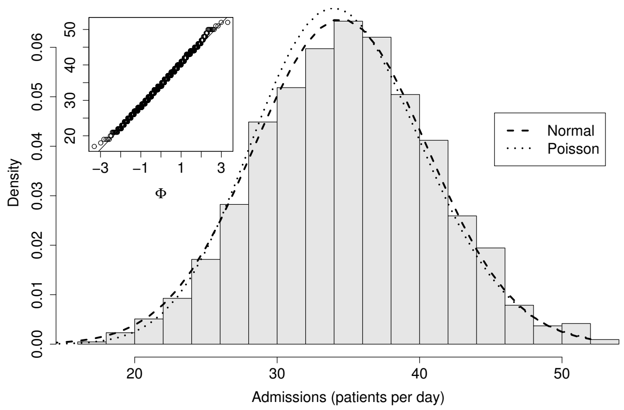
3) ARIMA Models: The Univariate Box-Jenkins methodology [15] was used to construct Autoregressive Integrated Moving-Average (ARIMA) models of admissions and occupancy. Two types of models were considered. The first is based on seasonal regression and ARIMA modelling of the residuals. The second is a Seasonal ARIMA (SARIMA) model.



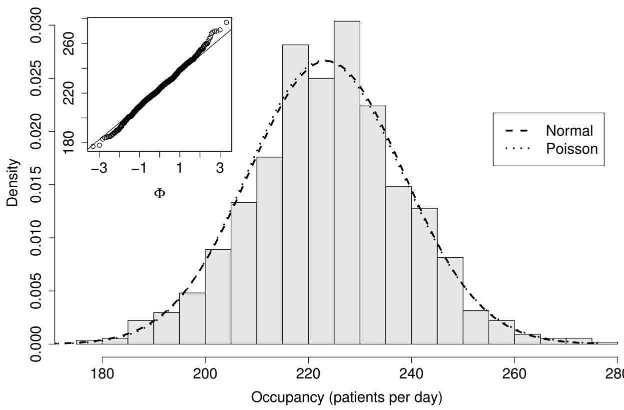
(a) Daily admissions, $\bar{x} = 34.5$, $S^2 = 36.9$.



(b) Daily occupancy, $\bar{x} = 223.5$, $S^2 = 224.3$.



(c) admissions



(d) occupancy

Fig. 1. Descriptive statistics for daily admissions and occupancy for the period 10/4/2004 to 25/3/2007 . Fig. 1(c) and 1(d) show histograms of admissions and occupancy, respectively, with the maximum-likelihood estimates for the Poisson and Normal distributions, and Standard Normal QQ-plots (insets).

a) Seasonal Regression (RegAR): Seasonal Regression assumes that seasonality is *deterministic*. The seasonal regression model was constructed by first detrending and deseasonalising the time series, using a least-squares regression of the dependent variable (admissions or occupancy) on time, and on indicator variables for day of the week and for month. While the month-of-the-year effect was significant and its inclusion in the regression (not shown) increased the R^2 scores, it did not improve forecast accuracy in the cross-validation stage (see Section III-C2), and therefore was not included in the final model. This is may be due to the fact that the daily effects were larger in magnitude than the monthly effects. The regression coefficients for admissions are shown in Table I, and for occupancy in Table II. The seasonal effect forecast $\hat{S}$ is expressed as

$$\hat{S}_t = \alpha t + \sum_{i=1}^{7} d_{i,t}\beta_i, \tag{3}$$

where t is time in days, d is an indicator variable for day-of-week, and α and β are the regression coefficients. Note that the model is parameterised in terms of an effect for each day of the week and no intercept, however the R^2 calculation used the mean as the null model.

	Estimate	Std. Error	p-value
Time (days)	0.003	0.001	≤ 0.001
Monday	32.943	0.570	≤ 0.001
Tuesday	32.784	0.570	≤ 0.001
Wednesday	33.397	0.571	≤ 0.001
Thursday	34.596	0.571	≤ 0.001
Friday	34.008	0.571	≤ 0.001
Saturday	32.287	0.570	≤ 0.001
Sunday	27.406	0.570	≤ 0.001

TABLE I
REGRESSION COEFFICIENTS FOR ADMISSIONS. $R^2 = 0.156$, ADJ. $R^2 = 0.151$.

The regression residuals are the difference between the observed value and the seasonal forecast

$$e_t = Y_t - \hat{S}_t. \tag{4}$$

The model is built by inspecting the regression residuals for autocorrelation using autocorrelation function (ACF) and partial autocorrelation function (PACF) plots. An ARIMA

	Estimate	Std. Error	p-value
Time (days)	0.014	0.001	≤ 0.001
Monday	224.965	1.279	≤ 0.001
Tuesday	221.438	1.280	≤ 0.001
Wednesday	216.605	1.280	≤ 0.001
Thursday	211.707	1.281	≤ 0.001
Friday	209.368	1.282	≤ 0.001
Saturday	203.000	1.277	≤ 0.001
Sunday	212.579	1.278	≤ 0.001

TABLE II
REGRESSION COEFFICIENTS FOR OCCUPANCY. $R^2 = 0.302$, ADJ. $R^2 = 0.297$.

TABLE III
ESTIMATED PARAMETERS FOR $\text{ARIMA}(5,0,0)$ MODEL OF REGRESSION RESIDUALS FOR ADMISSIONS, AND FOR $\text{ARIMA}(1,0,0)$ MODEL OF REGRESSION RESIDUALS FOR OCCUPANCY

	Admissions					Occupancy
	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_1
Estimate	0.0030	0.0753	0.0561	0.1109	0.0916	0.8024
Std. Err.	0.0303	0.0301	0.0302	0.0301	0.0303	0.0181

model is then fitted to the residuals, with the general form

$$\phi(B)e_t = \theta(B)a_t, \quad (5)$$

where $\tilde{z} = z - \mu$ is the value of the time series after subtracting the mean μ of the series, B is the backshift operator defined as $B\tilde{z}_t = \tilde{z}_{t-1}$, $\phi(B)$ and $\theta(B)$ are the polynomials in B for the autoregressive and moving-average components, respectively, and a_t is the error (*shock* in Box-Jenkins terminology) at time t.

The ACF/PACF plots of residuals of admissions regression showed weak autocorrelations for high lags (Fig. 2(a)). An ARIMA(5,0,0) model was fitted to the residuals, resulting in the model

$$(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3 - \phi_4 B^4 - \phi_5 B^5)e_t = a_t, \quad (6)$$

which when substituted into (4) yields the difference equation used for forecasting

$$\begin{aligned} Y_t = \hat{S}_t + \phi_1(Y_{t-1} - \hat{S}_{t-1}) + \phi_2(Y_{t-2} - \hat{S}_{t-2}) \\ + \phi_3(Y_{t-3} - \hat{S}_{t-3}) + \phi_4(Y_{t-4} - \hat{S}_{t-4}) \\ + \phi_5(Y_{t-5} - \hat{S}_{t-5}) + a_t. \quad (7) \end{aligned}$$

The estimated ARIMA coefficients are shown in Table III. Although the ACF showed small autocorrelations for longer lags, using higher-order models did not improve the fit.

Similarly, an ARIMA(1,0,0) model was fitted to the residuals of the occupancy regression (Fig. 2(b)), with the form

$$(1 - \phi_1 B)e_t = a_t, \quad (8)$$

and the resulting difference equation

$$Y_t = \hat{S}_t + \phi_1(Y_{t-1} - \hat{S}_{t-1}) + a_t. \quad (9)$$

The estimated coefficient is shown in Table III.

TABLE IV
ESTIMATED PARAMETERS FOR $\text{ARIMA}(0,0,0) \times (0,1,1)_7$ MODEL OF ADMISSIONS, AND FOR $\text{ARIMA}(1,0,0) \times (0,1,3)_7$ MODEL OF OCCUPANCY

	Admissions	Occupancy			
	Θ_1	ϕ_1	Θ_1	Θ_2	Θ_3
Estimate	0.9079	0.8099	0.8754	0.0188	0.0826
Std. Err.	0.0143	0.0202	0.0331	0.0368	0.0303

b) Seasonal ARIMA (SARIMA): Another approach to modelling seasonal time series is the Seasonal ARIMA (SARIMA) model. Such models are denoted $\text{ARIMA}(p,d,q) \times (P,D,Q)_s$, where p, d, and q are the non-seasonal AR, I, and MA coefficients, respectively, P, D, and Q are the seasonal coefficients, and s is the length of the season (here it is in days). SARIMA treats seasonality as *stochastically* varying, rather than deterministically varying. Only single-seasonality models were included (thus modelling the weekly effects but not the monthly effects).

The general SARIMA form [15, p. 332] is

$$\phi_p(B)\Phi_P(B^s)\nabla^d\nabla_s^D\tilde{z}_t = \theta_q(B)\Theta_Q(B^s)a_t, \quad (10)$$

where ∇ is the differencing operator defined as $\nabla\tilde{z}_t = \tilde{z}_t - \tilde{z}_{t-1}$, $\Phi_P(B)$ is the polynomial of P seasonal AR coefficients, and $\Theta_Q(B)$ is the polynomial of Q seasonal MA coefficients.

We examined several SARIMA models of admissions. The best model found, in terms of having lowest autocorrelation in the residuals and lowest variance of the residuals, was an $\text{ARIMA}(0,0,0) \times (0,1,1)_7$.

An $\text{ARIMA}(1,0,0) \times (0,1,3)_7$ was used to model occupancy. Substituting it into (10) yields

$$(1 - \phi_1 B)(\tilde{z}_t - \tilde{z}_{t-7}) = (1 - \Theta_1 B^7 - \Theta_2 B^{14} - \Theta_3 B^{21})a_t, \quad (11)$$

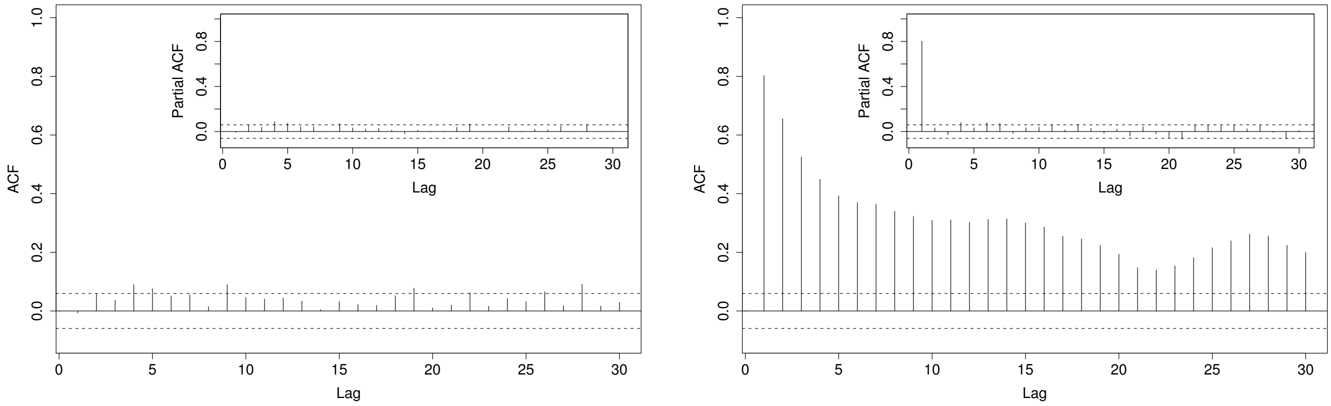
which can be expressed as the difference equation

$$\begin{aligned} \tilde{z}_t = \tilde{z}_{t-7} + \phi_1\tilde{z}_{t-1} - \phi_1\tilde{z}_{t-8} + a_t \\ - \Theta_1 a_{t-7} - \Theta_2 a_{t-14} - \Theta_3 a_{t-21}. \quad (12) \end{aligned}$$

The optimal parameters found for the SARIMA models of admissions and occupancy are shown in Table IV.

c) Modelling Discharge Proportions (AdmDis): The third approach to modelling occupancy combines models of admissions and discharges. Admissions are modelled using the previously described RegAR model. Discharges on any given day are dependent on that day's occupancy (prior to discharge). Therefore, discharges are modelled by performing a least-squares regression of the proportion of discharges on occupancy, using time and day-of-week as covariates. The regression coefficients are shown in Table V. The forecast occupancy $\hat{F}_{t+1}$ is the cumulative difference between admissions A_t and discharges D_t

$$\hat{F}_{t+1} = \hat{F}_t + A_t - D_t. \quad (13)$$



(a) Admissions

(b) Occupancy

Fig. 2. The autocorrelation function (ACF) and partial autocorrelation function (PACF) of the regression residuals for admissions and occupancy.

	Estimate	Std. Error	p-value
Time (days)	0.000	0.000	0.167
Monday	0.161	0.003	≤ 0.001
Tuesday	0.169	0.003	≤ 0.001
Wednesday	0.176	0.003	≤ 0.001
Thursday	0.173	0.003	≤ 0.001
Friday	0.191	0.003	≤ 0.001
Saturday	0.114	0.003	≤ 0.001
Sunday	0.074	0.003	≤ 0.001

TABLE V
REGRESSION COEFFICIENTS FOR DISCHARGE PROPORTIONS. $R^2 = 0.659$, ADJ. $R^2 = 0.656$.

d) Modelling Length of Stay (AdmLOS): The fourth and final model of occupancy models uses the RegAR model to model admissions, however, discharges are modelled using a model of patient length of stay (LOS). An empirical distribution of LOS was created from the data. Based on results from Renewal Theory (see, e.g., [16]), the distribution of LOS can be transformed into the Recurrence-Time distribution $g(x)$, which models *remaining* LOS using

$$g(x) = \frac{1 - F(x)}{\mu}, \tag{14}$$

where $F(x)$ is the cumulative distribution function (CDF) of LOS, and μ is the mean LOS, both estimated from the empirical LOS distribution.

The expected number of discharges D per day is the proportion of discharges per day, given by the probability $g(x = 0)$, multiplied by the previous occupancy. The forecast is again calculated using (13).

4) Number of Prior Data: Here we discuss the number of data required for building and using the models. We distinguish between two types of prior data. One is the number of previous days the ARIMA model uses for its forecast, also called *model order*, the other is the number of data used for estimating the model coefficients, also called *sample size*.

For the first type, the amount of prior data used depends on the model, since an ARIMA model forecasts future values based on the past. For example, an AR(1) model requires knowledge of the patient occupancy on one previous day, whereas a AR(10) requires 10 previous days. Models with higher orders may lead to lower forecast error since they take into account more of the past, however, there is also an increased risk of overfitting (equivalent to fitting too many parameters in a regression model). In practice, following Box-Jenkins methodology will lead to models with a suitable order, since examining the ACF/PACF plots (e.g., Fig. 2) shows the maximum lag at which there is still significant autocorrelation.

For the second type of prior data, as with any statistical estimation procedure, larger samples produce more stable estimates. Therefore, for a reasonable estimate of model coefficients (i.e., the ARIMA coefficients and the seasonal effects) we use all available data (since we use 10-fold cross-validation, this amounts to 9/10ths of the data in each fold).

Similarly, model order of the k-step moving-average and SES methods (Sections III-B1 and III-B2) is simply k and 1 respectively, and sample size for both is all available data.

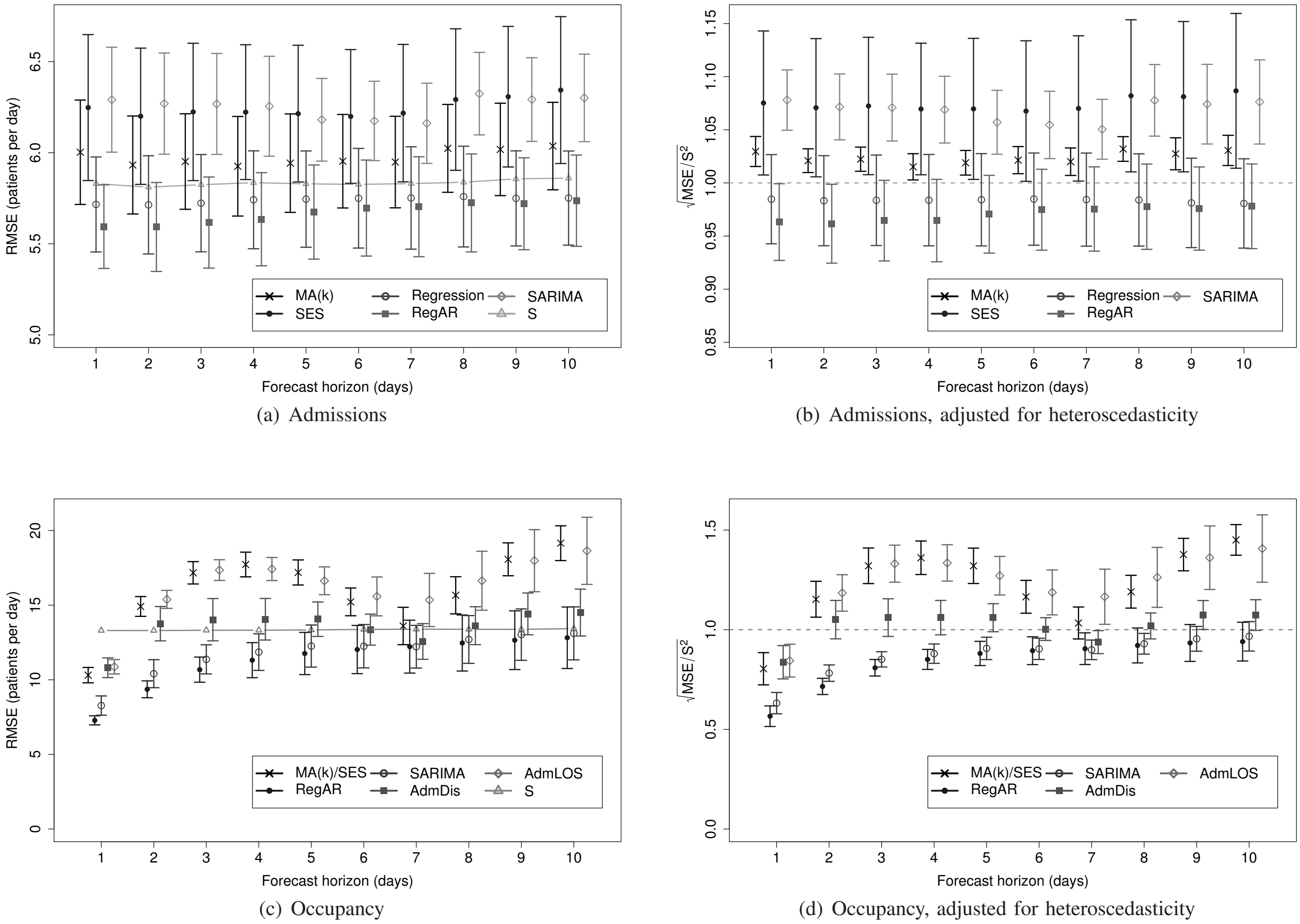
C. Model Selection

1) Measures of Model Quality: The models were compared using root mean-squared error (RMSE), defined as

$$\text{RMSE} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(Y_t - F_t)^2}, \tag{15}$$

where T is the number of days in the test set, and Y_t and F_t are the observed and forecast levels (admissions or occupancy) for day t, respectively.

2) Estimating the Forecasting Error: We must distinguish between a model's error in fit and its *error in forecast* (also called *generalisation error*). A high degree of fit can be achieved by incorporating many covariates. However, to be useful as a forecasting tool, a model must be able to generalise well, i.e., to forecast periods outside the one it was trained on. An *overfitted* model will fit the training data well, but forecast unseen data poorly (high forecast error) [17], [18].



(a) Admissions

(b) Admissions, adjusted for heteroscedasticity

(c) Occupancy

(d) Occupancy, adjusted for heteroscedasticity

Fig. 3. 95% confidence intervals for RMSE and normalised RMSE from 10-fold cross-validation of the admissions and occupancy models.

We used 10-fold *blocked cross-validation* (CV) to estimate the generalisation error [17], [19], [20]. The dataset is split into ten subsets (*blocks*). Nine are used for training the forecasting model (estimating model parameters) whereas one block is used to test the chosen parameters. This process is repeated ten times, each time using a different block as the test set. The natural ordering of days is kept within each block, thus maintaining the in-block autocorrelations. The estimated forecast RMSE is the square root of the average MSE (over the 10 test-blocks).

Ideally, we would like to estimate the standard error of the cross-validated RMSE, to determine whether two models have significantly different accuracy. However, this is not straightforward since the folds are not independent. Moreover, a general unbiased estimator of the variance of cross-validation does not exist [20]. For lack of a better alternative, the naïve variance estimator $S^2 = \frac{1}{n-1}\sum(x_i - \bar{x})^2$ was used.

Cross validation is only used for *selecting* the model of the alternatives, and to indicate how well the model generalises. Once the best model is chosen, it should then be trained on the *entire* data set before using it for real world forecasting, to avoid wasting any of the training data.

3) Model Performance: To provide a more intuitive description of model performance, we calculated two additional measures.

The 95% prediction interval (PI) of the RegAR model, as a measure of its precision. The 95% PI is the interval in which 95% of future values are expected to fall. The lth day ahead 95% prediction limit P_l is

$$P_l = 1.96 \times s_a \sqrt{1 + \sum_{j=1}^{l-1} \psi_j^2}, \tag{16}$$

where s_a is the estimated standard deviation of the shocks (given by the ARIMA estimation procedure), ψ are the L linear-filter coefficients (see [15, pp. 139–145]), and L is the maximal forecast horizon (here, 10 days). For an ARIMA(1,0,0) model, the filter coefficients are calculated recursively from the autoregression coefficients using

$$\psi_j = \begin{cases} 1 & j = 0, \\ \phi_1 \psi_{j-1} & 1 \leq j \leq L. \end{cases} \tag{17}$$

We also calculated the mean absolute-percentage error

(MAPE), defined as

$$\text{MAPE} = \frac{1}{T}\sum_{i=i}^{T}\left|\frac{Y_t - F_t}{Y_t}\right|, \tag{18}$$

using the same variable definitions as in (15).

D. Software

All analyses were performed using the R statistical environment, version 2.7.1 [21]. Supplementary information is available from the authors upon request.

IV. Results

A. Seasonal Effect

As shown in Tables I, II, the variation explained by day-of-week effects for admissions and occupancy is quite small: an R^2 of 0.151 and 0.297, respectively. For discharges (Table V), the R^2 is much higher (0.66), possibly due to different hospital discharge practices for different days of the week, such as fewer discharges over the weekend.

B. Models of Admission

Fig. 3(a) compares RMSE for the different models of admission, for different forecast horizons. Each of the ten folds has slightly different estimated variance. Fig. 3(b) shows the RMSE results normalised by the variance of each fold, thereby showing the proportion of unexplained variance. This is a better basis for model comparison than is RMSE alone since it is more robust to non-stationarity.

Good models have low unexplained variance. Of the models, the RegAR model performs best, however the proportion of unexplained variance is not significantly less than one. Other models perform consistently badly for all forecast horizons. This indicates that admissions are largely unpredictable — approximately “white noise” (a stationary Gaussian process).

C. Models of Occupancy

1) Model Accuracy: Fig. 3(c) shows the RMSE for forecast horizons of one to ten days ahead, and Fig. 3(d) shows the RMSE normalised by the variance. The results show a clear advantage of the RegAR and SARIMA models over the moving average and SES models. Although the RegAR seems to be slightly better than the SARIMA model, this effect is not likely to be significant. The proportion of unexplained variance remains below one for a forecast horizon of up to 7 days for RegAR and up to 8 days for SARIMA. The AdmDis and AdmLOS models have some forecasting ability for one day ahead, however, they are not useful beyond that.

Here, SES and MA(k) represent the same model. The optimal k for all 10 blocks was 1, whereas the optimal α value was 1 as well. Substituting the parameters back into (1) and (2), we see that the optimal SES model is identical to the optimal MA(k) model. Both models cannot capture the patterns of occupancy and are not useful for forecasting it.

An accurate prediction of LOS may increase the accuracy of the AdmLOS model. However, LOS depends on many covariates, such as patient age, comorbidities, staffing levels, and hospital congestion. (See [22] for various approaches to modelling LOS.) We found that including day-of-week and month covariates resulted only in negligible RMSE gain. Other covariates, such as patient age and comorbidities, are difficult to forecast as they require modelling of ageing and occurrence of disease in the population.

2) Model Performance: Fig. 4(a) shows the 95% prediction interval (PI) half-width for the RegAR occupancy model. It should be read as the forecasting precision — given the one day ahead occupancy forecast F_{t+1}, the 95% PI for the RegAR model is $F_{t+1} \pm 14$ patients per day. The actual coverage (*ex post*) of the 95% PI was checked using cross-validation. Averaged over the 10-day forecast horizon, it was approximately 94% (results not shown). Also shown is the PI for an ARIMA model (denoted AR) of identical order but without first accounting for the day-of-week effects. Fig. 4(b) shows the distribution of MAPE for the RegAR and AR occupancy model, for forecast horizons of 1 to 10 days. Clearly, including the day-of-week effects results in better forecasting ability, measured both in the PI and MAPE.

V. Discussion

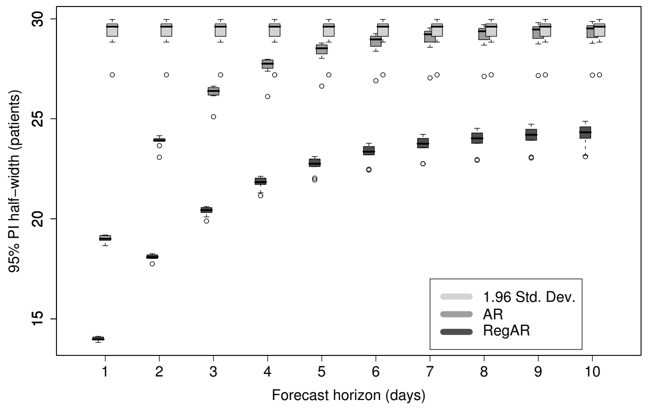
We have found that none of the admission models could produce useful forecasts for any forecast horizon (Fig 3(a), 3(b)). In line with [9], we conclude that admissions are largely unpredictable, with a only small amount of variance explained by the trend and day-of-week.

The best-performing of the occupancy models, RegAR and SARIMA, could explain most of the variance of occupancy for up to two days ahead, and some of the variance for up to one week ahead (Fig 3(c), 3(d)). Beyond that, they provide no useful information. For comparison, note that the average patient length of stay is approximately six days.

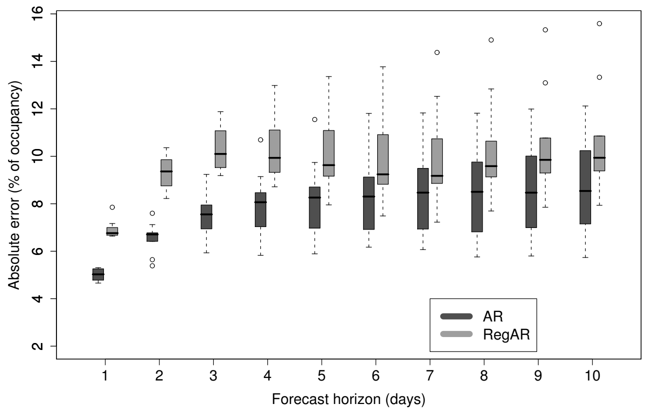
Litvak *et al.* [23] introduced the notion of *natural* versus *artificial* variability in patient flows. The former is due to the inherent randomness of patient arrivals, whereas the latter is due to suboptimal hospital decision-making. In these terms, emergency admissions seems naturally variable. The implicit hypothesis behind the AdmDis and AdmLOS models is that knowledge of the inputs to the hospital, together with predicted discharges, would lead to good models of occupancy. In other words, the hypothesis is that natural variability is the main driver of occupancy variability. However, as the results show, modelling occupancy itself is more useful. This suggests that variability in occupancy is likely due to both the natural variability in admissions and LOS, and to the artificial variability in hospital practices, such as staff levels and patient discharge procedures.

A. Model Precision and Implications

To better understand the practical implications of the prediction intervals discussed in Section IV-C2, assume that emergency occupancy is forecast using some method. Elective bookings are then allowed to use the remaining beds. An underestimate of true emergency demand would lead to a conflict between the bed requirements of emergency and



(a) 95% PI half-widths for the RegAR and AR models, and for a null model which is simply the average occupancy. The PIs are averaged over all the cross-validation folds.



(b) MAPE, over all CV folds.

Fig. 4. Results for RegAR model and an ARIMA model of same order but without day-of-week regression, denoted AR.

elective patients. If hospital management wishes to allow up to 9 conflicts between emergency and electives per year ($(1-0.95)/2 \times 365 \approx 9$), they would need to accept the upper limit of the 95% PI, and the corresponding expected number of empty beds. Taking the null occupancy model, a stationary Gaussian process, implies a reserve of approximately 29 beds per day (in a 350 bed hospital), for all forecast horizons. The RegAR model reduces this to approximately 20 beds, if electives can be booked 3 days in advance. Beyond 7 days lead time, the expected number of reserve beds is 24. Demands for greater certainty, in terms of rarer conflicts, require larger reserves; with the RegAR model, two conflicts per year require using 99% PIs with the expected reserve being 27 and 31 beds, for the 3 and 7 days ahead cases, respectively. Conversely, allowing for approximately one conflict every 3 weeks permits the use of a 90% PI and a smaller reserve: 17 and 20 beds, respectively.

B. Stationarity and Invertibility

Here we discuss the requirements of stationarity and invertibility [15, pp. 50–51]; for brevity, the numerical results are not included. An ARIMA (or SARIMA) model must satisfy the conditions of stationarity and invertibility. A (weakly) *stationary* process is a process that fluctuates about a constant mean, with roughly constant variance (homoscedasticity). An AR process of order p, with parameters $\phi_1, \cdots, \phi_p$ is stationary if and only if the zeros of the polynomial in B, $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$, lie *outside* the unit circle (in the complex plane), where B is the backward shift operator $Bz_t = z_{t-1}$. Conversely, An MA process of order q, with parameters $\theta_1, \cdots, \theta_q$, is *invertible* whenever the zeros of its polynomial $\theta(B)$ lie outside the unit circle. For an ARIMA(p, d, q) model, the AR(p) component must be stationary and the MA(q) component must be invertible.

All of the RegAR models used here are stationary (they have no MA component, therefore, there is no invertibility concern). The SARIMA admissions model is both stationary and invertible. However, for the SARIMA occupancy model, the last cross-validation test-set is non-invertible. This might indicate a structural break in that period, requiring a different model. Several SARIMA models were investigated, however, we found no invertible model for this block (within the cross-validation methodology). When trained on the *entire* data set, both the admissions and the occupancy SARIMA models are stationary and invertible.

C. Limitations and Practical Considerations

Here we consider some limitations of time-series analysis models and of ARIMA models in particular. First, the correct application of ARIMA models depends on stationarity of the time series being modelled (either inherently or through stationarisation). We cannot be certain that the time series is actually stationary, since there may be slowly-changing effects on time scales longer than the available data. Second, monthly effects were excluded from the regression since they did not reduce RMSE. These effects, together with other covariates, may have proven important had a larger sample size been used. Third, errors of model misspecification [24] were not considered. Fourth, ARIMA models only capture short-range dependencies; long-memory models were not considered. Fifth, estimated model parameters reflect the behaviour of the particular hospital, at a particular point in time. Therefore, such models cannot be applied blindly to all hospitals, instead requiring fine-tuning for each case, including possible periodic parameter re-estimation to compensate for policy changes. Finally, when applying the model, consideration should be given to patient mix and whether it changes over time. In this work we assume that patient mix is roughly stable over time.

VI. Conclusion

We have shown that the admissions from the emergency department to the wards are largely unpredictable. The models examined can forecast the resulting occupancy in the short term, however, they are not useful for forecast horizons of over one week. Hospital managers may use the models discussed here to gain insight into emergency patient flow and to plan elective admissions better. However, they should be aware of model limitations and especially of the inherent variability of emergency inpatient flow.

ACKNOWLEDGEMENT

Thanks to Peter G. Taylor from the University of Melbourne for discussion of renewal theory, to Caroline Brand from Melbourne Health for critical reading of early drafts, and to Paul Barfett from Melbourne Health for help with the data.

REFERENCES

[1] S. Ivatts and P. Millard, "Health care modelling: opening the 'black box'," *British J. Health Care Manag.*, vol. 8, pp. 251–255, 2002.
[2] ——, "Health care modelling: why should we try?" *British J. Health Care Manag.*, vol. 8, pp. 218–222, 2002.
[3] Department of Human Services, *Victorian Public Hospitals: Performance Monitoring Framework, Business Rules 2006–07*, Victoria, Australia, 2006.
[4] D. M. Fatovich, Y. Nagree, and P. Sprivulis, "Access block causes emergency department overcrowding and ambulance diversion in Perth, Western Australia," *Emerg. Med. J.*, vol. 22, pp. 351–354, 2005.
[5] G. W. Harrison, A. Shafer, and M. Mackay, "Modelling Variability in Hospital Bed Occupancy," *Health Care Manag. Sci.*, vol. 8, pp. 325–334, 2005.
[6] H. Xie, T. J. Chaussalet, and P. H. Millard, "A Model-Based Approach to the Analysis of Patterns of Length of Stay in Institutional Long-Term Care," *IEEE Trans. Info. Tech. Biomed.*, vol. 10, pp. 512–518, 2006.
[7] E. Akçali, M. J. Côté, and C. Lin, "A network flow approach to optimizing hospital bed capacity decisions," *Health Care Manag. Sci.*, vol. 9, pp. 391–404, 2006.
[8] W. T. Lin, "Modelling and forecasting hospital patient movements: Univariate and multiple time series approaches," *Int. J. Forecasting*, vol. 5, pp. 195–208, 1989.
[9] S. A. Jones, M. P. Joy, and J. Pearson, "Forecasting Demand of Emergency Care," *Health Care Manag. Sci.*, vol. 5, pp. 297–305, 2002.
[10] R. Champion, L. D. Kinsman, G. A. Lee, K. A. Masman, E. A. May, T. M. Mills, M. D. Taylor, P. R. Thomas, and R. J. Williams, "Forecasting emergency department presentations," *Aust. Health Rev.*, vol. 31, pp. 83–90, 2007.
[11] A. Earnest, M. I. Chen, D. Ng, and L. Y. Sin, "Using autoregressive integrated moving average (ARIMA) models to predict and monitor the number of beds occupied during a SARS outbreak in a tertiary hospital in Singapore," *BMC Health Serv. Res.*, vol. 5, 2005.
[12] N. Channouf, P. L'Ecuyer, A. Ingolfsson, and A. N. Avramidis, "The application of forecasting techniques to modeling emergency medical system calls in Calgary, Alberta," *Health Care Manag. Sci.*, vol. 10, pp. 25–45, 2007.
[13] *When does quality assurance in health care require independent ethical review?*, National Health and Medical Research Council, 2003. [Online]. Available: http://www.nhmrc.gov.au/ethics/human/conduct/guidelines/_files/e46.pdf
[14] R. J. Hyndman, A. B. Koehler, R. D. Snyder, and S. Grose, "A state space framework for automatic forecasting using exponential smoothing methods," *Int. J. Forecasting*, vol. 18, pp. 439–454, 2002.
[15] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time series analysis: forecasting and control*, 3rd ed. Prentice Hall, 1994.
[16] D. R. Cox, *Renewal Theory*. London, Great Britain: Methuen and Co., 1962.
[17] R. Kohavi, "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection," in *Proc. 14th IJCAI*, 1995, pp. 1137–1145.
[18] M. Mackay and M. Lee, "Choice of models for the analysis and forecasting of hospital beds," *Health Care Manag. Sci.*, vol. 8, pp. 221–230, 2005.
[19] T. E. Clark, "Can Out-of-Sample Forecast Comparisons Help Prevent Overfitting?" *J. Forecasting*, vol. 23, pp. 115–139, 2004.
[20] Y. Bengio and Y. Grandvalet, "No Unbiased Estimator of the Variance of K-Fold Cross-Validation," *JMLR*, vol. 5, pp. 1089–1105, 2004.
[21] R Development Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2007, ISBN 3–900051–07–0. [Online]. Available: http://www.R-project.org
[22] C. Vasilakis and A. H. Marshall, "Modelling nationwide hospital length of stay: opening the black box," *J. Oper. Res. Soc.*, vol. 56, pp. 862–869, 2005.
[23] E. Litvak, P. I. Buerhaus, F. Davidoff, M. C. Long, M. L. McManus, and D. M. Berwick, "Managing Unnecessary Variability in Patient Demand to Reduce Nursing Stress and Improve Patient Safety," *J. Qual. Patient Safety*, vol. 31, pp. 330–338, 2005.
[24] C. Chatfield, "Model uncertainty, data mining and statistical inference," *J. Royal Stats. Soc. A.*, vol. 158, pp. 419–466, 1995.

Gad Abraham received the BAppSci(Hons) degree in computer science from RMIT University, Melbourne, Australia, in 2005. He is currently a Master's student at the Department of Computer Science and Software Engineering at the University of Melbourne. His research interests include bioinformatics and mathematical modelling of health care systems.

Graham B. Byrnes was originally trained in applied mathematics in Australia. He now works principally in biostatistics and biological modelling, particularly related to cancer genetics. Since the beginning of 2008 he is employed as a biostatistician by the International Agency for Research on Cancer in Lyon, France.

Christopher A. Bain is a medical practitioner and information technologist by training, and is currently the Information Manager at the Western and Central Melbourne Integrated Cancer Service (formerly Manager of the Information and Analysis Team in the Redesigning Care and Patient Access Unit (RCPAU) at Royal Melbourne Hospital). His primary interest is in developing better health care management information systems. He is currently completing a PhD in regard to hospital management information systems in the School of Management at Edith Cowan University in Perth, Australia, and is a partner investigator on the ARC Linkage Project titled "Modelling patient flows through hospitals: optimising the effective use of resources".

View publication stats