MODELLING HOSPITAL ADMISSIONS AND OCCUPANCY INCLUDING METEOROLOGICAL INFORMATION

Paul Harper
School of Mathematics, Cardiff University, Cardiff, UK
harper@cardiff.ac.uk

John Minty
School of Mathematics, Cardiff University, Cardiff, UK
harper@cardiff.ac.uk

Sujit Sahu
Mathematical Sciences, University of Southampton, Southampton, UK
harper@cardiff.ac.uk

Bernard Baffour
Mathematical Sciences, University of Southampton, Southampton, UK
harper@cardiff.ac.uk

Christophe Sarran
Met Office
harper@cardiff.ac.uk

Julie Vile
School of Mathematics, Cardiff University, Cardiff, UK
harper@cardiff.ac.uk

ABSTRACT

Improved short-term predictions of hospital admissions and bed occupancy offer the potential to plan resource needs more accurately and effectively. The MetSim project explores the relationship between weather and health, building novel Bayesian models that are more sensitive to fluctuations in weather. Short-term forecasts (21-days ahead) of the numbers of admissions, categorised by age, gender and medical condition, are produced. In turn, coupled with predictions on length of stay and information on current occupancy, MetSim uses hazard ratios embedded within a simulation framework to provide forecasts of short-term bed needs. MetSim is a collaboration between Cardiff University, the University of Southampton, and the Met Office. In the UK, the Cardiff and Vale University Health Board, Southampton University Hospitals NHS Trust and Anuerin Bevan University Health Board have all guided the development of MetSim, provided data and piloted the tool.

KEYWORDS. Forecasting demand. Hospital capacity management. Weather. Simulation. Bayesian models.

1. Introduction

More than 2,000 years ago, Hippocrates first recognised that epidemics were related to seasonal changes in weather. However, it was only during the 1970s that research into connecting weather and health was taken seriously and, for the first time, meteorological variables were investigated to gain insight into the causes of increased mortality in winter and smaller increases in unusually hot weather (Keatinge 2002). Since then, the interest in the effects of weather on health has grown substantially, helped to some extent by raised awareness of global warming and concern about the public health impact of climate change. Knowledge on the influence of weather on health is valuable, and has the ability to contribute greatly to our understanding of epidemiology, the occurrence of accidents and injuries, and of public health issues. Examples of weather-health research from the literature include those relating to: extreme weather events (WMO, 2003); sunshine, such as skin cancer (Cancer Research UK, 2012) and Seasonal Affective Disorder (Garland, 2003); temperature, such as cold weather and mortality (Hajat et al., 2002); Thunderstorms, such as lightning strikes (Elsom, 2001) and leading to increased asthma attacks (Venables 1997, Dales et al. 2003, New Scientist, February 2006); and snow/ice leading to fractures (Smith and Nelson 1998).

The ability to predict weather offers the potential to provide valuable information that can be used in planning health services. For example, imagine a short-term hospital planning tool that was able to predict fluctuations in demand and bed occupancy for different specialities by including meteorological predictions alongside other known information such as day of the week. The relationship between weather and health is immediately evident in some specialities, for example respiratory medicine.

Figure 1 shows respiratory admissions data from Southampton General Hospital for both those aged 9 and under, and 80+. The top graph shows temperature over a five-year period. For the other 2 plots, the black lines indicate total weekly admissions and discharges, while the red line indicates occupancy. As the graphs indicate, young persons are highly susceptible to cold; admissions and discharges are sharply increased. Elderly patients too have increased rates of admissions when temperatures are low, although much less marked than young patients, but once admitted they are prone to stay a long time before discharge. It is that fact which leads to high numbers of elderly patients occupying beds and which gives such concern to hospital managers. Similar plots have been produced with data from other UK hospitals.

The MetSim project is a multidisciplinary collaboration involving academics (from OR and Statistics), meteorologists from the Met Office, and managers and consultants from hospitals. It is beneficial for managers of hospitals to have short-term forecasts of demand and occupancy. Of particular interest is the number of non-elective (emergency) patients. Our particular focus is on patients who are admitted into hospital, as opposed to attending Accident and Emergency. The objectives of the MetSim project are:

- To describe hospital occupancy in two parts: rate of admissions and length of stay.
- To explore relationship between weather patterns and hospital admissions.
- To describe different lengths of stay for different classes of patient.
- To build a web-based tool which provides hospitals with short-term (21-day) predictions on demand and corresponding bed occupancy.

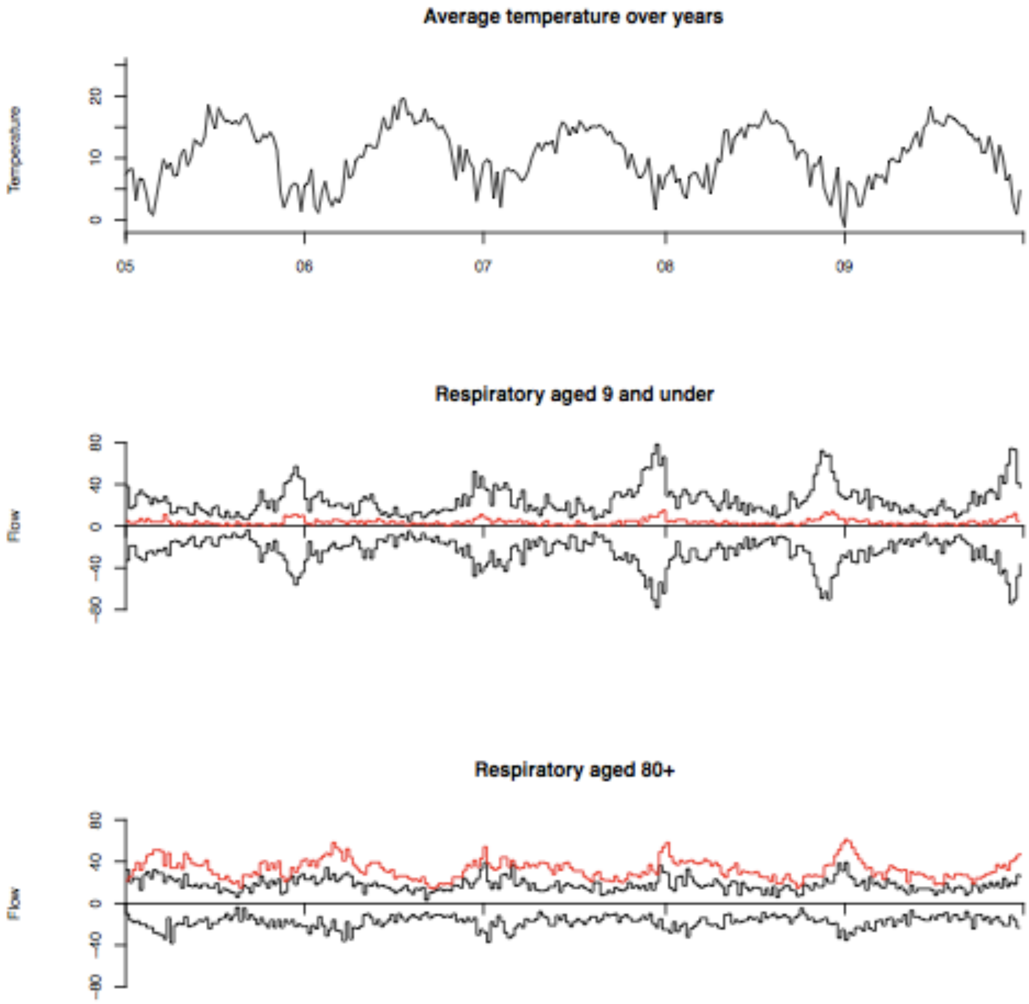

Figure 1: Respiratory patient flows at Southampton General Hospital

2. Datasets and design

When a hospital uses the tool for the first time, it provides rules on how patients are to be classified, and that is done only once. By contrast, a participating hospital regularly provides datasets on rows of patients. One set, labelled *historic*, covers an interval which is at least a year long and which strictly precedes the date on which the code was run, “the date of execution". The second dataset, labelled *current*, is simply a census of prevalent patients, namely patients recorded as present shortly before the start of the forecast. The Met Office provides datasets on weather, executes the machinery of the code and returns summary forecasts back to the hospital. Figure 2 illustrates the design of the tool and flow of decencies.

Anonymized patient admission and discharge data from participating hospitals have been linked to meteorological data provided by the Met Office. We summarise the data types below.

2.1. Historic hospital Data

For every admission/discharge of a patient over the course of a year the hospital records the age at admission, gender, broad class of treatment (medicine, surgery, trauma, paediatric or other), date of admission, and date of discharge. Hour of admission and discharge is optional. Ideally, the year of observation should be from 14 months ago to 2 months ago.

2.2. Current Hospital Data

For some day during the last week, the hospital gives a census of all prevalent patients. The items recorded are as for historic data except that there is, perforce, no date of discharge.

Meteorological Data

Over the entire time period, historic, current and forecast, the Met Office records the mean temperature on a given day and the minimum one week ago.

Temporal Data

Other variables needed are school holidays, public holidays, and day of the week. The historic hospital dataset is used to select models and estimate parameters. The current and forecast temperatures are then used to forecast admissions. The current hospital data are used only to simulate occupancy.

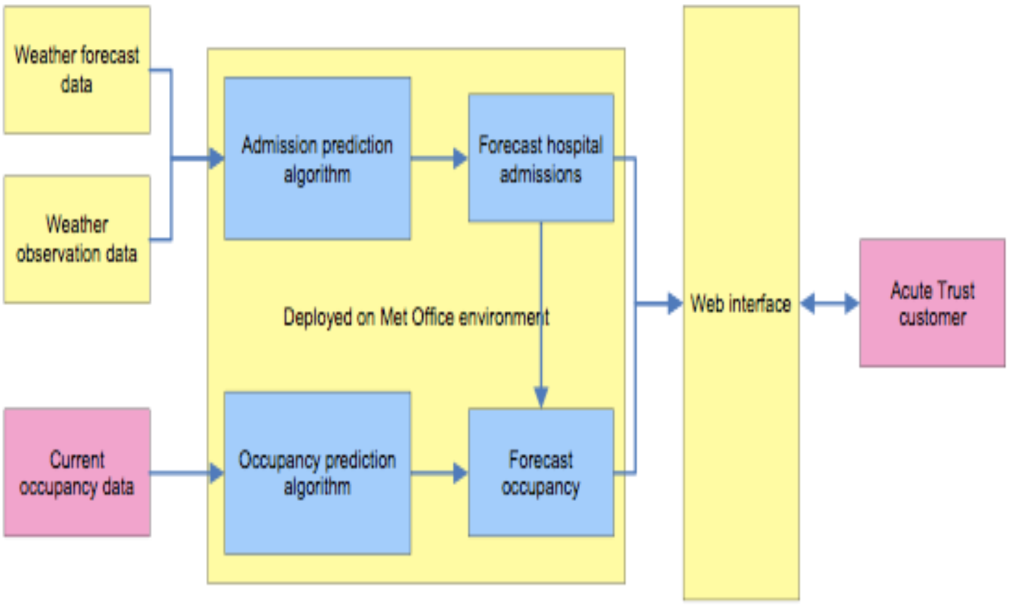

Figure 2: MetSim schematic

Historiography

We initially analysed datasets much larger than the ones in the final version of the model. The hospitals' historic datasets included the method of admission, full episodic progression of patient-spells and destination on discharge. The meteorological datasets included humidity, pressure, vapour pressure, rainfall and wind speed.

3. Forecasting Admissions

A comprehensive explanation of the forecasting element can be read in Sahu et al. (2014); a summary is provided here.

As anticipated, age is a significant explanatory variable; we partitioned patients into 0-17 as paediatric, 18-74 as adult and 75+ as elderly (on guidance from the hospitals). We used gender, as much for logistic (planning for single sex wards) as statistical reasons. Temperature is a significant explanatory variable as is current day of the week. Figure 3 illustrates that admissions are higher during weekdays than at weekends.

The data reveal that the number of daily admissions is naturally positively skewed. To overcome this in modelling, it is sufficient to take a square-root transformation. Having fitted a large number of models to the historic hospital dataset we reached the following conclusions:

A model, detailed below, which includes age, sex, day of the week, whether the day is a school holiday, mean daily temperature and minimum temperature a week ago is the best main effects model according to both the R^2 and AIC. Such a model is very parsimonious. We choose to fit a generalized linear model of admissions in preference to a time-series model. One purpose is to keep computations tractable. Also, once we have included day of the week as a factor, any time series effect lies predominantly in the weather, which is given to us. The data are positively skewed and we select a square-root transformation for the response variable number of admissions. We denote the rescaled response variable by Y_{ijt} where i is sex, j is age group and t time. The best fitted regression model is given by:

$$Y_{ijt} \sim N(\mu_{ijt}, \sigma^2)$$

where

$$\mu_{ijt} = \beta_0 + \alpha_i + \gamma_j + h(t) + w(t) + \lambda_m m(t) + \lambda_n n(t) + (\alpha{:}\gamma)_{ij} + (\gamma{:}n(t))_j$$

The initial term β_0 is an intercept. The factor $\alpha_i$denotes gender and is zero for female. The factor γ_j is the jth age group and is zero for paediatrics. The effect $h(t)$ is school holiday and is zero for non-holidays. The effect $w(t)$ is day of the week; it is zero for Sundays. On day t, the mean temperature is $m(t)$ and the minimum a week ago is $n(t)$; the coefficients are λ_m and $\lambda_n$respectively. We have interactive terms also (age-gender and interaction between age and minimum temperature a week ago).The model explains over 80% of the variation in daily admissions.

The model is found from historic data with observed weather temperatures. When predicting future admissions, we rely on weather forecasts. Accordingly, we regard actual future temperature as some linear function of forecast temperature, putting Bayesian uncertainty on the coefficients of the linear relation. Details of the Bayesian model can be seen in Sahu et al (2014).

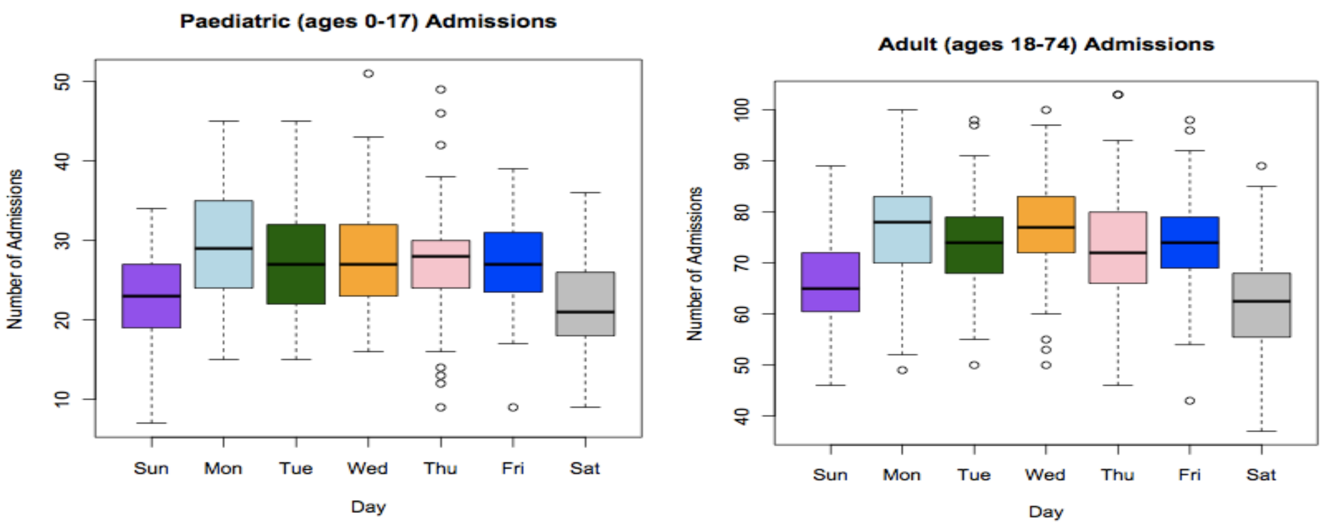

Figure 3: Boxplots of admissions by day

4. Length of stay

In the literature, there are models of lengths of stay for certain groups of patients, cardiac patients for example. Length of stay is a non-negative discrete number of days $n \geq 0$. But we prefer to build models of survival, since our patients may be somewhat heterogeneous. In this context, “survival" is not leaving hospital, whether by discharge, transfer or death. Let $h(n)$ denote the hazard rate for day n, namely the probability of exiting hospital on that day, conditional on having survived thus far. Let

$$K(n) = ln\frac{h(n)}{1-h(n)}$$

denote the log-odds of the hazard rate. For each fixed but arbitrary day n, we model $K(n)$ in terms of the explanatory variables age, gender, current day of the week and broad groupings of the patients' specialities (treatments). Figure 4 shows examples for non-paediatric length of stay on how the log-odds ratio $K(n)$ evolves. Furthermore it is shown that it depends also on age and current day of the week.

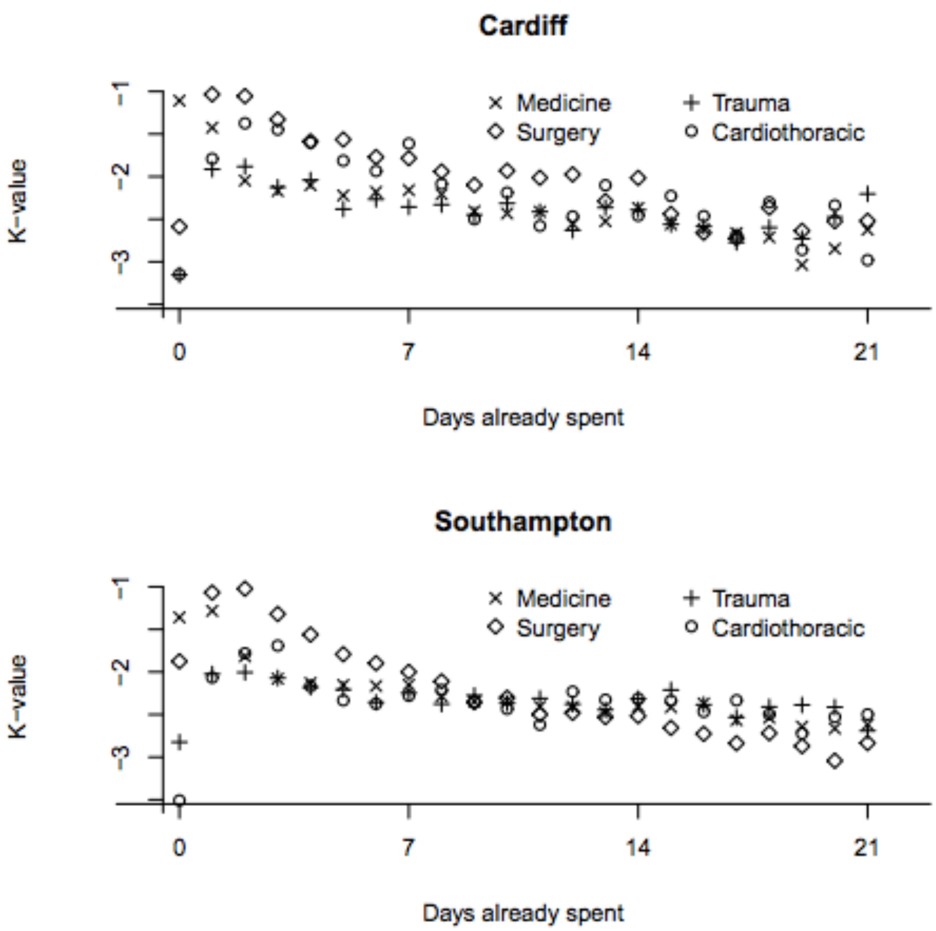

Figure 4: Log-odds of hazard rates (“K-value”, equivalently $K(n)$) for non-Paediatric patients across different specialities.

5. Simulation structure

The admissions model, as described in section 3, generates sampled streams of admissions over the 21 day timeline. It distinguishes patients by age-group and gender but not by the speciality supergroups. So for any age-group i and gender j the first step is to split streams of admissions, using multinomial probabilities p_{ijk} that a patient falls into the respective supergroup. For each i, j the sum $\sum_k p_{ijk} = 1$ where $1 \leq k \leq 2$ for paediatrics and $1 \leq k \leq 5$ for non-paediatrics. To estimate p_{ijk} we read the observed proportions in our historic dataset. Now we go through every row of admissions in the source file for i, j and generate a substream for each supergroup. We repeat for the same row of our source file to give a second simulated split.

Throughout the main simulation we consider patients at the finest level of categorization. As shown in Figure 5, for any age-group, gender and speciality supergroup we have three large sources of information: models of hazard rates of leaving hospital, simulated streams of admissions over the timeline, and a file which contains the distribution of prevalent patients by length of stay for some recent day. The timeline holds additional information such as bank holidays. The main aims of the simulation are, of course, to simulate streams of discharge and occupancy over the timeline. "Discharge" is a shorthand meaning all forms of exit. A secondary aim is to give a detailed profile of occupancy for a single future date during the interval of forecast: not just the total number of occupants forecasted but also how long these future occupants have themselves stayed.

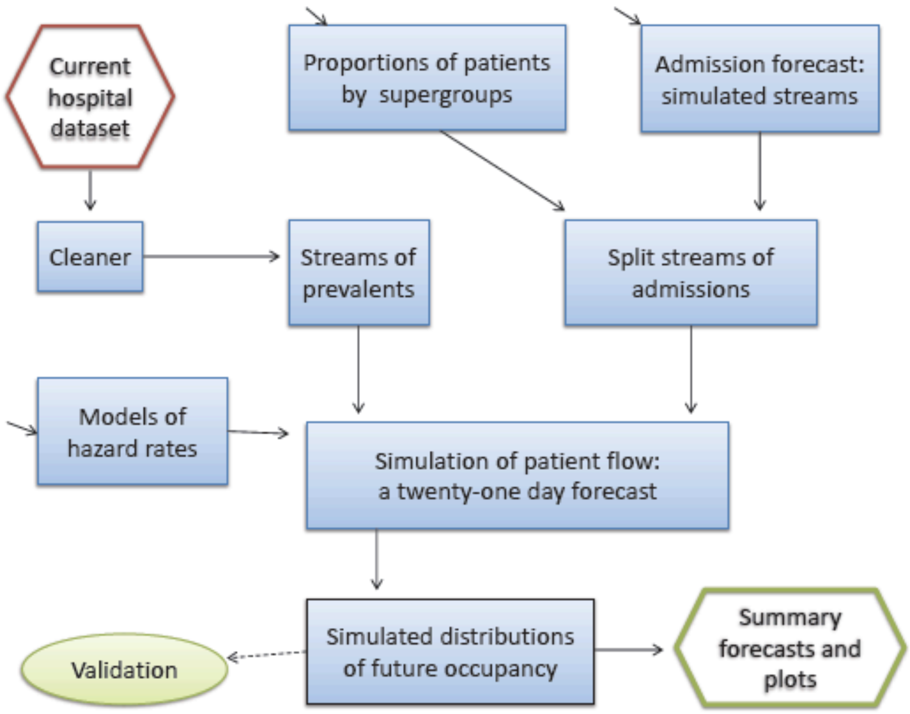

Figure 5: Simulating patient flow

We partition our chosen category of patient into cohorts by length of stay. Explicitly, for each day n, let

$$\boldsymbol{S_n} = (s_0, s_1, s_2 \ldots)$$

denote the counts of patients s_k who have already stayed $k \geq 0$ days. In particular, s_0 is the number of patients who have been admitted only recently.

The sum $\|\boldsymbol{S_n}\| = \sum_k s_k$ gives the total occupancy on day n. The simulation constructs $\boldsymbol{S_n}$ recursively along the timeline. Initially, we set s_0 to be 0; that is for the last trusted day, immediately before the start of our simulated admissions. The general method of constructing s_{n+1} from s_n is as follows.

1. Using binomial distributions $Bin(s_k, p_k)$ where p_k is the appropriate hazard rate, generate a vector:
$$\boldsymbol{D_n} = (d_0, d_1, d_2 \dots)$$
showing the number of discharges on day n.
2. Take the difference to find the number of remaining patients
$$\boldsymbol{S_n} - \boldsymbol{D_n} = (s_0 - d_0, s_1 - d_1, s_2 - d_2 \dots)$$
3. Slide one place to the right and insert the admissions; that gives the occupancy for the next day.
$$\boldsymbol{S_{n+1}} = (a, s_0 - d_0, s_1 - d_1, s_2 - d_2 \dots)$$

In practice, to speed up computations in step 1 of the algorithm, we compute tables of binomial probabilities before we start the iterations.

We thus obtain forecasts for admissions, discharges, occupancy and change in occupancy. For example, Figure 6 shows an illustrative forecast of admissions for the 21-day planning horizon, with 80% and 95% confidence intervals.

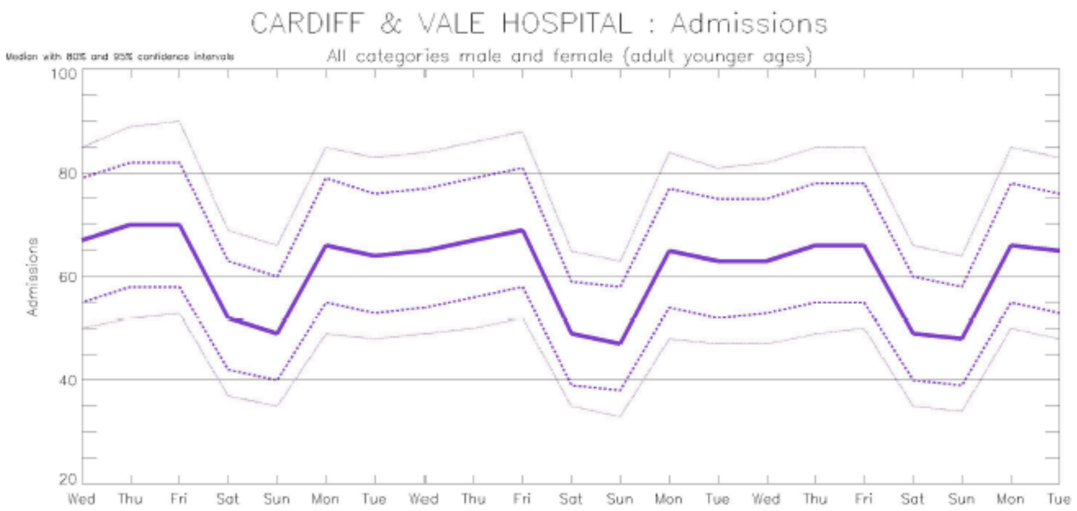

Figure 6: Illustrative output on admission forecasts

6. Some results to-date

MetSim has been piloted by Cardiff and Southampton hospitals, and is currently undergoing further testing and validation at the Royal Gwent hospital in South Wales. We present here a summary of some key findings to-date.

Prior to this tool, hospitals widely use one of two methods to forecast admissions and occupancy. The simplest model is the Persistence Model where the hospital looks at the figures for one year ago and adjusts for demographic trends. The other model is the Six-Week Moving Average. From preliminary validation and verification, our model seems at least comparable to the Six-Week Moving Average and both clearly outperform the Persistence Model. For example, Table 2 shows the relative root mean square error for Cardiff over a validation interval of 130+ days. We give the figures for Cardiff in preference to Southampton since we have a longer interval of validation.

For $0 \leq k \leq 20$ days ahead, the relative root mean square error is defined to be:

$$\frac{RMSE(y_k)}{RMSE(z_k)} - 1$$

where y_k and $z_k$refer respectively to the medians forecasted by our tool and the medians forecasted by the Six Week MA. On this measure, negative values are desirable, and that is what happens for 15 out of 21 days of forecasting. Other measures, such as percentage cover, confirm that the tool gives satisfactory forecasts that are overall better than the best of the existing forecasting tools used in the UK health service.

Table 1: Cardiff adult admissions

Days ahead	Relative RMSE				
0	0.039	7	−0.019	14	−0.009
1	0.004	8	−0.015	15	0.007
2	−0.014	9	−0.019	16	0.010
3	−0.026	10	−0.033	17	−0.006
4	−0.004	11	−0.005	18	−0.000
5	−0.022	12	0.035	19	−0.014
6	−0.013	13	0.009	20	−0.003

7. Conclusions

This paper outlines the underpinning methodology of the MetSim tool, designed to support hospital managers in predicting short-term demand and bed occupancy. Initially a Bayesian statistical model (full details are not included in this proceedings paper in the interests of space) is used to forecast demand for different categories/conditions of admissions. This tool is currently being piloted abut initial results (some of which are shown in section 6) at the time of writing this paper are promising. For example, over a one-year period at Cardiff Hospital, the Root Mean Square Error (RMSE) of 7-day ahead forecasts was just 4.8.

Demand forecasts are then fed into a simulation framework to produce corresponding bed occupancy predictions over the planning horizon. To do this, we simulate length of stay for the predicted admissions using hazard rates (such that the time a patient spends in hospital is modelled using 'survival' analysis techniques). The simulation is coded in C++ and sits on the server at the Met Office, Exeter, UK. A number of routines are run (e.g. that shown in Figure 5) typically for 1000's of iterations (and thanks to the power of the supercomputer at the Met Office, are executed within seconds), thus participating hospitals are provided on a daily basis with forecasts and associated confidence intervals.

Over the next few months we will streamline the system to have a web interface for ease of use and are currently performing a more comprehensive pilot at the Royal Gwent Hospital.

Acknowledgements

The authors are grateful to participating hospitals for discussion and data provided, especially to B Hendy from Southampton University Hospital Trust, to A Nelson, S Tarr and J Peters from Cardiff and Vale Health Board and to R Blackwell at Royal Devon and Exeter; also to supporting staff at the Met Office especially P Sachon and Y Clewlow, to G Pierce from Cardiff University and for grant EP/H010637/1 from the Engineering and Physical Sciences Research Council.

References

Cancer Research UK, 2012. http://www.cancerresearchuk.org. Accessed on the 1st May 2012.

Dales, R.E., Cakmak, S, Judek, S, Dann, T, Coates, F *et al* 2003. The role of fungal spores in thunderstorm asthma. *Chest* 123(3): 745-750.

Garland, C.F., 2003. More on preventing skin cancer: sun avoidance will increase incidence of cancers overall. *British Medical Journal* 327: 1228.

Hajat, S., Kovats, R.S., Atkinson, R.W. and Haines, A., 2002. Impact of hot temperatures on death in London: a time series approach. *Journal of Epidemiology and Community Health* 56: 367-372.

Keatinge, W.R., 2002. Winter mortality and its causes. *International Journal of Circumpolar Health* 61: 292-299.

Sahu SK, Baffour B, Harper PR, Minty JH, Sarran C 2014. A hierarchical Bayesian model for improving short-term forecasting of hospital demand by including meteorological information. *J R Statist Soc A*. 177(1): 39-61.

Smith, R.W. and Nelson, D.R., 1998. Fractures and other injuries from falls after an ice storm. *American Journal of Emergency Medicine* 16(1): 52-55.

Venables, KM, Allitt, U, Collier, CG, Emberlin, J, Greig, JB et al (1997) “Thunderstorm-related asthma – the epidemic of 24/25 June 1994”. Clinical and Experimental Allergy 27(7): 725-736.

WMO (World Meteorological Organization) (2003). Extreme weather events might increase. July 2nd 2003. http://www.wmo.int/index-en.html. Accessed on 1st May 2012.