University of Hertfordshire UH

Citation for the published version:

Ordu, M., Demir, E., & Tofallis, C. (2019). A comprehensive modelling framework to forecast the demand for all hospital services. The International Journal of Health Planning and Management. DOI: 10.1002/hpm.2771

Document Version: Peer reviewed accepted Version

Title: A comprehensive modelling framework to forecast the demand for all hospital services

Short Running Title: A comprehensive modelling framework

Author List: Muhammed Ordu[1], Eren Demir[1], Chris Tofallis[1]

Corresponding Author: Muhammed Ordu, University of Hertfordshire, Hertfordshire Business School, AL10 9EU, Hatfield, United Kingdom, m.ordu@herts.ac.uk

Institutional Affiliations: [1] University of Hertfordshire, Hertfordshire Business School, AL10 9EU, Hatfield, United Kingdom

Abstract

Background: Due to increasing demand hospitals in England are currently under intense pressure resulting in shortages of beds, nurses, clinicians and equipment. To be able to effectively cope with this demand management needs to accurately find out how many patients are expected to use their services in the future. This applies not just to one service but for all hospital services.

Purpose: A forecasting modelling framework is developed for all hospital's acute services, including all specialties within outpatient and inpatient settings and the A&E department. The objective is to support the management to better deal with demand, and plan ahead effectively.

Methodology/Approach: Having established a theoretical framework, we used the national episodes statistics dataset to systematically capture demand for all specialties. Three popular forecasting methodologies, namely ARIMA, exponential smoothing and multiple linear regression were used. A fourth technique known as the seasonal and trend decomposition using loess function (STLF) was applied for the first time within the context of healthcare forecasting.

Results: According to goodness of fit and forecast accuracy measures, 64 best forecasting models and periods (daily, weekly or monthly forecasts) were selected out of 760 developed models, i.e., demand was forecasted for 38 outpatient specialties (first referrals and follow-ups), 25 inpatient specialties (elective and non-elective admissions), and for A&E.

Conclusion: This study has confirmed that the best demand estimates arise from different forecasting methods and forecasting periods (i.e. one size does not fit all). Despite the fact that the STLF method was applied for the first time, it outperformed traditional time series forecasting methods (i.e. ARIMA, exponential smoothing) for a number of specialties.

Practice Implications: Knowing the peaks and troughs of demand for an entire hospital will enable the management to: 1) effectively plan ahead, 2) ensure necessary resources are in place (e.g. beds, staff), 3) better manage budgets, ensuring enough cash is available, 4) reduces risk.

Keywords: Hospital demand, time series analysis, forecasting hospital services

1. Introduction

The National Health Service (NHS) in England has experienced considerably increased demand over the last 10 years. Approximately a 26% increase for accident and emergency (A&E) from 2006/07 to 2017/18 financial year, and the number of attendances and admissions to outpatient and inpatient specialties has increased by around 27% and 32%, respectively (National Health Services England, 2018a and 2018b).

Severe demand causes A&E departments to run under intense pressure which results in shortages of beds, nurses, clinicians and equipment. The increasing waiting times and length of stay experienced by A&E departments in the UK has negatively affected the daily functioning of A&E services. It is observed that waiting times and length of stay have been increasing and the 4-hour target (the percentage of patients spending 4 hours or more in hospital should be less than 5%) determined by the government has not been achieved since the financial year 2014-15 (National Health Services England, 2018a).

The increasing demand for hospital services beds makes it even more difficult to manage their limited number, not to mention the burden and additional workload on staff. Doctors and nurses are expected to treat more patients, thus putting patient safety at risk, which may further lead to inadequate discharges of patients. The key decision makers for each department will need to better understand the future demands to plan ahead for the needs of their local population. This includes planning for human resources, department expansion (or reduction), bed capacity requirements, and medical instrument needs.

In this respect many studies have been conducted using time series analysis to forecast patient demand. Gul and Celik (2018) carried out an exhaustive literature review and specified that the vast majority of studies are aimed at forecasting A&E departments, for example, Champion et al. (2007), Jones et al. (2008), Sun et al. (2009), Kam et al. (2010), Marcilio et al. (2013), Aboagye-Sarfo et al. (2015), Rosychuk et al. (2016), Juang et al. (2017) and Jiang et al. (2018). In addition, a number of studies used forecasting methods for different purposes. For example, Blaisdell et al. (2002) and Moustris et al. (2011) forecast hospitalizations for paediatric patients with asthma. Moreover, Joy and Jones (2005) was interested in forecasting hospital bed demands, whereas Toerper et al. (2016) predicted the required hospital beds for patients admitted from the cardiac catheterization laboratory. Schweigler et al. (2009) focused on estimating the bed occupancy rate in an A&E department, whereas length of stay was predicted by Hussey and Guo (2005) in a child residential treatment, Wrenn et al. (2005) in an emergency department, Garg et al. (2011) in a hospital, Levin et al. (2012) for a paediatric intensive care unit, Li et al. (2013) for Cholecystitis patients, Hachesu et al. (2013) for cardiac patients, Comber et al. (2014), Gul and Guneri (2015) in an A&E department, Papi et al. (2016) in a hospital and Barnes et al. (2016) for discharge prioritization.

A limited number of studies has also been carried out for hospital level demand forecasting using a wide range of methods (Boutsioli, 2010; Boutsioli, 2013). These hospital models are very high level, and are restrictive in that they assume that arrival patterns across all specialties are the same. We are therefore motivated in the present study to relax such assumptions in order to model individual aspects of system behaviour more accurately. This bespoke approach inevitably leads to more accurate forecasts of demand for each part of the hospital. In turn this will enable improved allocation of resources.

A typical NHS Trust in England provides services across 25 specialties, for example cardiology, respiratory medicine, ophthalmology, gynaecology, maternity, general surgery, and geriatric medicine. There are huge variations between these specialties providing services within outpatient and inpatient settings. The arrival pattern for first referrals (those coming in to outpatients for the first time) is hugely different to follow-ups (those coming in for routine check-ups). The same applies for inpatient elective and non-elective admissions.

Furthermore, what is the appropriate period of forecasting (daily, weekly, monthly or even hourly periods)? It might sound like a simple issue, but the choice of the period does matter and could potentially make a difference in forecast accuracies. A prediction period for one specialty (say in an outpatient setting) may produce accurate forecasts but can be ineffective for other specialties (or services). According to the literature, no study has examined this phenomenon at the entire hospital level (including inpatient, outpatient and A&E) to recommend key decision makers accordingly.

As a result, an accurate and reliable long term hospital forecasting modelling framework is required to evaluate and respond to the needs of local populations, both currently and in the future. This study will therefore enable hospital management to deal with the demand for A&E,

outpatient and inpatient services (for all specialties individually), thus providing the opportunity to effectively plan ahead for resource requirements (e.g. doctors and nurses).

The objectives and contributions of this study are as follows:

1) The development of a hospital level forecasting modelling framework (i.e. theory) for the management to adapt and use accordingly. To the best of our knowledge a study at this scale has never been conducted before.
2) The development of a procedure (and process) for the hospital management to determine the optimal periods (i.e. daily, weekly, and monthly) to deal with demand for outpatient, inpatient and A&E services.
3) The testing, validation and comparison of the most widely used forecasting methodologies for all specialties individually, with the aim of assisting hospital management to better plan for all hospital services.

2. Theory

Figure 1 illustrates the modelling framework we have developed as part of this study, which is made up of four steps to forecast demand for all hospital's services.

1) Extract daily, weekly and monthly activity for each specialty (or department) within outpatient (first referrals and follow-ups), inpatient (elective and non-elective) and A&E services. For instance, if the hospital provides services across 25 specialties, then a total of 303 extractions would need to be carried out, i.e., Outpatient = 25 specialties x 2 patient types (first referrals and follow-ups) x 3 data periods (daily, weekly and monthly) = 150; Inpatient = 25 specialties x 2 patient types (elective and non-elective) x 3 data periods = 150; A&E = 3 (i.e. 3 data periods only). Note that in this example 303 would be the maximum number of data extractions and in the majority of instances it will be lower, as some specialties may have a very small number of activities, e.g.,

within an Ophthalmology department there are usually few inpatient elective admissions, where the majority of patients would be treated in an outpatient setting and discharged on the same day, resulting in no data extractions. This category of patients can be bundled under the "Other" category.

2) For each of the extracted data, divide the data into training and validation sets. A typical rule of thumb is a 75/25 split, for example, if the entire data period is 33 months, then the first 25 months can be the training set and the remaining 8 months validation.
3) Develop a set of models for each of the extracted data using relevant forecasting techniques (and software). Assuming the analyst is comparing 4 forecasting methods for the 303 extracted data, then you would expect to develop 1,212 forecasting models in total.
4) A comparative analysis of the periods and the forecasting methods needs to be carried out to determine which forecasting method and period provides the most accurate predictions. This needs to be established for each of the extracted data using an appropriate forecasting accuracy measure, such as the mean absolute scaled error.

2.1 The Data

This study was carried out for the Princess Alexandra Hospital located in Harlow, England, working in collaboration with a number of managers, i.e., Director of Service Redesign and Transformation, Director of Planning, Director of Clinical Services, Director of Finance, Assistant Director of Finance, and Finance Consultant. We used the national Hospital Episode Statistics (HES) dataset. The HES data set contains personal, medical and administrative details for all patients admitted to, and treated in, NHS hospitals in England. There are more than 80 million records for each financial year. A financial year is from 1st April to 31st March the following year. The HES data set captures all the consultant episodes of patients during their stay in hospital. During a hospital stay, a patient might encounter several successive episodes, collectively known as a spell.

The data were provided in a txt format and the necessary steps were taken to import the data into Microsoft SQL Server version 12.0, so that database programming could be carried out to prepare the data for analysis. Initial checks were made to ensure that the data sets provided contained encrypted NHS numbers for matching purposes. The data period is from 01/04/10 to 31/03/13 (three financial years).

The total number of observations in the A&E dataset over the data period in England is 65 million records, with 15 million inpatient admissions, and 175 million outpatient attendances. We extracted all inpatient, outpatient and A&E data sets corresponding to the Princess Alexandra Hospital, which had 248,910 A&E arrivals, 996,134 outpatient attendances and 191,462 inpatient admissions over the data period. A breakdown of all activity for each specialty is illustrated in Table 1.

The inpatient and outpatient datasets were further partitioned into 27 distinct specialties to ensure all specialties within the hospital are considered as part of our forecasting modelling framework. For each specialty activity related data based on various periods - namely daily, weekly, and monthly inpatient admissions (elective/non-elective) and outpatient attendances (first referrals and follow-ups) - was extracted and analysed accordingly.

3. Method

Many forecasting methods have been used in hospital demand forecasting in the literature. The autoregressive integrated moving average (ARIMA), exponential smoothing (ES) and multiple linear regression were found to be the most widely used techniques (Ordu et al., 2019). A new method, known as the seasonal and trend decomposition using loess (STL) developed by Hyndman and Athanasopoulos (2014) was also tested, which has never been applied within the healthcare context before. STL separates the time series datasets into seasons and trends, a typical phenomenon experienced by most healthcare providers and services. The STL function (STLF) could therefore be effective in forecasting demand for hospital services and provide new insights. Our proposed approach in Figure 1 will therefore be tested using these four forecasting techniques.

This study was accepted and approved by the Ethics Committees with Delegated Authority in the University of Hertfordshire.

ARIMA

The autoregressive integrated moving average (ARIMA) method is a forecasting technique that generates forecasts by means of autocorrelations in the time series (Hyndman and Athanasopoulos, 2014). The ARIMA method has three parameters (p, d and q) where p denotes the order of autoregression, d is the order of differencing and q is the order of the moving average (DeLurgio, 1998,). Firstly, stationarity analysis is carried out (in the data preparation section). In model selection, the values of AR (p) and MA (q) in an ARIMA method are determined.

The AR (p) is the autoregression, which implies that Y_t depends on previous observations in the data. The MA (q) is the moving average, which relates Y_t to errors for previous terms in the data. In the estimation section, computer programs automatically compute the best parameters of ARIMA models. To do this, it generates the initial values of the parameters and the best model is found iteratively. The method of least squares or maximum likelihood estimation is used in order to determine an ARIMA model. In this study, the auto.arima () package in R developed by Hyndman and Khandakar (2008) is applied in order to select the best ARIMA models. This function uses the Akaike's Information Criterion (AIC) as a criterion in order to find the orders of AR (p) and MA (q) in ARIMA models. This function runs iteratively until the lowest AIC is found. The diagnostics section shows whether the ARIMA model selected in the estimation section is an appropriate model or not by using the portmanteau test, which examines whether the residuals of the model is white noise or not. If the model is not adequate, a new model should be selected (Makridakis et al., 1998). The estimated values are produced by the selected and verified ARIMA model in the forecasting step. At the final step, forecast accuracy is calculated.

Exponential Smoothing

In exponential smoothing (ES) the closest observation has more impact in forecasting the next period compared to the furthest observation, i.e., weights exponentially decrease as the observations get older (Makridakis et al., 1998).

There are 15 types of ES model if the error terms are not taken into account. There are in total 30 ES models with the additive error and the multiplicative error for each model (Hyndman, Koehler, Ord and Snyder, 2008). An exponential smoothing method is selected. The data is trained and then a validation process is carried out. This process is repeated until the all ES methods are applied. Using the ets() function in R developed by Hyndman and Khandakar (2008), the best ES model is selected taking into account the lowest AIC value (Hyndman and Khandakar, 2008).

Multiple Linear Regression

Multiple linear regression (MLR) seeks to find a relationship between independent (explanatory) and dependent variables. In other words, one variable is forecast using two or more independent variables as can be seen in the equation below (Makridakis, Wheelwright and Hyndman, 1998, p. 241).

$$Y = b_0 + b_1X_1 + b_2X_2 + \cdots + b_kX_k + e$$

Y is the dependent variable, X_k are explanatory variables, b_0 is a constant term, b_k is the coefficient of the explanatory variable and e is the error term. The parameters are estimated using the least squares method in order to obtain the minimum sum of squares of the residuals. (Makridakis, Wheelwright and Hyndman, 1998, p. 250). After estimating the coefficients of the model, F-Test, p-value, R^2 and adjusted R^2 are calculated in order to better understand whether the regression model is significant or not.

In the literature, the existing studies applying a MLR method have used dummy variables (e.g. days of the week, months of the year) for independent variables. For example, the stepwise linear regression model for the daily estimation includes days of the week, months of the year, and variables related to holidays.

STLF

The STL method is a reliable decomposition method and uses the loess (i.e. locally estimated scatterplot smoothing) to decompose the time series (Hyndman and Athanasopoulos, 2014).

The STLF method is a forecasting technique which uses a non-seasonal forecasting method after decomposition of the time series using the STL method. To decompose the time series, the following equation is used for an additive decomposition:

$$y_t = \hat{S}_t + \hat{A}_t$$

where $\hat{A}_t = \hat{T}_t + \hat{E}_t$, $\hat{A}_t$ is the seasonally adjusted component at period t, $\hat{T}_t$ is the trend-cycle component at period t, and $\hat{E}_t$ is the error component at period t (Hyndman and Athanasopoulos, 2014). To decompose the time series, the following equation is used for a multiplicative decomposition:

$$y_t = \hat{S}_t \hat{A}_t$$

where $\hat{A}_t = \hat{T}_t \hat{E}_t$, $\hat{A}_t$ is the seasonally adjusted component at period t, $\hat{T}_t$ is the trend-cycle component at period t, and $\hat{E}_t$ is the error component at period t (Hyndman and Athanasopoulos, 2014).

After that, a non-seasonal forecasting technique (e.g. non-seasonal ARIMA method or Holt's method) is used to estimate the time series. The estimated values are then re-seasonalized by using "the last year of the seasonal component" (Hyndman, 2016). In this study, the stlf () package in R developed by Hyndman and Khandakar (2008) is applied in order to select the best STLF models.

3.1. Choosing the best forecasting methods and periods

Forecast accuracy refers to the goodness of fit of the developed forecasting model for the out of sample data. The important issue is to select the 'best' forecasting method. A number of metrics are available for this purpose. We have chosen to use the mean absolute scaled error (MASE) method which has the advantage that if zero occurs in the observations, MASE avoids the infinities which occur with mean absolute percentage error (MAPE) (Hyndman and Koehler, 2006). MASE is based on a simple quantity that managers can comprehend, namely the average prediction error (irrespective of sign). MASE is a ratio which compares this with the corresponding value from using the naïve forecasting method as a benchmark. In the MASE, the numerator is the mean absolute error of the forecasting method and the denominator is the mean absolute error of the naïve method, i.e. when the forecast is the previous observation. The denominator is therefore the same for all methods studied. Hence, the MASE compares the errors with those from the naïve method.

4. Results

Using the modelling framework illustrated in Figure 1, we developed forecasting models for all the hospital services. A total of 768 forecasting models were developed, 12 for A&E, 456 for outpatient, and 300 for inpatient services (details are provided below). This process can be extremely laborious, repetitive and time consuming, where an analyst could easily make a mistake. To avoid this a programme was developed in R to automate the whole process. The R code extracts relevant data (from the prepared and structured big data), develops the forecasting model using each of the four methods for each specialty, then selects the best method that produces the lowest MASE value.

A&E

The average number of A&E admissions per year at Princess Alexandra Hospital is 82,535 patients over the study period. The A&E is a 22 bed department operating 24/7, treating on average 227 patients per day (standard deviation [SD] ± 28.71). Figure 2 shows the percentage changes in the number of A&E admissions by comparing one month against the same month in the previous year. Admissions showed increases in 3-year data (i.e. 2010 to 2013), despite observing a decrease in January, March and July. For example, the demand increased by 3.26%, 4.27% and 6.59% in October 2010 to 2011, 2011 to 2012 and 2012 to 2013, respectively. We also observe a year on year increase in A&E admissions, for example, the demand increased by 2.08%, 11.39% and 1.09% in November 2010 to 2011, 2011 to 2012 and 2012 to 2013, respectively, and by 14.91% in the same month 2010 to 2013. This figure illustrates some evidence of the winter demand pressures experienced by PAH A&E services (particularly in October, November, December and February).

We used all the four methods under three different periods, i.e. daily, weekly and monthly. Hence 12 methods were developed to forecast A&E demand. The MLR produced the lowest MASE for monthly data for both the training and the validation sets. According to the adjusted R^2 measure, monthly MLR produced the best goodness of fit, with approximately 60% of the variation in demand estimates explained using 7 independent variables. Amongst the variables, "Month" and "Holiday" (e.g. bank holidays, New Year's Day) variables were found to have the highest impact in forecasting demand. The significant variables ($p < 0.05$) were January, February, March, April, August, and Holiday variables.

The estimated coefficient for January was found to be negative (i.e. -581.64), meaning that A&E activity in January is decreasing, whereas the holiday variable coefficient was positive (1519) where demand for A&E services during the holiday period is increasing. This may seem like a surprise, but according to Figure 2 year on year A&E admissions at Princess Alexandra Hospital in January are decreasing, thus our findings are in agreement with the dataset. As expected the holiday variables are significant, meaning closure of primary care services (e.g. general practitioners) during the holiday season has a negative impact on A&E services.

Outpatient services

The average number of first and follow up attendances per year at PAH is around 69528 and 155952 patients, respectively. We notice a large variation at the specialty level, where trauma and orthopaedics (average of 213 patients per week) and General surgery (average 140 per week) has the highest number of first referrals, followed by Gynaecology (132/week), Ophthalmology (106/week), Ear, Nose and Throat (94/week), and General medicine (78/week). Clinical oncology experiences the lowest number of first referrals (16/week).

Harlow, the town where Princess Alexandra Hospital is located has a high birth rate and it is no surprise that the Obstetrics specialty has the highest number of weekly follow ups (507/week), followed by Trauma & Orthopaedics (430/week), Ophthalmology (329/week), Anaesthetics (250/week), and Gynaecology (198/week).

A total of 456 forecasting models were developed across 19 outpatient specialities at PAH (19 specialties x 2 (first and follow up referrals) x 3 periods (daily, weekly and monthly) x 4 forecasting methods (ARIMA, ES, MLR and STLF), hence 456. Table 2 (top half) summarizes the best forecasting models and periods selected for both first and follow up referrals for all the 19 outpatient specialties.

According to Table 2, multiple linear regression (MLR), forecasting on a daily period produced the lowest MASE value (similarly for follow-up referrals too). The best forecasting methods for all outpatient specialties are summarized below:

- 11 MLR for first referrals and 7 MLR for follow up referrals
- 1 ARIMA for first referrals and 6 ARIMA for follow up referrals
- 3 ES for first referrals and 4 ES for follow up referrals
- 4 STLF for first referrals and 2 STLF for follow up referrals

The best forecasting periods (see Table 2) for all outpatient specialties are summarized below:

- 14 daily estimations for first referrals and 12 daily estimations for follow up referrals
- 1 weekly estimation for first referrals and 3 weekly estimations for follow up referrals
- 4 monthly estimations for first referrals and 4 monthly estimations for follow up referrals

Inpatient services

Over the study period the average number of elective and non-elective admissions per year at PAH is around 30,588 and 33,132 patients, respectively. Trauma and orthopaedics (71/week), General surgery (70/week), urology (69/week) and clinical haematology (53/week) has the highest number of average inpatient elective (i.e. booked) admissions, with the Paediatrics specialty the lowest (9/week). Obstetrics (155/week), Geriatric Medicine (136/week) and General Medicine (124), followed by General Surgery (66) have the highest number of non-elective admissions, whereas Others has the lowest (17 per week).

According to the bottom half of Table 2 (inpatient forecasting results) the multiple linear regression method produced the most accurate forecasts for elective inpatient admissions, whereas ARIMA for non-electives. A total of 192 forecasting models was developed for elective admissions (across 16 specialties) and 108 models for non-elective admissions (across

9 specialties). Note that not all specialties within inpatient services had patient admissions. For example, there are only 16 specialties with patient activity in elective inpatient services, and 9 specialties for non-electives (out of a total of 27 specialties). Where a specialty has a very small number of patient activity, it is categorised within the "Others" group.

The findings can be summarised as follows:

- 8 MLR for elective admissions and 1 MLR for non-elective admissions
- 5 ARIMA methods for elective and 6 ARIMA for non-elective
- 3 ES methods for elective admissions and 1 ES method for non-elective admissions
- 0 STLF method for elective admissions and 1 STLF method for non-elective admissions

In terms of forecasting periods the findings are as follows:

- 8 daily estimations for elective admissions and 5 daily estimations for non-electives
- 5 weekly estimations for electives and 1 weekly estimation for non-electives
- 3 monthly estimations for electives and 4 monthly estimations for non-electives

The findings suggests that the best demand estimates are based on different forecasting methods and periods. In total, 64 best forecasting models are selected out of the 768 models, i.e., 38 forecast demand for outpatient specialties, 25 for inpatient specialties, and 1 for A&E.

Validation of the predicted demand for all the hospital services

We compared the best forecasting result and actual value by using a paired *t* test which is determined as a formula in the equation below,

$$t_0 = \frac{\bar{d} - \mu_d}{S_d/\sqrt{K}}$$

where $\bar{d}$ denotes average observed differences between actual values and predictions; μ_d is the mean difference; S_d denotes the standard deviation and K is the number of input data set (Banks et al., 2005, p. 377). The *t* test value ($|t_0|$) is compared against the critical value, and if it is less than or equal to t critical values ($t_{\alpha/2,K-1}$) at the 95% significance level, then the predicted demand is validated. Thus we would endorse the use of the model. This process was repeated for each of the selected 64 models and we confirm that in all cases the *t* test value was less than the critical value.

We also carried out a visual inspection of the actual activity against the predictions. Figure 3 illustrates the actual values vs. the estimated demand for inpatient elective and non-elective admissions for each specialty. After a visual inspection there is a very close agreement between the actual values and the estimates. In Figure 3 some of the specialties have four bars and some two bars. For example, General Surgery has a large amount of patient activity as elective and non-elective admissions, therefore necessary forecasting models were developed and compared accordingly, whereas Urology has very few non-elective admissions, thus no models were developed.

5. Discussion

The lack of accurate demand estimates in a modern healthcare system is one of the main reasons why hospitals are unable to effectively plan ahead, negatively affecting hospital performance, e.g. increased waiting lists and waiting times. Improving the forecasting of future demand is essential for the planning and delivery of healthcare services, particularly in relation to supporting better decision making on service needs.

For the hospital management, forecasting demand for a single department, single disease, specialty, or a service (e.g. A&E) is inadequate. Understanding the demand for the entire hospital is what the management needs and without a sound theoretical framework (or a process to do so), the data analyst could easily end up generating inaccurate forecasts. For instance, the NHS heavily relies on the use of Excel spreadsheets to forecast demand. Without a systematic approach (i.e. the theoretical framework described above) and range of methods to test with, it is inevitable that demand estimates will be both inaccurate and misleading.

As a result of this major gap in the literature and the desperate need of hospital service managers, we developed a forecasting modelling framework where hospitals could easily adapt and capture accurate demands for individual services at a sufficient level of detail. The framework includes a step by step guide as to how the data should be extracted and prepared using various periods; how the data should be divided into training and validation samples; how the forecasting models should be developed using appropriate software for individual specialties (or departments and diseases), and finally how to select the optimal model for each specialty.

This framework is tested by extracting the entire activity of a mid-size NHS Trust in England using the national hospital episodes statistics dataset. According to the goodness of fit and forecast accuracy measure MASE, the best 64 forecasting models and periods are selected out of 760 developed models, i.e., 38 forecasted demand for outpatient specialties (first referrals and follow-ups), 25 for inpatient specialties (elective and non-elective admissions), and 1 for A&E.

This study indicates that forecasting hospital demand under different periods will generate more accurate results. This study has confirmed that the best demand estimates are based on different forecasting methods and forecasting periods (i.e. one size does not fit all). In addition, we also tested with a forecasting methodology that has never been considered within the healthcare setting before (i.e. the STLF method). It proves that the STLF method outperformed traditional time series forecasting methods (i.e. ARIMA, exponential smoothing) for a number of specialties, despite the method being applied for the first time. For instance, according to MASE, the STLF method was superior for the non-elective admissions in the general medicine specialty. In addition, this study highlights that hospital managements will need to take into consideration different forecasting periods to better estimate hospital demands.

No modelling framework (or a forecasting exercise) is ever perfect and the results need to be interpreted cautiously. For example, the average adjusted R squared estimates for MLR models is around 0.6, i.e., 60% of the variation is explained by the significant variables. This suggests that there is plenty of room for improvement and additional variables (e.g. weather variables from the meteorological office) can be obtained to enhance the predictive power of the models.

In an environment with severe financial constraints, coupled with increasing demand for acute services, there is a greater emphasis and need for evidence based decision making models. Reliable and accurate forecasting frameworks are welcomed by hospital service managers to better understand the needs of their local populations.

6. Practice Implications

In this rapidly changing environment forecasting healthcare demand plays an essential role for hospital service managers, so that they are able to plan and implement relevant strategies accordingly. Every decision made by the hospital manager - whether to acquire new beds, equipment, expansion of the service, or a business plan to set up a new service - hinge on the accuracy of the forecast demand. Reliable forecasts enable key decision makers to foresee and prepare for future challenges.

Given the amount of data generated by a mid-size hospital in England (in the region of thousands of observations per day), it is a challenge to store, extract, prepare and develop a forecasting modelling framework that covers individual services for an entire hospital.

This research was conducted with a mid-size NHS Trust in England, in collaboration with Directors of Finance, Strategy and Planning, and Turnaround Director. When a hospital is in financial trouble (e.g. in serious debt) a Turnaround director is usually appointed for a short period of time (e.g. 1 year) to turn the hospital around. At the time of this research, the hospital had little idea about its demand for all its various services. This modelling framework was developed and forecasts were generated for individual services, where the management were able to use these estimates to allocate resources (e.g. staff, beds, budget, etc.) appropriately.

There are numerous findings and practical implications of this research for the hospital:

1) The demand estimates for all the hospital services enabled the management to schedule staff (particularly consultant and nurses) and beds for all services rather than just focusing on a single service (or department and disease). This helped the management to observe trends, seasonality and busy periods (e.g. A&E), and ensure that an adequate supply of equipment, staff and other needs are in place throughout the year.
2) A significant improvement on the staff management side of operation. For example, the forecast activity for most specialties in inpatient admissions were fairly stable, with a small year on year increase. But a dramatic increase in maternity admissions was observed (Obstetrics specialty). This suggested that with the current level of staffing the service will not be able to cope with this expected demand over the next 12 months. Therefore, additional staff (made up of midwives and nurses) are essential to ensure the service functions appropriately. If the forecasting had not been carried out for each of the individual services, it wouldn't have been possible to capture this significant trend.
3) Most NHS hospitals struggle with their finances, either due to poor management and/or budget cuts from the Department of Health. Knowing the peaks and troughs of demand for the entire hospital will enable the management to better manage their budgets, ensuring enough cash is available; if not, the hospital can find itself in arrears. Poor cash flow can have severe consequences, creating a risk situation for patients, staff, and the NHS.
4) Updating the modelling framework with new data on a regular basis (e.g. quarterly) will enable the management to judge whether there is likely to be a surge of patients in

the near future (just as in the maternity example above). It thus offers the opportunity to provide better care for patients at the point of need within existing budgets.

References

1. Aboagye-Sarfo P, Mai Q, Sanfilippo FM, Preen DB, Stewart LM, Fatovich DM. A comparison of multivariate and univariate time series approaches to modelling and forecasting emergency department demand in Western Australia. Journal of Biomedical Informatics 2015;57,62-73.
2. Barnes S, Hamrock E, Toerper M, Siddique S, Levin S. Real-time prediction of inpatient length of stay for discharge prioritization. Journal of the American Medical Informatics Association 2016;23,e2–e10.
3. Blaisdell CJ, Weiss SR, Kimes DS, Levine ER, Myers M, Timmins S, Bollinger ME. Using seasonal variations in asthma hospitalizations in children to predict hospitalization frequency. The Journal of Asthma: Official Journal of the Association for the Care of Asthma 2002;39(7),567-575.
4. Boutsioli Z. Forecasting the stochastic demand for inpatient care: the case of Greek national health system. Health Services Management Research 2010;23,116-120.
5. Boutsioli Z. Estimation of unpredictable hospital demand variations in two Piraeus public hospitals, Greece. Journal of Hospital Administration 2013;2,126-137.
6. Champion R, Kinsman LD, Lee GA, Masman KA, May EA, Mills TM, Taylor MD, Thomas PR, Williams RJ. Forecasting emergency department presentations. Australian Health Review 2007;31,83-90.
7. Combes C, Kadri F, Chaabane S. Predicting hospital length of stay using regression models: Application to emergency department. Paper presented at: 10th International Conference on Modeling, Optimization & Simulation; November 5-7, 2014, Nancy, France.

8. DeLurgio SA. Forecasting Principles and Applications. New York: McGraw – Hill; 1998.
9. Garg B, Sufyan Beg MM, Ansari AQ, Imran BM. Soft computing model to predict length of stay of patient. In: S. Dua, S. Sahni, & D. P. Goyal (eds.). Paper presented at: 5th International Conference on Information Intelligence, Systems, Technology and Management; March 10-12, 2011, Gurgaon, India.
10. Gul M, Celik E. An exhaustive review and analysis on applications of statistical forecasting in hospital emergency departments. Health Systems 2018.
11. Gul M, Guneri AF. Forecasting patient length of stay in emergency department by artificial neural networks. Journal of Aeronautics and Space Technologies 2015;8(2),43–48.
12. Hachesu PR, Ahmadi M, Alizadeh S, Sadoughi F. Use of Data Mining Techniques to Determine and Predict Length of Stay of Cardiac Patients. Healthcare Informatics Research 2013;19(2),121–129.
13. Hussey DL, Guo S. Forecasting length of stay in child residential treatment. Child Psychiatry and Human Development 2005;36(1),95–111.
14. Hyndman RJ. *Package 'forecast'*. https://cran.r-project.org/web/packages/forecast/forecast.pdf. Published January 18, 2019. Accessed May 26, 2017.
15. Hyndman RJ, Athanasopoulos G. Forecasting Principles and Practice. Australia: Otexts; 2014.
16. Hyndman RJ, Khandakar Y. Automatic Time Series Forecasting: The forecast Package for R. Journal of Statistical Software 2008;27.
17. Hyndman RJ, Koehler AB. Another look at measures of forecast accuracy. International Journal of Forecasting 2006;22,679-688.
18. Hyndman RJ, Koehler AB, Ord JK, Snyder RD. Forecasting with Exponential Smoothing: The State Space Approach. Berlin, Germany: Springer; 2008.

19. Jiang S, Chin KS, Tsui KL. A universal deep learning approach for modelling the flow of patients under different severities. Computer Methods and Programs in Biomedicine 2018;154,191-203.
20. Jones SS, Thomas A, Evans RS, Welch SJ, Haug PJ, Snow GL. Forecasting daily patient volumes in the emergency department. Academic Emergency Medicine 2008;15,159-170.
21. Joy MP, Jones S. Predicting bed demand in a hospital using neural networks and ARIMA models: A hybrid approach. Paper presented at: European Symposium on Artificial Neural Networks, April 27-29, 2005, Bruges, Belgium.
22. Juang WC, Huang SJ, Huang FD, Cheng PW, Wann SR. Application of time series analysis in modelling and forecasting emergency department visits in a medical centre in Southern Taiwan. BMJ open 2017;7(11).
23. Kam HJ, Sung JO, Park RW. Prediction of daily patient numbers for a regional emergency medical center using time series analysis. Healthcare Informatics Research 2010;16,158-165.
24. Levin SR, Harley ET, Fackler JC, Lehmann CU, Custer JW, France D, Zeger SL. Real-time forecasting of pediatric intensive care unit length of stay using computerized provider orders. Critical Care Medicine, 2012;40(11),3058–3064.
25. Li JS, Tian Y, Liu YF, Shu T, Liang MH. Applying a BP neural network model to predict the length of hospital stay. Paper presented at: 2nd International Conference on Health Information Science, March 25-27; 2017, London, UK.
26. Makridakis S, Wheelwright SC, Hyndman RJ. Forecasting Methods and Applications. New York, USA: John Wiley & Sons; 1998.
27. Marcilio I, Hajat S, Gouveia N. Forecasting daily emergency department visits using calendar variables and ambient temperature readings. Academic Emergency Medicine 2013;20,769-777.

28. Moustris KP, Douros K, Nastos PT, Larissi IK, Anthracopoulos MB, Paliatsos AG, Priftis KN. Seven-days-ahead forecasting of childhood asthma admissions using artificial neural networks in Athens, Greece. International Journal of Environmental Health Research 2011;22(2),93-104.
29. National Health Services England. (2018a). *A&E Attendances and Emergency Admissions and Bed Availability and Occupancy*. Retrieved June 20, 2018, from https://www.england.nhs.uk/statistics/statistical-work-areas/ae-waiting-times-and-activity/
30. National Health Services England. (2018b). *NHS Inpatient Elective Admission Events and Outpatient Referrals and Attendances*. Retrieved June 20, 2018, from https://www.england.nhs.uk/statistics/statistical-work-areas/hospital-activity/quarterly-hospital-activity/qar-data/
31. Ordu M, Demir E, Tofallis C. A decision support system for demand and capacity modelling of an accident and emergency department. Health Systems 2019.
32. Papi M, Pontecorvi L, Setola R. A new model for the length of stay of hospital patients. Health Care Management Science 2016;19(1),58–65.
33. Ripley B, Venables B, Bates DM, Hornik K, Gebhardt A, Firth D. *Package 'MASS'*. https://cran.r-project.org/web/packages/MASS/MASS.pdf. Published November 1, 2018. Accessed December 26, 2018.
34. Rosychuk RJ, Youngson E, Rowe BH. Presentations to emergency departments for COPD: A time series analysis. Canadian Respiratory Journal 2016;1-9.
35. Schweigler LM, Desmond JS, McCarthy ML, Bukowski KJ, Ionides EL, Younger JG. Forecasting models of emergency department crowding. Academic Emergency Medicine 2009;16(4),301–308.

36. Sun Y, Heng BH, Seow YT, Seow E. Forecasting daily attendances at an emergency department to aid resource planning. BMC Emergency Medicine 2009;9(1).

37. Toerper MF, Flanagan E, Siddique S, Appelbaum J, Kasper EK, Levin S. Cardiac catheterization laboratory inpatient forecast tool: a prospective evaluation. Journal of the American Medical Informatics Association 2016;23,e49–e57.

38. Wrenn J, Jones I, Lanaghan K, Congdon CB, Aronsky D. Estimating Patient's Length of Stay in the Emergency Department with an Artificial Neural Network. Paper presented at: *AMIA Annual Symposium; 2005, Washington, USA*.