Short term hospital occupancy prediction

Article *in* Health Care Management Science · March 2007

DOI: 10.1007/s10729-006-9000-9 · Source: PubMed

Mark W Isken
Oakland University

Health Care Manage Sci (2007) 10:47–66
DOI 10.1007/s10729-006-9000-9

Short term hospital occupancy prediction

Steven J. Littig · Mark W. Isken

Received: 11 January 2006 / Accepted: 9 October 2006 / Published online: 28 November 2006

Abstract Inpatient census, or occupancy, is a primary driver of resource use in hospitals. Fluctuations in occupancy complicate decisions related to staffing, bed management, ambulance diversions, and may ultimately impact both quality of patient care and nursing job satisfaction. We describe our approach in building a computerized model to provide short-term occupancy predictions for an entire hospital by nursing unit and shift. Our model is a comprehensive system built using real hospital data and utilizes statistical predictions at the individual patient level. We discuss the results of piloting an early version of the model at a mid-size community hospital. The primary focus of the paper is on the development and methodology of a second generation of the predictive occupancy model. The results and accuracy of this new model is compared to a variety of other predictive methods based on tests using 2 years of actual hospital data.

Keywords Hospital occupancy · Forecasting · Census prediction · Patient flow · Decision support system

1 Introduction

"Is there a way for us to anticipate today how busy we will be tomorrow?" That question, posed to us by an administrator at a hospital we will call Study Hospital was the genesis for the project described here—to build a comprehensive patient level model and data based system for providing short-term occupancy forecasts by nursing unit, by shift, for an entire hospital. Inpatient census, or occupancy, is a primary driver of resource use in hospitals. Appropriate nurse staffing levels are directly affected by census along with the level of care required by each patient. Patients also generate demand for a myriad of ancillary services including laboratory, pharmacy, physical therapy, radiology, housekeeping and surgical services. Improved short-term information has the potential to improve the matching of hospital resources to fluctuating demand levels for a number of hospital business processes such as patient placement and admission control policies [1], managing ancillary services [2], reducing ambulance diversions [3–5], reducing periods of under and over staffing [6], and management of inpatient bed capacity [7].

Hospital occupancy is impacted by a number of patient input streams and patient care processes. Previous studies have looked at trying to predict hospital length of stay for certain patient populations or to predict emergency room arrivals. This work differs from previous research in that we attempt to address the problem of whole hospital census prediction by simultaneously modeling all streams affecting patient occupancy. Furthermore, we were forced to tackle many of the thorny issues involving the realities of hospital information systems in order to implement an early version of this model at the Study Hospital.

Most hospitals manage occupancy with rudimentary census snapshot reports along with ad-hoc use of data from multiple information systems to augment their managerial intuition. In this era of electronic medical records, fast and inexpensive computing, and widespread availability of business intelligence platforms, more sophisticated computerized models are feasible and would help hospital

S. J. Littig
Improvement Path Systems, Inc., Farmington Hills, MI, USA
e-mail: littig@improvementpath.com

M. W. Isken
Department of Decision and Information Sciences,
Oakland University, Rochester, MI, USA

M. W. Isken (✉)
School of Business Administration, Oakland University, 317
Elliott Hall, Rochester, MI 48309, USA
e-mail: isken@oakland.edu

administrators better plan for ways to meet patient demand. Fulfilling this objective requires spanning the disciplines of healthcare informatics, information systems and management science. If management science based models are to be effectively implemented in hospitals, integration of modeling and information systems is almost unavoidable.

The objective for this paper is to lay out the structure of our methodology, present data and modeling details and report on preliminary results based on actual use of this approach in a mid-size community hospital. Both the breadth and depth of this modeling application are rather ambitious and we wish to give the reader some sense of this rather than focusing in great detail on any one aspect of this project. The uniqueness of this work is in its comprehensive scope in conjunction with patient level detailed modeling. Early versions of this approach were in use at Study Hospital while the research reported here was ongoing. We will discuss the numerous subproblems and challenges related to data acquisition, data processing, model building, and model maintenance. Much work remains to be done to make hospital wide real time census prediction a reality in today's complex healthcare delivery systems, however, we believe the potential benefits are significant. Sections 2 and 3 discuss the problem and review related work on hospital occupancy prediction. Section 4 provides a high level description of our approach to occupancy prediction. Sections 5 and 6 discuss the two main developmental efforts: the predictive occupancy database (POD) and predictive occupancy model (POM), respectively. Model accuracy is discussed in Section 7 and error analysis in Section 8. The paper concludes with a summary of outstanding research and development issues.

2 Background

There are several challenges in managing hospitals in the United States including a widespread nursing shortage, constrained hospital reimbursement rates, an aging population requiring ever more medical resources, increasing expectations on the quality of care, and an inherently labor-centric business that makes it difficult to use technology to improve productivity. Nursing shortages and job dissatisfaction are widespread [8, 9]. Medical errors are more likely to occur when staff shortages spread nurses too thin [10, 11]. Emergency departments are increasingly required to shut their doors and divert patients due to full census and bed management problems [12, 13]. Great efforts have been made, and continue to be made, to computerize and digitize patient records and process-related data in hospitals. Needed are practical decision support tools that leverage this information to provide hospital management and clinical providers with relevant detailed operational statistics and analysis.

Every shift of every day within a hospital requires managers to make decisions regarding staff allocation. In addition, nursing and registration staff must make decisions regarding patient placement and transfers throughout the day, diversion status of the emergency department and scheduling of inpatient surgical cases. Each decision implicitly requires the manager to make assumptions regarding the expected census for some unit or group of units in the short-term future. Predicting patient occupancy is difficult due to significant fluctuations in occupancy from day to day and week to week. Short-term predictions today are virtually always made through human intuition supplemented by available status reports from standard hospital information systems. The quality of information is fairly low and requires users to assimilate many divergent pieces of information to attempt to estimate future occupancy patterns. Given that a 200-bed hospital with an active emergency department might average 600 admission/discharge/transfer (ADT) events per day, it is not surprising that manually integrating this information is a very difficult task. Since accurate short-term predictions are difficult to come by within the hospital, most decisions affected by future occupancy are delayed until the solution is obvious. For example, occupancy on the weekends generally drops from Friday night through Monday morning. Often, this provides an opportunity to close a unit or flex staffing downward to meet demand. Because of the uncertainty in demand however, nurse managers are often unwilling to take the risk in closing beds until it is obviously safe to do so. Conversely, expensive on-call agency or contract nurses will not be called in until it is clear that there is an active bed shortage. More proactive use of this nursing pool could help accelerate nursing care to facilitate discharges and avoid the bed shortage in the first place.

2.1 Study hospital

Study Hospital (SH) consists of 226 licensed beds in 13 units. The scope of the study currently includes 11 units of the facility representing 211 licensed beds and an additional 39 short-stay beds. The 15 beds not modeled in this Phase 1 study were on the Pediatric and Special Care Nursery units both of which have unique and low utilization. These 11 units were all modeled individually and represent the main inpatient units in the hospital: two critical care units, two surgical units, two medical units, two labor/delivery units, one short-stay unit, and one telemetry unit. A total of 14 unit groups (overlapping) were defined based on various combinations of the 11 units to summarize certain bed types and to represent the way in which beds are used within the hospital. For example, 'All Surgical Beds' is made up of 2 of the 11 units, 'All Main Inpatient Beds' is made up of the seven primary units serving the inpatients. This created a total of 25 distinct units and unit groups representing patient beds at the hospital that were modeled.

SH is a mid-sized community hospital with a typical mix of patient admission sources: surgical, emergency, and direct admissions. We classify a 'direct admission' to be any patient who was not admitted to the hospital via a surgical procedure or through the emergency department. This catch all category includes patients who are admitted: directly from their doctor's office, or kept overnight after an scheduled outpatient procedure, or transferred from another hospital, etc. For the baseline study period SH had about 23,300 admissions to a licensed bed. There were a total of approximately 56,500 visits to the emergency department of which about 11,800 were admitted to an inpatient bed. There were approximately 17,300 surgical cases which generated about 5,800 admissions to beds. The remaining 5,700 admissions were direct admits.

Although SH is frequently at very high occupancy, occupancy still fluctuates over a wide range. During a 1-year analysis period, occupancy ranged from 0 open beds to 52 open beds on their 127-bed medical/surgical units. During a single week it is not uncommon to see occupancy change by 30–40 beds from the low period of the week to the high period. The *90% width* of the occupancy range is the difference between the 95th percentile of occupancy and the 5th percentile of occupancy on that unit or unit group. This 90% width can be interpreted as the range of occupancy that administrators must manage to account for the occupancy present on 90% of the shifts. Therefore 10% of the shifts (or about 2 of the 21 shifts a week) will typically be outside this range. Shifts are defined to be 8 h long starting at midnight, 8 A.M., and 4 P.M.. Table 1 shows this statistic for several groups of beds at SH.

3 Review of occupancy prediction research

Predicting hospital occupancy has been and continues to be a popular topic of research. There have been several studies that have explored various time-series based forecasting models for hospital admissions and occupancy [14–22]. The motivation for much of this work has been longer-term bed allocation decisions or to characterize admission volumes or patterns. A number of researchers have presented various methodologies for prediction of discharges and hospital length of stay (LOS). Barie, Hydo, and Fischer [23] explore prediction of prolonged length of stay in a surgical intensive care unit (ICU) for critically ill surgical patients using various patient classification, or scoring, methods. Mounsey et al [24] explored predictors of LOS in ICU following coronary artery surgery to predict "fast track" patients requiring and ICU bed for less than 24 h. Walczak, Pofahl, and Scorpio [25] report the first steps toward a comprehensive LOS prediction model for purposes of supporting resource allocation decisions. They developed neural network models to predict hospital LOS for pediatric trauma patients and patients with acute pancreatitis. One goal of their work was to use variables that are easily available within 10 min of presentation in the emergency department (ED). Pofahl et al [2] used a neural network to predict LOS>7 days for patients with acute pancreatitis using demographic and clinical patient characteristics. Buchman et al [26] compared logistic regression to neural networks for predicting chronicity (defined as LOS>7 days) in an SICU using detailed clinical data. The neural networks outperformed logistic regression on a variety of bases. Zernikow et al [27] used multiple linear regression and neural networks to predict LOS in preterm neonates. They used 14 variables representing first day of life data and found very similar performance between neural networks and regression. Izenberg [28] used neural networks to predict trauma mortality based on data from patient presentation in the ED. They had very promising results with 91% accuracy on close to 300 patients with 78% sensitivity for 49 deaths. Toumpoulis et al [29] showed that both the linear and logistic regression forms of the Euro-SCORE could be used to predict pos-operative length of stay for CABG patients.

Table 1 Summary of underlying occupancy variation by hospital area

Area	Occupancy by Shift				
	Occupancy by Shift		Range of Avg Occupancy		
	Average	St. Dev	5th %	95th %	90% Width
3 Critical Care Unit (3CCU)	7.9	1.3	5.3	9.8	4.5
4 North (Surgical Unit)	28.9	5.2	18.5	34.9	16.4
Monitored Units	40.4	3.4	33.9	44.9	11.0
All Surgical Beds	48.5	8.1	32.4	57.6	25.3
All Medical Beds	62.5	4.9	53.6	68.4	14.8
All Med/Surg Beds	111.0	11.0	89.4	124.3	34.9
In-Patient Beds (seven units)	151.4	12.5	127.8	167.6	39.9
Entire Hospital	191.6	16.8	160.6	215.5	54.9

There has been a flurry of recent work in modeling LOS and patient flow with mixed exponential distributions and multi-compartment models [30–38] as well as on decision support systems for various aspects of bed management [1, 6, 39–42]. With respect to the role of hospital occupancy modeling, Seymour [43] notes that modeling has practical application in healthcare, the literature is coherent, gaining momentum, more publications are needed focusing on practical application, there are many barriers to implementation including medical antipathy to mathematics, software needs to made easily available, and there exists a need for software to be compatible with data collection systems routinely used within the health care sector. As can be seen from the work cited above, past research has addressed specific subsets of overall hospital occupancy prediction. However, gaps in the research remain and to the best of our knowledge there exist no research studies or practical implementations of hospital wide short-term occupancy predictions.

4 The big picture

Our approach to short term occupancy prediction is based on simple patient flow equations, a predictive occupancy database (the POD) and a set of predictive occupancy models (collectively called the POM). The predictive occupancy models are built on the observation that patient occupancy is a function of three basic flow processes:

1. Patients arriving to the hospital—They enter through a variety of methods (emergency, surgery, doctor's office transfer, transfer from other hospital) and are assigned to an initial bed.
2. Patients transferred within the hospital—During a patient's stay, he or she may move from a bed on one unit to a bed on another unit as their recovery proceeds.
3. Patients discharged from the hospital—Patients will be discharged from the hospital upon their recovery, their death, or their transfer to another facility.

Different versions of length of stay will play a prominent role throughout this paper. The total amount of time a patient spends in the hospital during a single hospital visit is called the *hospital length of stay* (HLOS). During a single hospital visit, a patient may spend time in a series of nursing units. The amount of time spent by a patient in each nursing unit visited is a *unit length of stay* (ULOS). During one hospital visit a patient may stop at the same nursing unit more than once, resulting in multiple ULOSs. If HLOS (ULOS) is prefaced by the word *current*, it refers to the amount of time spent in the hospital (unit) up until the current point in time. If HLOS (ULOS) is prefaced by *total*, we are referring to the entire stay in the hospital (unit).

4.1 Occupancy flow equations

Consider a group of n beds in a hospital. These beds can be one unit or a group of units (i.e., all medical/surgical beds). Time is represented by discrete hourly intervals. Time period i is the interval (t_i, t_{i+1}). At t_i, the beginning of time period i, each bed on the unit is either unoccupied or occupied by a patient. Define occupancy on some unit A to be

$N(i, A) \equiv$ number of occupied beds at time t_i on unit A.

The occupancy at the start of the next time period $i+1$ can be computed as

$$N(i+1, A) = N(i, A) + a(i, A) - d(i, A)$$

where

$a(i, A)$≡New arrivals to unit A during time period i (*inflow*)

$b(i, A)$≡Departures from unit A during time period i (*outflow*)

The expected occupancy on unit A at time $i+n$ is given by

$$E(N(i+n, A)) = N(i, a) + E\left(\sum_{j=0}^{n-1} a(i+j, A)\right) - E\left(\sum_{j=0}^{n-1} b(i+j, A)\right).$$

Accurate inflow and outflow forecasts should result in accurate occupancy forecasts.

4.2 Estimating patient inflow

Determining the expected patient inflow to a unit by hour requires estimates of volume for four main streams of patients: emergency arrivals, scheduled arrivals, direct arrivals (non-emergent, non-scheduled), and transfers into the unit from other areas. The methods used to estimate these inflow streams will differ as will the information sources used. Emergency and direct arrival forecasts are performed using time series models along with multinomial logistic regression models (for prediction of arrival unit). Predictions for scheduled arrivals use schedule related data to make forecasts on actual arrivals (schedules are not assumed to be deterministic) using multinomial logistic regression models again to predict the specific unit to which the patient will arrive.

4.3 Estimating patient outflow

The POM uses information *about each individual patient* within logistic regression models to make distributional projections about their outflow timing over the next 3–4 days. By assuming independence across patients, these can be added together to arrive at a departure distribution by hour for the entire unit. Data elements used in the models include attending physician, procedures, admitting diagnosis, gender, age, current HLOS and ULOS as well as a number of others which will be described in subsequent sections. Use of logistic regression models was based on computational ease, ability to include large numbers of predictor variables, and most importantly the ability to interpret the output of the models as probability predictions. For example, such models could help directly answer the question "What is the probability that this patient will be gone from the current unit within the next 24 h given what we know about the patient 6 A.M. today?" This construct proved useful at the patient level and also provided an easy method to aggregate the individual results into the larger aggregate model.

Figure 1 shows the actual cumulative distribution function at Study Hospital for HLOS for patients with two common surgical procedures—hip arthroscopy and discectomy. Note that from the time of arrival on the unit, there is virtually no chance of either of these patients leaving the unit via a discharge or transfer within 24 h. By the time the LOS reaches 60 h, 80% of the arthroscopy patients will have departed but only 4% of the discectomy patients will have departed. Notice also that the distribution functions have a step like nature due to the practical realities of the timing of discharges from hospitals during the day. Complications such as these cannot be assumed away if a POM is to be used in practice.

Figure 2 displays a graphical representation of the full POM concept. The POM itself can be organized into two high-level model groups: Patient Inflow and Patient Outflow. Both of these groups consist of a multitude of submodels to make up the overall inflow and outflow predictions. The inflow predictions separate patients based on their 'arrival path' into the hospital: emergency, surgical, or direct. The outflow predictions predict when and where patients will leave their current unit and bed. These submodels are then integrated to produce the overall occupancy predictions. The underlying modeling methodology for each submodel is noted in Fig. 2. The submodels are supported by and used in conjunction with the predictive occupancy database (POD). The POD collects and integrates many separate data streams to be accessed by the POM and is used for both periodic model parameter estimation as well as real time predictions.

4.4 Pilot projects and original POM

In conjunction with the SH, several pilot projects were identified where predictive information would be useful and an initial version of the computerized POM concept was developed and implemented. This version became known as alpha-POM (APOM). We introduce APOM in this paper for two primary reasons. It is useful to compare its accuracy to the subsequent model, dubbed BPOM, described in this paper. Although APOM is still a complex computational model, it has far less data requirements than BPOM and is therefore easier to implement. Structurally, APOM is similar to BPOM in that it is a patient-centric model of the form described in the equations of Section 4.1. The primary difference between APOM and BPOM is in the scope of input variables, submodels, and variable preprocessing. Generally speaking, APOM can be implemented using only basic information about the timing of a patient's arrival to the unit whereas BPOM requires a richer data source of patient demographics (for example: insurance provider, previous history, physician, surgical status, etc.). A quantitative comparison between these two modeling methodologies therefore allows us to explore the tradeoff between ease of implementation and accuracy of the predictive results.

The second reason we introduce APOM, is to demonstrate that this concept is feasible. The APOM predictions were distributed daily at the SH for a 6 month pilot program. Each morning reports were distributed to nursing units and administrators predicting the hospital occupancy over the next six shifts. An example of a report is shown in Fig. 3.

The daily pilot report predicted the number and status (Open, Tight, or Critical) of beds that would be available on the Med/Surg units over the next 2 days. The purpose of the pilot was to evaluate the accuracy of the predictions supplied by APOM and to begin to explore the options available to nurses and administrators when a critical bed shortage was anticipated. The pilot program had mixed success. The predictions were generally deemed to be helpful and reasonably accurate and dozens of people requested to be on the daily distribution list. However, using the predictive information to suggest an "optimal" policy was difficult and it was a challenge to gain the trust required for dramatic policy decisions such as closing down a unit for the weekend. Some of these implementation issues are discussed further in Section 9.

5 The predictive occupancy database (POD)

The numerous submodels that make up the BPOM rely on a large number of information systems and specialized databases for both parameter estimation and real-time data. The POD creation is a multi-step process which includes extraction, transformation and loading (ETL) of data from several hospital information systems, creation of new fields with high predictive potential, creation of population history and daily prediction tables, computation and inclusion of *hazard probability* fields, and creation of final tables for use in BPOM. Each of these steps is now described in further detail.

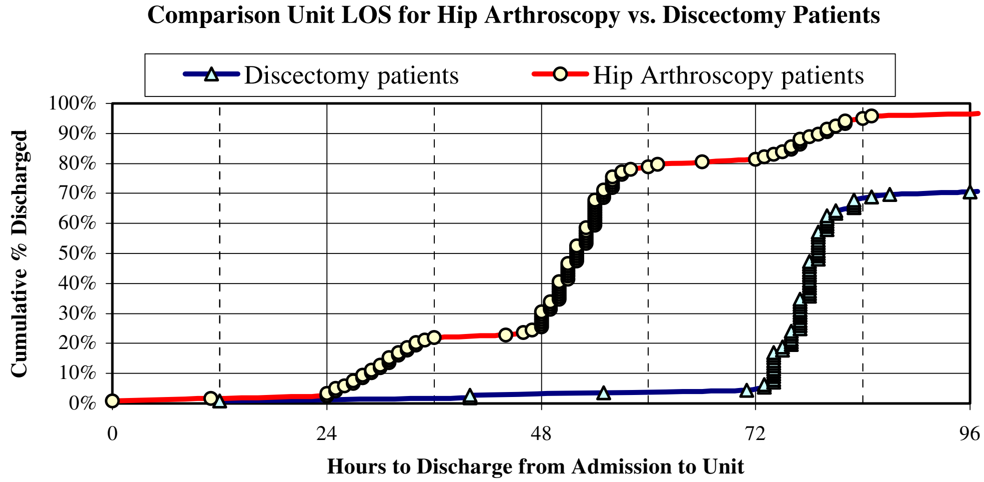

Fig. 1 Comparison of departure likelihood for two classes of patient

5.1 ETL from hospital information systems

The hospital information systems (HIS) used to construct the current version of the POD included databases from the emergency department, financial, surgical scheduling, and inpatient tracking systems. Each HIS has its own database design and data coding methods. Being a real-world system, there are fields with missing or invalid information. Field data types and naming conventions are often inconsistent across systems. Other field names are duplicated across systems causing difficulty in merging the different databases. In practice, every individual hospital has its own unique raw data sources and associated complications. To facilitate the modeling process, a data warehouse which we called the

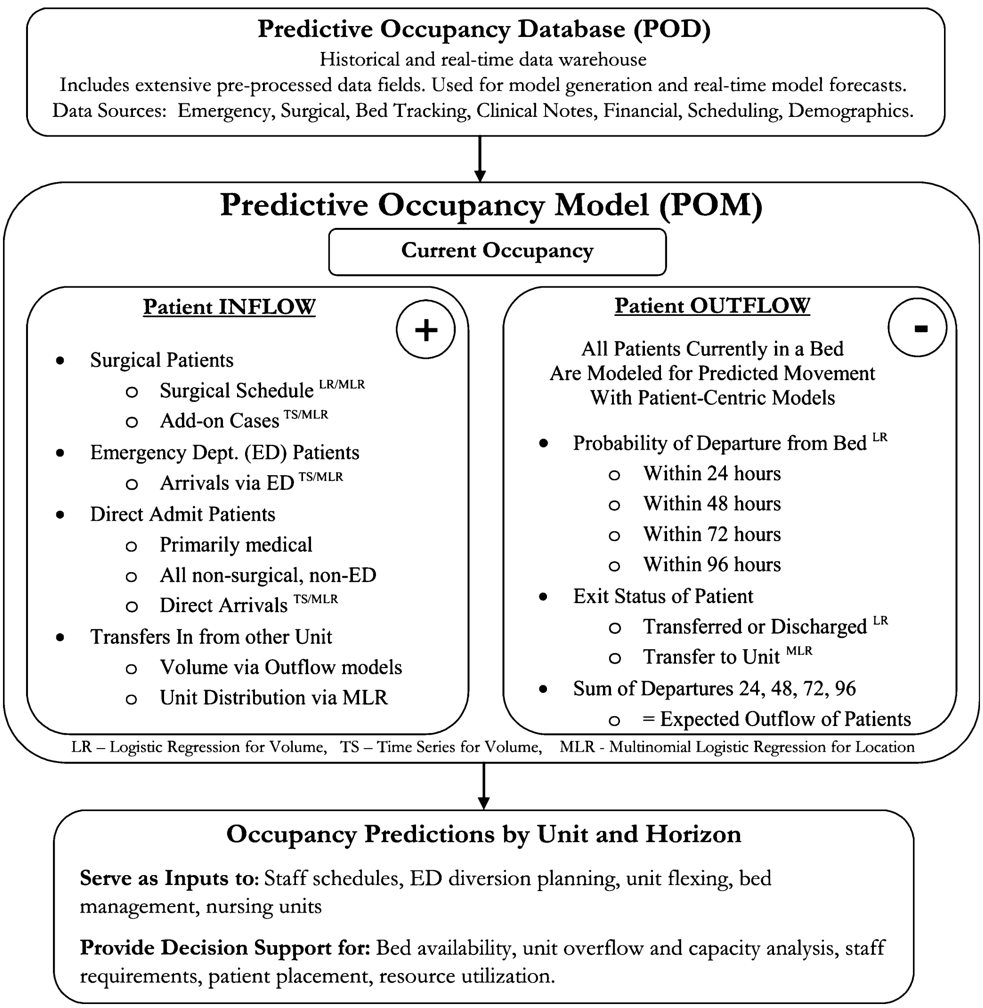

Fig. 2 High level view POD/POM

Fig. 3 Occupancy prediction report from pilot study

Med/Surg

	Monday			Tuesday		
	Day	Aft.	Mid.	Day	Aft.	Mid.
Med/Surg beds - available	127	127	127	127	127	127
Med/Surg beds - "predicted" utilization	109	107	118	122	118	124
Occupancy	**86%**	**84%**	**93%**	**96%**	**93%**	**98%**
Status	**Open**	**Open**	**Tight**	**Critical**	**Tight**	**Critical**

POD was designed to serve as a repository for data from the various HIS databases. The data warehouse allowed us to create a standardized database structure designed specifically to support the predictive modeling effort. Standardized field names, data types and coding schemes were developed. After this step, the POD was sufficient for creating all submodels of the BPOM (such as time series models for admission streams) except for the patient outflow predictions. These required additional data transformations which are discussed in the following sections.

5.2 Field selection and creation

There were over 200 raw data fields in the four data sources that were potential candidates for inclusion in the POD. Preliminary correlation analysis was done to eliminate fields with duplicate informational or little predictive value. Further conversion and modification needed to be done for a large number of fields to enhance their eventual statistical strength in the models.

5.2.1 Free text fields

Several of the critical modeling fields, such as chief complaint for emergency patients, were free text. These free text fields contain some of the most valuable information about an individual patient's condition when trying to predict their length of stay. Unfortunately, they do not map into a statistical model in any straightforward manner. This is a common problem in using real data. After reviewing our options for converting this information into a useable modeling field, we implemented a mapping technique to identify the most significant phrases. In the interest of space we won't go into the details here but significant predictive text phrases were identified and categorized. For example, the phrase 'carotid artery endarterectomy' was identified as the most significant free text procedure phrase. Patients with that exact partial phrase written in their procedure text had a total length of stay more than 60% higher than the average patient. This modeling step resulted in a reduction of unique field values for the free text fields from thousands to less than one hundred unique values.

5.2.2 Computed fields

Further work was done to calculate additional fields that were not directly in any of the raw data sources. Calculated fields included such things as patient readmitted within 7 days, patient readmitted within 30 days, last unit visited, and time since departure from operating room. The final result was 94 pre-processed data fields of information in the POD containing information about each patient that visited the facility over a 2-year period. Some of these fields were constant for the entire patient visit such as gender, initial diagnosis, or age upon admission. Other fields change value during the patient's stay such as current length of stay in hours or last unit visited. These dynamic variables were updated each day for every patient in the database. The 94 fields were a combination of numeric, indicator, and categorical data types. Even these field types posed challenges for the statistical modeling and were eventually converted to more usable forms in subsequent processing steps.

5.3 Prediction support tables

A table of historical occupancy was created to document the bed location of each patient by hour that visited the facility over two years. From this table two key modeling tables were created. The first, the 6 A.M. Prediction Table (6AMP), contained one record for each patient in some bed at 6 A.M. each day. Since occupancy predictions were to be made each day at 6 A.M., the *forecast time*, it was essential to know which beds were occupied at 6 A.M., and by which patient. The second, the Patient Unit Visit Summary (PUVS) contained information about the entire unit stay and had one record for each visit a patient makes to each unit during their hospital stay. This table was used to generate population length of stay distributions for specific patient types on a specific unit.

After development of the basic 94 variable POD, some initial modeling was done to investigate the predictive potential of the POD fields. The statistical approach chosen for the individual patient outflow submodels was logistic regression (LR for short). In order to use LR, categorical fields with large numbers of unique values had to be

Table 2 Comparison of LOS distributions

Patient Diagnosis	Hours of Unit LOS		Percent of Patients Who Leave	
	Mean	St. Dev	In 48 h	In 72 h
OsteoArthritis Hip	78	37	16	54
Cellulitus	84	55	43	62

transformed into a set of binary indicator fields. Unfortunately this led to an explosion in the number of fields and precipitated the development of *hazard probabilities* along with extensive preprocessing in attempt to keep the number of input fields manageable.

5.4 Adding hazard probability fields

The outflow models are used to predict the likelihood that an individual patient will leave the unit in the next 24, 48, 72 and 96 h given the current ULOS and the other patient specific variable values at each forecast time. Patient groups with similar mean ULOS can have very different length of stay distributions. See Table 2 for an example comparing patients with diagnoses of ‘osteoarthritis hip’ and ‘cellulitus’. The two ULOS distributions are both unimodal, skewed right and have similar means. Yet the probability that one patient type leaves in less than 48 h is 2.5 times greater than for the other patient type. In order to take such distributional information into account within LR models, we developed an approach to convert selected raw POD fields into hazard probabilities for use with LR. A sketch of the basic approach follows. To help illustrate the approach, we will use the example of a regular surgical nursing unit called 4N.

1. Using the PUVS table, a distribution fitting software library was used to find the best fit probability distribution for the total ULOS for every category in each categorical predictor field where the sample size was at least 50. A Chi-square test for goodness of fit was used [44]. If no distributional match with sufficient statistical significance was found, a table of empirical distribution probabilities was created for that category. All small sample size categories (less than 50 observations) were combined into an ‘Other’ category and fit to either a theoretical or empirical distribution.
2. We then step through the 6AMP table and compute departure probabilities for each patient for a number of categorical fields. For example, assume that the attending physician is Dr. Smith for some patient. Taking the current ULOS and using the appropriate distribution of total ULOS for Dr. Smith, use (1) to compute the probability the patient will leave in the next 24 h given the current ULOS (where cLOS is current ULOS, *f(x)* is the pdf and *F(x)* is the cdf of the ULOS distribution). This hazard ratio provides the probability of departure in the next 24 h. This is repeated for 48, 72, and 96 h thus converting one categorical field (in this example, attending physician) to four numeric probability fields: MD24, MD48, MD72, and MD96. These can readily be included in LR models and have higher information content than simple binary coding of a categorical variable. This hazard probability approach was taken for a total of 20 fields which are listed in Table 3 (along with the number of unique values for each field). Only the significant categorical values were fit to distributions. Note that the three fields in Table 3 which begin with the word “Path” are custom fields which are a string based representation of the sequence of units visited by each patient.

$$P(24) = \frac{\int_{\mathrm{cLOS}}^{\mathrm{cLOS}+24} f(x)dx}{1 - F(\mathrm{cLOS})} \tag{1}$$

In all a total of 5,069 distributions were fit for 4N (see Table 4). Since only about 17% of LOS distributions required an empirical fit, we feel that this method accurately captured the underlying population of the LOS distribution for each field category.

Table 3 Fields used to derive hazard probabilities

Field Name	Number of Significant Categories
Admitting Diagnosis Code	161
Admitting Physcian	444
Attending Physician	443
Bed Group	90
First Surgical Case—Procedure	541
First Surgical Case—Surgeon	719
Last Surgical Case—Procedure	122
Last Surgical Case—Surgeon	122
Clinical Guideline	127
Free Text Diagnosis	351
Free Text Diagnosis HLOS	376
Free Text Procedure	199
Free Text Procedure HLOS	222
Primary Insurance	222
Free Text Discharge Note	143
Path_UptoNow_StopAgg	97
Path_UptoNow_StopRaw	95
Path_UptoNow_StopType	98
Physician Service	445
Pending Transfer Location	52
Total Categories	5,069

We feel the hazard fields were an innovative approach to our modeling needs. They conveniently combined the various predictor categories (such as surgeon ID) with a quantitative acknowledgement of the impact the current LOS has on the residual LOS. We would like to be able to make an estimate that the hazard fields “provided an x% improvement on our ability to forecast occupancy” but we have yet to make that determination. Some first order model snooping indicated that hazard fields frequently made it into the final models but it is premature to interpret that with any specificity or significance. Given the effort required to preprocess these fields both for this and future POMs, it is important to determine how beneficial they are to the overall modeling objective. The ‘black box’ nature of this model does make it difficult to assess the impact of any one variable or set of variables but in future research we will attempt to develop methodologies that will allow us to estimate the quantitative benefits of the hazard field approach.

5.5 Create modeling tables

5.5.1 Field type conversions

Numeric fields such as number of visits in last 30 days, or number of surgical procedures during this visit needed no conversion for use in LR models. Twenty of the 94 raw data fields were converted to numeric hazard probability fields as described above. The remaining categorical fields were transformed into binary or numeric fields through the following process:

1. Calculate mean, standard deviation, and sample size of the ULOS for each unique value in each categorical field.
2. Compute the significance probabilities for the difference between the overall LOS mean and standard deviation and the mean and standard deviation of the LOS for each unique field value. Sort this list by statistical significance.
3. Select the most significant field value. Store this in a separate table. Remove this category from the field and recompute the overall mean and standard deviation of the entire field with it removed. Repeat steps 1 and 2 until there are no further significant field categories.
4. All field values identified as significant spawn a new binary field.

While this resulted in an entirely numeric set of preprocessed data for LR modeling, there was a dramatic increase in the number of fields to potentially send to the LR parameter fitting routine. For example, for unit 4N, of the 94 original fields, 52 were categorical fields. After step 4, those 52 fields produced 465 unique significant field/category combinations.

Table 4 Summary of distribution fitting for unit LOS

Distribution Family	Distributions—Number Fit by Family
LogLogistic	971
Empirical	849
ExtValue	484
BetaGeneral	473
Expon	456
Pearson5	443
InvGauss	441
Gamma	180
Lognorm	178
Pareto	92
Weibull	88
Triang	86
Logistic	77
Rayleigh	76
Erlang	58
Pareto2	49
Normal	43
ChiSq	11
Erf	10
Student	4
Total Distributions	5,069

5.5.2 Selecting final modeling fields

At this stage of the process there were still many hundreds of potential predictor fields for each of the LR models. Due to software limitations and concerns about computation time, we decided on a limit of 245 potential predictor variables. The database application only allows tables to have 255 total fields and several fields were required for the five independent variables and various record tracking identifiers for the patient, unit, and unit stop number. We chose to include all the hazard fields and all the direct quantitative fields. Obviously, this was a simplifying assumption that chosen because: a) we had put a great deal of effort into the hazard variables and felt they would have strong predictive content; and b) even after this automatic inclusion, there were still about 120 open slots for candidate predictor variables coming from the categorical and binary fields. We recognize that improved final selection of variables is an area for future research and perhaps a relatively easy avenue for future model improvement.

One last heuristic step was required to select those binary and categorical variables (with specific values) that demonstrated the greatest predictive potential. For each patient outflow model to be created (for example ‘Patient leaves in next 24 h’) all potential categorical and binary predictor variables were evaluated against the independent variable (the 0/1 actual output representing if the patient left within 24 h or not). The mean and standard deviation of the residual LOS on the unit (RULOS) was computed for each observed value of the variables. The mean of each variable value was compared using a standard *t*-test to the overall mean RULOS for all variables. Lower *p* values indicated that that mean RULOS for that variable value was significantly different than the overall mean. This was theorized to be a promising predictive variable for predicting RULOS since it contains information distinctly different than the overall mean RULOS. The *p* values were sorted by statistical significance and our relatively simple heuristic selected the top 120 predictor variables with a statistical significance of at least $p < 0.05$ and a sample size of at least 50.

There certainly could be substantial multicolinearity in the various predictor fields available to the modeling phase at this stage and our heuristic is susceptible to adding fields for consideration that provide little additional information to the ones previously added. However, it provided a relatively simple mechanism for reducing the number of independent variables for statistical modeling from many hundreds to the top 120 non-hazard fields (for each of the LR models). Intelligent variable selection throughout this process is a ripe area for improving model results.

The resulting POD is an extensive data source rich in information, cleansed of many data impurities, standardized for consistency, and supplemented with computed fields. It is very important to note that the POD was constrained to only include information that would be available for daily real-time predictions. For example, important patient classification fields like DRG or ICD9 were not included since they are not completed until chart review after patient discharge.

6 Predictive occupancy model

As illustrated in Fig. 2, the BPOM consists of numerous submodels for predicting the patient inflow and the patient outflow for each nursing unit over the prediction horizon. Our definition of prediction horizon in this study was 1–12 shifts ahead (8–96 h) from the time of the forecast.

6.1 Predicting patient inflow

All patient inflow models were based on several years of data collected at the SH and involved different methodologies for each inflow stream—emergency department admissions, direct admissions, surgical cases, and transfers between units.

6.1.1 Emergency department (ED) admissions

Past research [22] has shown time series to be an effective tool for ED admission forecasting. Our admission predictions for ED were based on time series models whose independent variables included day of week, seasonal factors, short-term trends (to pick up flu season spikes), and long-term trends (growth at the hospital). The total number of expected admissions to the hospital through the ED for the 24, 48, 72, and 96 h horizons was predicted. The aggregate numbers were then divided into expected initial unit location using multinomial logistic regression (MLR) models.

6.1.2 Direct admissions

These are patients who arrive neither through the ED nor through the operating room. Much like the ED submodel, a time series model was used to generate the total number of expected admissions which was distributed by unit using a MLR model.

6.1.3 Surgical admissions

This submodel predicted the number of patients who would be admitted to each unit based on the surgical schedule for the days within the forecast horizon. Schedule information and pre-admission patient information is used to determine the likelihood that each scheduled surgical patient is admitted to the hospital using a LR model. Individual scheduled patient admission probabilities are summed to arrive at total expected admissions. We also used the scheduled case time and scheduled case length to predict when during the day these patients would arrive on the unit. The total number of expected admissions is modified with a correction factor for cancellations and add-on cases (based on historical cancellation rates at SH). Finally, a MLR model is used to predict the initial admission unit for each patient.

6.1.4 Transfers in from other units

Predictions for transfers in come directly from the outflow models described below. The number of expected patient transfers into the unit over the forecast horizon was determined by summing the total predicted number of patient transfers within the hospital and proportionally distributing them to the various units based on a MLR model which predicts the transfer in unit based on the transfer out unit.

6.2 Patient outflow

The patient outflow models are all LR models predicting the probability of departure of each patient from each unit and were based on two years of data. Twelve months of data were used for model development and 12 months were set aside for testing of the model. For the patient outflow submodels, five separate models were created for each unit including four models for the probability of departure from the unit in the next 24, 48, 72, and 96 h, respectively, and one model for the probability that the departure is a transfer to another unit as opposed to discharge from the hospital. This resulted in 55 separate LR models (11 units * five models per unit). All predictor variables were either indicator (0/1) variables, pure quantitative variables, or probabilities (from hazard fields). There were no interaction terms used due to computational and complexity reasons. Each model was intended to predict the movement of each patient in a bed at 6 A.M.

Table 5 shows the model coefficients for two patient outflow LR models predicting the probability of patient outflow from unit 4 North at 24 and 72 h. While there is not space available to discuss these models in detail or to describe all the variable definitions that made it into these two models, several highlights are worth noting. A broad range of data sources are represented in the final model. Common sense tells us that diverse data sources will contain information useful to evaluating patient status. Hazard variables are liberally represented in the models. As a positive generalization, a reader experienced with hospital operations will note that nearly all the variables in the models 'make intuitive sense' as things that would have an impact on patient length of stay. On the negative side, a careful examination of the variables and coefficients reveals several instances of messy 'black box' model behavior with multiple variables perhaps being included where one would suffice. An example of this is two fields from the Clinical database measuring Anticipated Discharge for the Gone in 72 h model. One hazard field with a horizon of 72 h has a coefficient of 3.127 while a second hazard field built off the same field but with a horizon of 96 h has a coefficient of −3.368. While no general conclusions can be made from Table 5, at least for these models practical intuition is supported both by the variables included and the data sources represented.

An interesting area for future research will be to evaluate the importance of different data sources and the impact of the pre-processing steps to transform variables. On one hand, a broad range of variable types and sources is a welcome result since it confirms the intuitive assumption that many different factors do impact length of stay. On the other hand, a broad range of variable types and sources is a suboptimal result since it indicates that the modeling requirements may be complex and the implementation difficult due to diverse data demands.

It should be noted that there is likely significant colinearity between the predictor variables in all these outflow models. Furthermore, most of our resultant models had between 30 and 60 significant predictor variables in the final model produced through stepwise regression. It is possible that a more parsimonious model may produce comparable accuracy with far less data requirements. However, our objective with this phase of the research was to build the most accurate point-estimating 'black box' as possible for the entire hospital. In this phase, we are not making any determinations about the relative impact of any predictor variable on the departure probability nor are we creating error bands around our prediction. In many ways, these models are comparable to neural networks. Improvements in final variable selection, model parsimony, and prediction of forecast error confidence bands are topics for future work.

After computing the individual patient probabilities for departure, the sum of all patient probabilities on a given unit each day at 6 A.M. was the prediction for the expected number of patients who would depart that unit within the next 24, 48, 72, or 96 h. The sum of the discharge/transfer probabilities for all patients predicted the number transferred off the current unit to another unit in the hospital. This transfer total was applied to the 'transfer in' model above to find the expected location of each transfer. The predicted patient outflow totals were distributed over the 24 h period using the observed hourly distribution of discharges from that unit. Typically most patients are discharged between 9 A.M. and 5 P.M. on a unit but discharges do occur throughout the entire day. By using this approach, the predicted patient outflow pattern reflected this reality and the hourly predictions could be used for either bed availability or staffing needs.

6.3 Aggregating results and predicting future occupancy

The current occupancy on a unit plus the expected inflow of patients minus the expected outflow patients over the forecast horizon provides an initial BPOM prediction. As a final adjustment to the model we created one linear regression meta-model per unit using the current occupancy, the predicted inflow, the predicted outflow and various system variables that described factors that were not captured at the individual patient level. The predictors that went into this final linear regression model reflect the operational dynamics of the hospital as system—independent of individual patients. Some of these variables included in the final BPOM meta-model were overall hospital occupancy, current hospital occupancy by unit, the day of week, and recent error rates of BPOM predictions. It was felt by the staff that high or low occupancy in the hospital can affect the discharge patterns for patients. During high occupancy there is more motivation to discharge patients and when certain units or unit groups are near capacity, typical patient flow patterns are disrupted. This might invalidate our assumption of independence between patients for discharge probability and so by including occupancy in a meta-model we attempt to adjust for this. Discharge patterns are also affected by the day of week. Once a patient stays through Friday they are less

Table 5 Sample logistic regression models

			Departure Model Coefficients	
Data Source	Field Description	Field Type	Gone In 24 h	Gone In 72 h
ADT	Bed Control Service	Cat->Ind	−0.273	−0.484
ADT	Bed Control Service	Haz48	1.161	
ADT	Free Text Diagnosis vs. HLOS	Haz72		1.258
ADT	Free Text Diagnosis vs. ULOS	Haz24	1.378	
ADT	Free Text Diagnosis vs. ULOS	Haz48		0.752
ADT	Free Text Procedure vs. HLOS	Haz24	1.697	
ADT	Free Text Procedure vs. HLOS	Haz72		5.2
ADT	Free Text Procedure vs. HLOS	Haz96	−1.268	−3.481
ADT	Free Text Procedure vs. ULOS	Haz24	1.445	
ADT	Attending Physician	Haz48		1.207
ADT	Backwards Moves	Value		0.882
ADT	Checked in before 11 A.M.	Ind	0.148	0.235
ADT	CLOS<8 h	Ind	0.378	
ADT	Current Patient LOS in EC	Value	−0.022	−0.021
ADT	Current Patient LOS in ICUs	Value	−0.005	−0.006
ADT	Current Patient LOS in regular units	Value		0.001
ADT	Current Patient LOS in short stay units	Value		0.008
ADT	Intra Unit Move Last 48 h	Ind	−0.231	−0.384
ADT	Num Backward Moves (low acuity to high)	Value		0.572
ADT	Unit Path Summary to Date	Haz24		−1.068
ADT	Unit Path Summary to Date	Haz48	−1.118	
Clin	Anticipated Discharge	Haz48	4.23	
Clin	Anticipated Discharge	Haz72		3.127
Clin	Anticipated Discharge	Haz96	−5.226	−3.368
Clin	Free Text Diagnosis vs. HLOS	Haz24	1.15	
Clin	Scheduled Discharges	Haz24	4.986	2.886
Clin	Scheduled Discharges	Haz96	−4.222	−2.19
EC	EC arrival brought in by category ‘2’	Cat->Ind	0.17	0.205
EC	EC insurance payor ‘25’	Cat->Ind		−0.238
EC	EC insurance service group ‘2’	Cat->Ind		0.248
Fin	Admission Source ‘1’	Cat->Ind	−0.254	
Fin	Admitting Physician	Haz24		1.194
Fin	Admitting Physician	Haz24	2.824	
Fin	Admitting Physician	Haz72	−6.876	4.097
Fin	Admitting Physician	Haz96	10.016	
Fin	Diagnosis	Haz96	−1.601	
Fin	Insurance Code	Haz48	1.339	
Hist	LOS Ratio last 180 days	Value	−0.09	
Hist	Num Visits last 360 days	Value	−0.063	−0.073
Hist	Previous Visit>5 days	Ind		−0.273
OR	Evening OR Case	Ind	−0.356	
OR	Had OR Case within 48	Ind	−0.28	−0.443
OR	Num Procedures Done in OR case	Value	−0.091	−0.141
OR	OR First Case Procedure vs. PostOp ULOS	Haz48		2.891
OR	OR First Case Procedure vs. PostOp ULOS	Haz72		−11.544
OR	OR First Case Procedure vs. PostOp ULOS	Haz96		11.358
OR	OR First Case Procedure vs. ULOS	Haz24	1.087	
OR	OR First Case Procedure vs. ULOS	Haz48		0.575
OR	OR First Case Surgeon vs. HLOS	Haz72		0.855
OR	OR First Case Surgeon vs. PostOp ULOS	Haz24	1.862	
OR	OR First Case Surgeon vs. PostOp ULOS	Haz72		4.008
OR	OR First Case Surgeon vs. PostOp ULOS	Haz96	−3.222	−8.003
OR	OR First Case Surgeon vs. ULOS	Haz24		−3.47
OR	OR First Case Surgeon vs. ULOS	Haz72		−4.503

Table 5 (continued)

Data Source	Field Description	Field Type	Departure Model Coefficients Gone In 24 h	Gone In 72 h
OR	OR First Case Surgeon vs. ULOS	Haz96	4.13	8.32
OR	OR Case Class	Cat->Ind	0.259	0.481
OR	OR Case Length	Value	−0.23	−0.179
OR	OR Last Case Procedure vs. HLOS	Haz96		5.02
OR	OR Last Case Procedure vs. PostOp HLOS	Haz24		3.738
OR	OR Last Case Procedure vs. PostOp HLOS	Haz48		2.167
OR	OR Last Case Procedure vs. PostOp HLOS	Haz72		−11.592
OR	OR Last Case Procedure vs. PostOp ULOS	Haz24		−2.448
OR	OR Last Case Procedure vs. PostOp ULOS	Haz48	1.304	
OR	OR Last Case Procedure vs. PostOp ULOS	Haz96		8.524
OR	OR Last Case Procedure vs. ULOS	Haz24	−5.353	−1.953
OR	OR Last Case Procedure vs. ULOS	Haz96	6.231	1.644
OR	OR Patient Position	Cat->Ind	0.431	0.281
OR	OR Procedure	Cat->Ind	−1.006	−1.868
OR	OR Service	Cat->Ind		0.669
OR	OR Surgeon	Cat->Ind		0.804
OR	Pended Bed Control Service '1'	Cat->Ind	−0.795	−0.298
OR	Pended Bed Control Service '10'	Cat->Ind		−0.558
OR	Pended Physician Service Group '0'	Cat->Ind	−0.614	
	Constantfs		−5.618	−9.786

Data Sources: ADT—Admission/discharge/transfer; Clin—Clinical system; Fin—Financial; Hist—Historical Visits; OR—Surgical Scheduling
Field Types: Cat->Ind—Categorical converted to Indicator; Haz*X*—Hazard probability with X hour horizon; Ind—Indicator; Value—Actual value

likely to be discharged until Monday based on various internal factors (lab availability, doctor availability for discharge orders, etc.) Recent prediction errors were included also as a proxy for continuously adaptive parameter updating. This final meta-model produced predictions up to 4 days out for all 11 units and the 14 additional unit groups defined by hospital personnel.

7 Prediction assessment

7.1 Assessment of model predictions

In order to quantify the accuracy of the BPOM predictions and to explore their practical value we wanted to compare BPOM predictions with prediction methods that are currently used in the hospital to anticipate future occupancy levels. Though we know of no formal predictive occupancy models used daily within hospitals at this time, a variety of human heuristics are employed both consciously and subconsciously. Hospital employees are well aware of the need to manage occupancy swings. Just a few of the common daily questions administrators face include: Should we open the swing unit for more beds? Is the census going to stay low enough to reduce weekend staffing? To what units should we direct incoming patients to balance nursing workloads? Are we going to have a lack of beds that may shut down the ER? When will new beds be opening up? We discovered through conversations with hospital administrators that daily predictions are made intuitively by each administrator through basic information and knowledge about the hospital's operation. We were able to identify several basic categories of these human heuristic models and to develop quantifiable prediction models based on each of these human methods.

7.1.1 The budget heuristic

Staffing decisions and schedules are booked several weeks in advance. The average occupancy by shift and day of week from the previous 8 week period is used to project the occupancy and staffing needs by shift and day of week for the future period. Budget projections account for varying occupancy on different days of the week but otherwise these predictions are constant from one week to the next until they are recomputed after the next 8-week update.

7.1.2 The now heuristic

From our discussions with hospital personnel we believe this prediction method probably best reflects current practice. The method simply states that whatever the occupancy level is at the time of the forecast is anticipated to continue to be the same for the following shifts. It could

Table 6 Heuristic predictions

Day	Sunday		Monday			Tuesday			Wednesday			Thu
Shift	Day	Aft	Mid	Day	Aft	Mid	Day	Aft	Mid	Day	Aft	Maid
Horizon	1	2	3	4	5	6	7	8	9	10	11	12
OccDelta Computations	−1.2	−2.8	−0.3	0.9	−0.5	1.4	3.0	1.6	2.9	3.8	3.1	4.0
Predictions for Each Method												
OccDelta Predictions	22.8	21.2	23.7	24.9	23.5	25.4	27.0	25.6	26.9	27.8	27.1	28.0
Budget−8 Week Predictions	27.3	24.0	25.9	27.3	24.0	25.9	27.3	24.0	25.9	27.3	24.0	25.9
Now Predictions	24.0	24.0	24.0	24.0	24.0	24.0	24.0	24.0	24.0	24.0	24.0	24.0

go up or down from here but predicting it will stay the same is both easy and presumably unbiased. The Now prediction methodology probably serves as the basis for most decisions in the hospital with the possible enhancement of some version the OccDelta method described below that accounts for weekly patterns.

7.1.3 The OccDelta heuristic

This method is more sophisticated than the Now method and represents an optimistic scenario of an experienced manager for predicting occupancy several shifts or days ahead. It measures the average change in occupancy by unit from one shift to future shifts over the weekly pattern of 21 shifts. The OccDelta method uses a 3 month window for computing the parameter estimates. Note that for a 4-day out prediction there are 12 parameters, or deltas, which are used along with the current occupancy to generate a forecast. The deltas change from shift to shift over the 21 shifts of the week. Therefore, while a human may have a vague intuition of these weekly patterns it is very unlikely they would develop a forecast as accurate as OccDelta without computational assistance.

7.1.4 The heuristics in action

Consider the following example for the surgical unit 4N. Table 6 contains predictions for the next 12 shift horizons using the various human heuristic methods made at 6 A.M. on a Sunday morning in January when the current occupancy was 24 patients. The top line in Table 6 displays the 12 computed parameters for Sunday morning OccDelta predictions. The bottom three lines of the table display the actual predictions for the three human methods. Note that the OccDelta method predicts a drop of 2.8 patients (to 21.2 patients) from Sunday at 6 A.M. to the Sunday afternoon shift (4 P.M. to midnight) and an increase of 3.0 patients by the Tuesday day shift (up to 27.0 patients).

Table 7 Comparison of standard deviation of forecast error

	BPOM Results—Standard Deviation of Forecast Error for Unit=Surgical Beds								
	Standard Deviation of Forecast Error by Unit/Horizon					Percent Reduction in Error via BPOM compared to Other Methods for Unit/Horizon			
Horizon	APOM	BPOM	Budget	Now	OccDelta	APOM	Budget	Now	OccDelta
1	2.27	2.21	7.46	3.58	2.82	3	70	38	22
2	4.17	3.53	7.53	5.92	4.76	15	53	40	26
3	4.50	4.03	8.78	6.20	5.11	10	54	35	21
4	4.96	3.89	7.41	7.59	5.46	22	48	49	29
5	6.05	4.32	9.21	9.16	6.58	29	53	53	34
6	6.11	4.37	8.75	9.11	6.82	28	50	52	36
7	6.12	3.83	7.40	9.79	6.89	37	48	61	44
8	6.92	4.42	9.24	10.99	7.77	36	52	60	43
9	6.93	4.45	8.78	10.70	7.94	36	49	58	44
10	7.07	3.90	7.44	10.73	7.85	45	48	64	50
11	8.09	4.31	9.28	11.42	8.34	47	54	62	48
12	8.02	4.35	8.82	11.09	8.49	46	51	61	49
Whole Hospital (h=12)	17.2	8.68	16.94	23.02	14.91	50	49	62	42

Table 8 Comparison of correct predictions for surgical beds

	Forecast Horizon—Number of Shifts Ahead Predicted											
Prediction Class	1	2	3	4	5	6	7	8	9	10	11	12
BPOM Correct Days	335	293	262	284	264	265	286	261	264	287	269	265
BPOM Out of Range Days	16	58	89	67	87	86	65	90	87	64	82	86
APOM Correct Days	339	274	262	246	212	211	210	183	161	151	152	123
APOM Out of Range Days	12	77	89	105	139	140	141	168	190	200	199	228
BPOM — # of Additional days Correctly Predicted over APOM	−4	19	0	38	52	54	76	78	103	136	117	142
BPOM—Percent Reduction vs. APOM in number of days incorrectly predicted and staffed	−33%	25%	0%	36%	37%	39%	54%	46%	54%	68%	59%	62%
Budget Correct Days	189	189	161	189	142	160	188	142	160	188	143	160
Budget Out of Range Days	162	162	190	162	209	191	163	209	191	163	208	191
BPOM—Number of Additional days Correctly Predicted over Budget	146	104	101	95	122	105	98	119	104	99	126	105
BPOM—Percent Reduction vs. Budget in number of days incorrectly predicted and staffed	90%	64%	53%	59%	58%	55%	60%	57%	54%	61%	61%	55%

7.2 Modeling methodology and prediction result analysis

All of the statistical models were tested by a simulated run over one full year of actual data that was not used for estimation of the model parameters. Due to end of horizon effects for the patients still in the hospital at the end of our available dataset (our data ended on December 31), the last 2 weeks of the year were not used for prediction which resulted in 351 prediction days that were modeled with BPOM, APOM, and the three human heuristic models. Recall that APOM was the first version of the model and the BPOM is the new version described in this paper. While conceptually similar, the APOM used a limited set of variables and did not include hazard probabilities in the outflow models. On each prediction day, for each unit, we created a simulated prediction at 6 A.M. for the next 12 shifts. Combining the days (351), forecast horizons (12), prediction methods (5), and unit groups (25) resulted in over 500,000 total predicted unit occupancies for comparison to actual occupancy results. Needless to say we cannot

Table 9 Overview of success window predictions for selected unit groups

		Number of Horizons (max 12) where BPOM reduces the number of days inaccurately predicted by at least X% when compared to other methods											
		10%				25%				50%			
Success Width (# patients)	Area	APOM	Budget	Now	Occ Delta	APOM	Budget	Now	Occ Delta	APOM	Budget	Now	Occ Delta
±2	3 Critical Care Unit	9	9	12	12	5	5	12	10	1	1	9	6
±3	4 North	9	12	12	12	5	12	12	9	1	1	5	1
±3	Monitored	4	3	12	7	1	1	8	1	1	1	0	0
±5	SurgicalBeds	10	12	12	12	9	12	12	11	5	12	11	6
±5	MedicalBeds	11	12	12	12	7	12	11	10	4	1	5	3
±5	MedSurg	9	12	12	12	6	12	11	9	1	1	1	1
±10	House	12	12	12	12	9	12	12	12	5	7	12	2

Table 10 Critical and open bed status predictions for 'surgical beds'—horizons 7 and 8

	Prediction Result Classification	BPOM Results(%)	APOMResults		BudgetResults		NowResults		OccDeltaResults	
			%	BPOM Ratio	%	BPOM Ratio	%	BPOM Ratio	%	BPOM Ratio
Critical Beds>56 patientsHorizon 7	Correct Prediction—Overall	82	75	1.08	77	1.06	71	1.16	79	1.04
	Actual Critical Shifts Predicted Correctly	35	19	1.87	0	n/a	11	3.11	32	1.12
	Critical Shift Predictions that were Critical	68	41	1.68	n/a	n/a	21	3.19	56	1.23
Open Beds<42 patientsHorizon 8	Correct Prediction—Overall	89	77	1.15	70	1.26	66	1.34	75	1.19
	Actual Open Shifts Predicted Correctly	70	21	3.39	17	4.07	21	3.39	53	1.33
	Open Shift Predictions that were Open	82	64	1.28	32	2.58	26	3.11	49	1.68

cover all the results. We have selected a sample of the prediction results for reporting that best demonstrates the scope of analysis and some of the successes, failures, and key findings of this research.

Besides the sheer quantity of predictions posing a difficulty in analyzing the results, choice of metrics is also an issue. The predictions can be used in a variety of ways within the facility and a "successful" prediction depends on how that prediction will be used and what practical actions it will generate. We have included results for three distinct methods of gauging prediction accuracy: 1) standard deviation of prediction error, 2) percent within a specified error window, and 3) percent of predictions correctly forecasting occurrence of some threshold occupancy (high or low).

7.2.1 Standard deviation of prediction error

Table 7 shows the standard deviation of the prediction error by forecast horizon for the unit group 'Surgical Beds'. BPOM produced significant reduction in forecast error over all other methods tested. The reduction was most dramatic for the longer-term horizons (5 to 12 shifts out). All of the longer-term horizons had dramatic BPOM improvement (reductions in the standard deviation of prediction error between 30 and 60%) and most of the short-term horizons did as well. The APOM fared reasonably well in the one to three shift horizons but has been vastly improved on by BPOM for the four shift and longer horizons. The Budget method is particularly poor for short-term estimates since it does not adjust the prediction as other methods do—its error rate is typically double that of BPOM. We also include at the bottom of Table 4 the results from the most ambitious prediction made in this phase: the 12-horizon ahead, whole hospital prediction. BPOM predictions produced a 42–62% reduction in standard deviation compared to the three human heuristics. OccDelta is the best "human method" as expected but is still beaten soundly by BPOM. BPOM standard deviations are actually fairly stable from 4 to 12 shifts out while the other methods all significantly degrade over time.

7.2.2 Prediction success window

A second way of looking at BPOM accuracy is to use a *success window* methodology. In practice, the predictions can be used to help manage staffing levels or determine the need for unit openings or closings. If the prediction is 'close enough' the policy decision made using the prediction was correct. If the actual observed occupancy is outside the range of the prediction success window, then a different choice may have been made and the occupancy prediction was unsuccessful. For example, if the target patient-to-nurse staffing ratio is 6:1 and the occupancy is 30 patients, then the appropriate staffing level is five nurses. As long as the actual occupancy is between 27 and 33 patients then the choice of five nurses was correct. A summary of the findings are presented in Tables 8 and 9.

We will use the 8-shift ahead forecast as an example. BPOM produced errant forecasts on 90 days, APOM produced errant forecasts on 168 days, and the Budget method produced errant forecasts on 209 days. Since Budget is typically used to set staffing, BPOM presents an opportunity to improve the 8-shift (2 1/2 day) ahead prediction for appropriate staffing levels on an additional 119 out of 351 days.

Table 9 provides a more comprehensive overview of BPOM accuracy versus the other methods. This table shows the results for seven different unit groups summarized across all 12 horizons using the success window metric. Recall that 12 prediction horizons are made each day. Table 9 shows the number of those predicted horizons for which BPOM reduced the number of days with inaccurate predictions by 10, 25, and 50%. The largest possible number is 12 and large numbers represent better performance by BPOM over the other methods. Consider a specific example. Refer to the 'Surgical Beds' row for the 10% APOM comparison. It shows ten horizons where BPOM reduced the number of out of range days by 10% or more.

Note that BPOM is at least a 25% improvement over all three human methods for most of the units (often for all 12 horizons). The notable exception would be monitored beds where BPOM beats Budget and OccDelta by 25% for only one of the 12 horizons. With respect to success windows, BPOM is a clear improvement over the other heuristics. The strongest improvement came for the surgical units which is intuitive since these patients have more specific information available about why they are in the hospital. The weakest BPOM performance was for Monitored Beds and 3 Critical Care. This is interesting as these patients are the most acute which we initially presumed may be a strength of BPOM.

7.3 Critical bed prediction and open bed prediction

Another way of looking at the prediction results involves making a bed status prediction based on a 'threshold occupancy' for the unit or unit group. Full units can cause operational difficulties and bottlenecks in patient flow from other units, from the operating recovery rooms, or from the emergency department. Conversely, if a unit in the hospital empties out and has many open beds, that unit can be temporarily reduced in size (by consolidating patients and closing a wing for example) and staff costs can be saved or consolidated to other areas. Therefore, there were two predictions of interest about the status of each unit group: Open or Critical. All of the numerical thresholds that would trigger open or critical status for each unit group were set in consultation with hospital personnel. BPOM significantly outperformed all other methods for the vast majority of unit groups and horizons in predicting open or critical status. We have highlighted one specific example which shows several strengths and weaknesses of the BPOM critical/ open bed predictions in Table 10. This example is for the

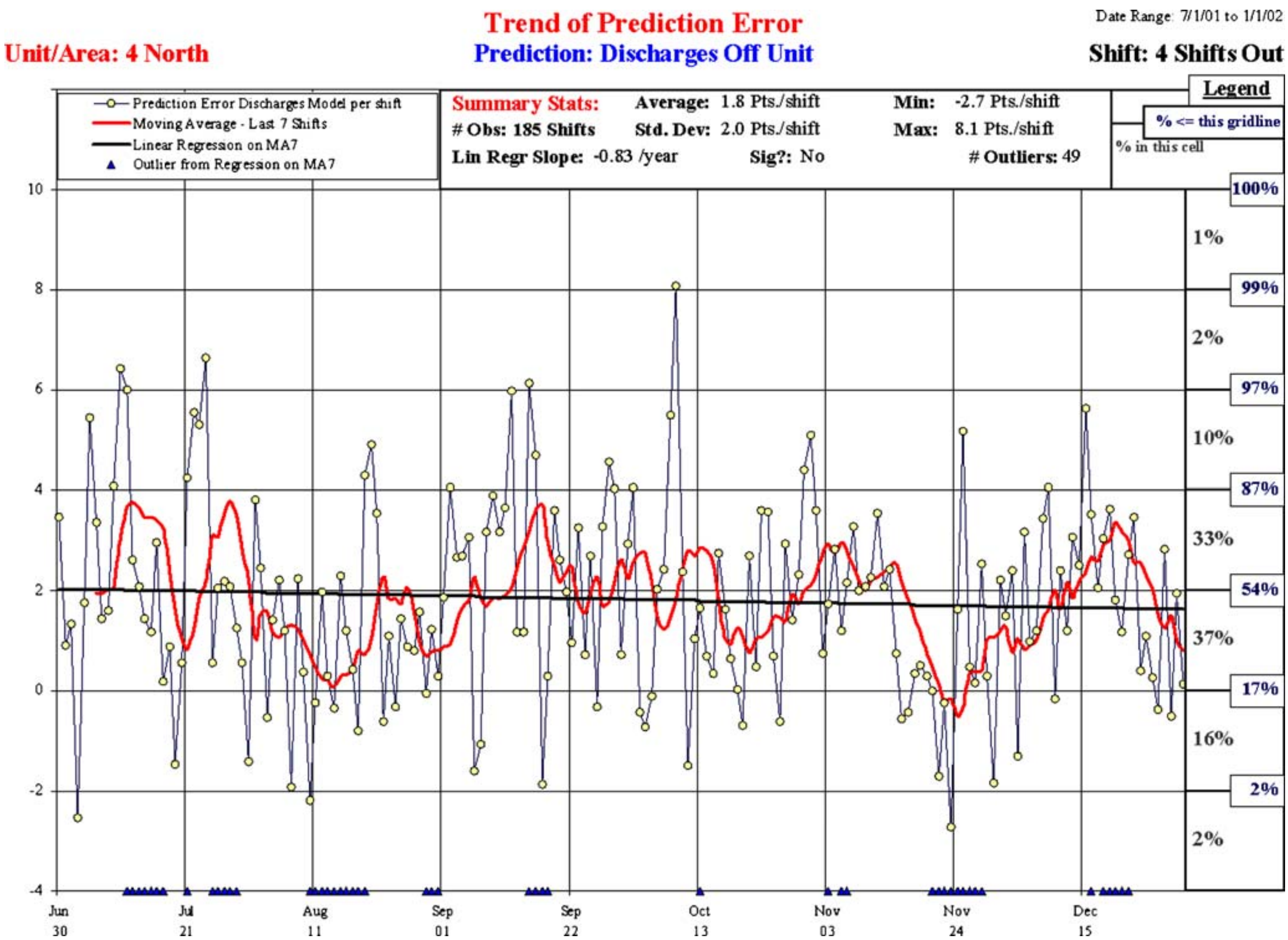

Fig. 4 Prediction errors over time

unit group 'Surgical Beds' for a horizon h=7 Critical bed prediction and h=8 Open bed prediction.

For h=7, 68% of BPOM predictions of critical status actually ended up being critical status (56 patients or greater) compared to APOM's 41%, Now's 21%, and OccDelta's 56%. Note that the budget method never predicts critical status due to its fairly static nature. For h=8, 70% of actual open shifts were correctly predicted by BPOM (42 patients or less). The closest competitor is OccDelta which correctly predicted 53% of open shifts eight horizons ahead.

The conclusion from the analysis of threshold event predictions is that once again BPOM outperforms the other methods. However, it is worth noting the relative competitiveness for some of OccDelta's predictions compared to BPOM. For example BPOM only produced between a 4 and 23% improvement for 'critical bed' threshold predictions and between a 19 and 68% improvement for 'open bed' threshold predictions depending on the metric. This indicates that for some predictive applications in the hospital it may be sufficient to employ easy to implement computational models sacrificing some accuracy in place of a more comprehensive and accurate POM that is far more complex to implement. This is another intriguing area of research.

8 Error analysis

Due to the enormous number of data sources and submodels contributing to the overall occupancy predictions it can be difficult to pinpoint the source of prediction errors. We developed a software tool to quantify error rates of the various model components and submodels such as the number of emergency admissions, transfers between units, or hospital discharges. Units or unit groups are analyzed by shift or hour of day and any of the shift horizons from next shift to 12 shifts out. The output of the error analysis tool is a set of reports detailing the prediction error and actual values for each patient flow component of the main model. By comparing predicted with actual error distributions we discovered bi-modal distributions and cases of highly non-normal error distributions. Both of these situations indicated that further refinements in the modeling efforts for that subcomponent were needed.

Another error report shows the historical trend of the prediction error over time—see Fig. 4. Since hospital patient flow patterns and admissions vary throughout the calendar year, it is important to look at these trends to determine which of the model components are susceptible to changes over time or to seasonal effects. Furthermore, the trend reports allow us to watch the degradation in model accuracy over time and to anticipate a 'model shelf life' after which the model parameters should be recomputed. In this example, the prediction error of the Patient Discharges from unit 4N with a four-shift horizon is relatively stable over time.

9 Challenges and conclusions

Implementing a real-time predictive model in practice involves significant theoretical, computational, and practical challenges. Having explored all three of these challenges to varying degrees, it is clearly a massive undertaking. Even implementing a simple model requires coordination with a variety of real-time data sources, creation and delivery of daily reports to decision makers, along with flexible policies which support effective use of the predictive information.

Theoretically, development of an accurate predictive model for forecast horizons over multiple days requires complex statistical modeling for predicting the detailed patient flow (location and timing) of hundreds of patients through the facility each day. Furthermore, forecasts must be generated for new arrivals via the emergency room, operating room, and physician offices. It involves calculation and aggregation of many statistical submodels and is ultimately a 'black box' due to its detailed inner complexity. We believe we have demonstrated the theoretical and practical feasibility of a model based on patient level data and flow equations. Ongoing modeling research includes comparison of complex and simpler models, extending the forecast horizon length, error analysis, and improving areas and units where the predictions were weak.

Computationally, a patient centric POM requires live access to a patient tracking system as well as other hospital information systems. Certainly the recent trends in information system infrastructure are supportive of this type of open real-time access though the challenges are formidable. There are also significant computational challenges in the area of model development. An information rich POD is essential to a successful POM. Their creation requires a robust and repeatable methodology across the development spectrum: data cleaning and transformation of the raw hospital data sources, creation of the POD tables, generating the submodels and aggregating them into the final model, reporting the results, and simulated testing of the POM results. Furthermore, any practical implementation of the POM concept would require quick and repeatable generation of the models and data sources so they could be updated to reflect recent trends in the hospital.

The challenge now is to create practical and innovative applications that utilize this or similar predictive models and deliver information to decision makers in ways they can use. In fact, a common question that arose during our pilot projects at SH was 'how do we best use this raw prediction?' Having a good idea of what is going to happen tomorrow does not necessarily lead to an obvious action today. The real benefits of the POM may not be realized until it serves as a foundational input for some very useful ‘killer applications.’ These applications would tackle some of the thorny utilization problems that have plagued hospitals for years. Rather than simply making a prediction as the POM does, they would use the POM information to provide recommendations and decision support on patient movement and placement through the facility in order to achieve key objectives. Some of the objectives we could envision the POM supporting through specifically tailored applications include: minimizing ambulance diversions, minimizing patient transfers, maximizing and balancing nursing utilization, and reducing occurrences of critical bed status.

This has not been a neat and tidy research study focusing on one particular modeling aspect of census prediction. Instead, it exists in that ephemeral world between management science and information systems. This is where the action is for modern decision support. We hope to have illuminated a number of areas calling for future research while at the same time placing such research questions in the broader context of real time decision support. Having patient centric probabilistic forecasts does open up new modeling possibilities but also highlights the many practical and unavoidable challenges that must be confronted when using real time information systems to drive model based decision support systems. There are many aspects of this project calling for more research and we hope that other management scientists can both attack those individual areas as well as to improve on the overall framework of this approach.

We have shown that accurate occupancy forecasts up to 4 days out are technologically feasible. The next challenge is to find innovative ways to leverage such information to aid hospital decision makers. We are confident that knowing what might happen in the future is a valuable asset and that having that advance knowledge will enable hospitals to provide more efficient and better quality care to their patients.

Acknowledgement This research was partially supported by the National Science Foundation under Grant no. DMI-0128556.

References

1. Clerkin D, Fos PJ, Petry FE (1995) A decision support system for hospital bed assignment. Hosp Health Serv Adm 40:386–400
2. Pofahl W, Walczak S, Rhone E, Izenberg S (1998) Use of an artificial neural network to predict length of stay in acute pancreatitis. Am Surgeon (The American surgeon.) 64:868–872
3. Cameron P, Scown P, Campbell D (2002) Managing access block. Aust Health Rev 25:59–68
4. Fatovich DM, Nagree Y, Sprivulis P (2005) Access block causes emergency department overcrowding and ambulance diversion in Perth, Western Australia. Emerg Med J 22:351–354
5. Schull MJ, Lazier K, Vermeulen M, Mawhinney S, Morrison LJ (2003) Emergency department contributors to ambulance diversion: a quantitative analysis. Ann Emerg Med 41:467–476
6. Ledersnaider DL, Channon BS (1998) SDM95—reducing aggregate care team costs through optimal patient placement. J Nurs Adm 28:48–54
7. Mackay M (2001) Practical experience with bed occupancy management and planning systems: an Australian view. Health Care Manage Sci 4:47–56
8. Nevidjon B, Erickson JI (2001) The nursing shortage: solutions for the short and long term. Online J Issues Nurs 6(1)
9. Ulrich BT (2005) How RNs view the work environment: results of a national survey of registered nurses. J Nurs Adm 35:389
10. McGillis Hall L, Doran D, Pink GH (2004) Nurse staffing models, nursing hours, and patient safety outcomes. J Nurs Adm 34:41–45
11. Sochalski J (2004) Is more better?: the relationship between nurse staffing and the quality of nursing care in hospitals. Med Care 42:67–73
12. Epstein SK, Slate DH (2001) The Massachusetts college of emergency physicians ambulance diversion survey. Acad Emerg Med 8:526–527
13. Armitage M, Flanagan D (2001) Improving quality measures in the emergency services. J R Soc Med 94(Suppl 39):9–12
14. Cerrito PB (2005) Visits to the emergency department as transactional data. J Healthc Manag 50:389–395
15. Erbas B, Rj H (2005) Sensitivity of the estimated air pollution-respiratory admissions relationship to statistical model choice. Int J Environ Health Res 15:437–448
16. Farmer RD, Emami J (1990) Models for forecasting hospital bed requirements in the acute sector. J Epidemiol Community Health 44:307–312
17. Kao EPC, Pokladnik FM (1978) Incorporating exogenous factors in adaptive forecasting of hospital census. Manage Sci 24:1677–1686
18. Lapierre SD, Goldsman D, Cochran R, DuBow J (1999) Bed allocation techniques based on census data. Socio-econ Plann Sci 33:25 (14 pages)
19. Lin WT (1989) Modeling and forecasting hospital patient movements: univariate and multiple time series approaches. Int J Forecast 5:195–208
20. Pang CY, Swint JM (1985) Forecasting staffing needs for productivity management in hospital laboratories. J Med Syst 9:365–377
21. Postma MJ, Ruwaard D, Jager, HJ (1995) Projecting utilization of hospital in-patient days in The Netherlands: a time-series analysis. IMA J Math Appl Med Biol 12:185–202
22. Tandberg D, Qualls C (1994) Time series forecasts of emergency department patient volume, length of stay, and acuity. Ann Emerg Med 23:299–306
23. Barie P, Hydo L, Fischer E (1996) Utility of illness severity scoring for prediction of prolonged surgical critical care. J Trauma 40:513–519
24. Mounsey J, Griffith M, Heaviside D, Brown A, Reid D (1995) Determinants of the length of stay in intensive care and in hospital after coronary artery surgery. Br Heart J 73:92–98
25. Walczak S, Pofahl WE, Scorpio RJ (2003) A decision support tool for allocating hospital bed resources and determining required acuity of care. Decis Support Syst 34:445
26. Buchman T, Kubos K, Seidler A, Siegforth M (1994) A comparison of statistical and connectionist models for the prediction of chronicity in a surgical intensive care unit. Crit Care Med 22:750–762
27. Zernikow B, Holtmannspotter K, Michel E, Hornschuh F, Groote K, Hennecke K (1999) Predicting length-of-stay in preterm neonates. Eur J Pediatr 158:59–62
28. Izenberg S, Williams M, Luterman A (1997) Prediction of trauma mortality using a neural network. Am Surgeon 63:275–281

29. Toumpoulis IK, Anagnostopoulos CE, DeRose JJ, Swistel DG (2005) Does EuroSCORE predict length of stay and specific postoperative complications after coronary artery bypass grafting? Int J Cardiol 105:19–25
30. el-Darzi E, Vasilakis C, Chaussalet T, Millard PH (1998) A simulation modelling approach to evaluating length of stay, occupancy, emptiness and bed blocking in a hospital geriatric department. Health Care Manage Sci 1:143–149
31. Harrison GW (2001) Implications of mixed exponential occupancy distributions and patient flow models for health care planning. Health Care Manage Sci 4:37–45
32. Mackay M, Lee M (2005) Choice of models for the analysis and forecasting of hospital beds. Health Care Manage Sci 8:221–230
33. Mackay M, Millard PH (1999) Application and comparison of two modelling techniques for hospital bed management. Aust Health Rev 22:118–143
34. Marshall A, Vasilakis C, El-Darzi E (2005) Length of stay-based patient flow models: recent developments and future directions. Health Care Manage Sci 8:213–220
35. Marshall AH, McClean SI, Millard PH (2004) Addressing bed costs for the elderly: a new methodology for modelling patient outcomes and length of stay. Health Care Manage Sci 7:27–33
36. Marshall AH, McClean SI, Shapcott CM, Hastie IR, Millard PH (2001) Developing a Bayesian belief network for the management of geriatric hospital care. Health Care Manage Sci 4:25–30
37. Marshall AH, McClean SI, Shapcott CM, Millard PH (2002) Modelling patient duration of stay to facilitate resource management of geriatric hospitals. Health Care Manage Sci 5:313–319
38. McClean SI, Millard PH (1998) A three compartment model of the patient flows in a geriatric department: a decision support approach. Health Care Manage Sci 1:159–163
39. Bagust A, Place M, Posnett JW (1999) Dynamics of bed use in accommodating emergency admissions: stochastic simulation model. BMJ 319:155–158
40. McClean S, Millard PH (1995) A decision support system for bed-occupancy management and planning hospitals. IMA J Math Appl Med Biol 12:249–257
41. Sorensen J (1996) Multi-phased bed modeling. Health Serv Manag Res 9:61–67
42. Wyatt S (1995) The bed occupancy management and planning system. Lancet 345:243–245
43. Seymour DG (2001) Health care modelling and clinical practice. Theoretical exercise or practical tool? Health Care Manage Sci 4:7
44. Law AM, Kelton WD (2000) Simulation modeling and analysis, 3rd edn. McGraw-Hill, Boston