Academic Emergency Medicine
Official Journal of the Society for Academic Emergency Medicine

ORIGINAL RESEARCH CONTRIBUTION

Predicting Emergency Department Inpatient Admissions to Improve Same-day Patient Flow

Jordan S. Peck, MS, James C. Benneyan, PhD, Deborah J. Nightingale, PhD, and Stephan A. Gaehde, MD, MPH

Abstract

Objectives: The objectives were to evaluate three models that use information gathered during triage to predict, in real time, the number of emergency department (ED) patients who subsequently will be admitted to a hospital inpatient unit (IU) and to introduce a new methodology for implementing these predictions in the hospital setting.

Methods: Three simple methods were compared for predicting hospital admission at ED triage: expert opinion, naïve Bayes conditional probability, and a generalized linear regression model with a logit link function (logit-linear). Two months of data were gathered from the Boston VA Healthcare System's 13-bed ED, which receives approximately 1,100 patients per month. Triage nurses were asked to estimate the likelihood that each of 767 triaged patients from that 2-month period would be admitted after their ED treatment, by placing them into one of six categories ranging from low to high likelihood. Logit-linear regression and naïve Bayes models also were developed using retrospective data and used to estimate admission probabilities for each patient who entered the ED within a 2-month time frame, during triage hours (1,160 patients). Predictors considered included patient age, primary complaint, provider, designation (ED or fast track), arrival mode, and urgency level (emergency severity index assigned at triage).

Results: Of the three methods considered, logit-linear regression performed the best in predicting total bed need, with a receiver operating characteristic (ROC) area under the curve (AUC) of 0.887, an R^2 of 0.58, an average estimation error of 0.19 beds per day, and on average roughly 3.5 hours before peak demand occurred. Significant predictors were patient age, primary complaint, bed type designation, and arrival mode ($p < 0.0001$ for all factors). The naïve Bayesian model had similar positive predictive value, with an AUC of 0.841 and an R^2 of 0.58, but with average difference in total bed need of approximately 2.08 per day. Triage nurse expert opinion also had some predictive capability, with an R^2 of 0.52 and an average difference in total bed need of 1.87 per day.

Conclusions: Simple probability models can reasonably predict ED-to-IU patient volumes based on basic data gathered at triage. This predictive information could be used for improved real-time bed management, patient flow, and discharge processes. Both statistical models were reasonably accurate, using only a minimal number of readily available independent variables.

ACADEMIC EMERGENCY MEDICINE 2012; 19:1045–1054 © 2012 by the Society for Academic Emergency Medicine

From the New England Veterans Engineering Resource Center (JSP, JCB), Boston, MA; the Emergency Medicine Service at the Boston VA Healthcare System (SAG), Boston, MA; the Engineering Systems Division at the Massachusetts Institute of Technology (JSP, DJN), Cambridge, MA; and the Healthcare Systems Engineering Institute at Northeastern University (JCB), Boston, MA.

Received December 16, 2011; revisions received April 30, 2012; accepted June 2, 2012.

Preliminary results from this study were presented at the Society for Health Systems 2011 conference and at the INFORMS Healthcare 2011 conference.

This research was supported by the Department of Veterans Affairs through the Veterans Integrated Service Network 1 (New England) Veterans Engineering Resource Center and the VA Office of Systems Redesign. Any views expressed herein represent those of the authors and do not necessarily reflect those of the Veterans Health Administration.

The authors have no potential conflicts of interest to disclose.

Supervising Editor: Eddy Lang, MD.

Address for correspondence and reprints: Jordan S. Peck, MS; e-mail: jspeck@mit.edu.

© 2012 by the Society for Academic Emergency Medicine
doi: 10.1111/j.1553-2712.2012.01435.x
ISSN 1069-6563
PII ISSN 1069-6563583

La Predicción en el Servicio de Urgencias de los Ingresos Hospitalarios Para Mejorar del Mismo día el Flujo de Pacientes

Resumen

Objetivos: Evaluar tres modelos que usan información generada durante el triaje para predecir, en tiempo real, el número de pacientes del servicio de urgencias (SU) que serán ingresados posteriormente en una unidad hospitalaria (UH) e introducir una nueva metodología para implementar estas predicciones en el marco hospitalario.

Método: Se comparó tres métodos simples para predecir el ingreso en el triaje del SU: una opinión de experto, una probabilidad condicional de Bayes ingenua y un modelo de regresión lineal generalizado con una función *logit* (*logit-lineal*). Durante 2 meses, se recopiló datos de las 13 camas del SU del Boston VA Healthcare System, que atiende aproximadamente 1.100 urgencias al mes. Se pidió a las enfermeras del triaje que estimasen la probabilidad de cada uno de los 767 pacientes valorados durante ese periodo de dos meses de ser ingresado tras su tratamiento en el SU, y que los situasen en una de las seis categorías que iban desde probabilidad baja a alta. La regresión *logit-lineal* y los modelos de Bayes ingenuos también fueron desarrollados utilizando datos retrospectivos y también fueron utilizados para estimar las probabilidades de ingreso para cada paciente que acudió al SU en los dos meses del periodo, durante las horas de triaje (1,160 pacientes). Los factores predictivos considerados incluyeron la edad del paciente, el motivo de consulta principal, el destino (cama estándar del SU o de consulta rápida), modo de llegada y nivel de urgencia (nivel de gravedad asignado por el triaje).

Resultados: De lostres métodos considerados, la regresión *logit-lineal* fue la mejor en predecir la necesidad total de camas, con un área bajo la curva (ABC) ROC de 0,887, una R^2 de 0,58, una media de error estimado de 0,19 camas al día y una media aproximada de 3,5 horas anterior a ocurrir el pico de demanda. Los factores predictivos significativos fueron la edad del paciente, el motivo de consulta principal, el tipo de cama asignada y el modo de llegada ($p < 0{,}0001$ para todos los factores). El modelo bayesiano ingenuo tuvo un valor predictivo similar, con un ABC de 0,841 y una R^2 de 0,58, pero con una media de diferencia de aproximadamente 2,08 por día en el total de camas necesarias. La opinión de la enfermera experta de triaje también tiene cierta capacidad de predicción, con una R^2 de 0,52 y una media de diferencia de 1,87 por día en el total de camas necesarias.

Conclusiones: Los modelos de probabilidad simples pueden predecir razonablemente el volumen de pacientes del SU que necesitará hospitalización a través de datos básicos recopilados en el triaje. Esta información predictiva podría ser usada para mejorar el manejo de camas a tiempo real, el flujo de pacientes y los procesos de alta. Ambos modelos estadísticos fueron razonablemente certeros, y usaron sólo un número mínimo de variables independientes fácilmente disponibles.

Emergency department (ED) crowding is a major problem nationally and occurs when there is a mismatch between the demand and supply of the resources needed to evaluate, treat, and discharge patients from the ED. Resource constraints may be related to resources controlled within the ED, such as nurse and provider staffing, or from resource constraints external to the ED such as the availability of radiology or laboratory capacity or the availability of open inpatient beds. Availability of inpatient beds to receive ED patients is arguably the single most important factor related to ED flow problems.[1–7]

Organizational solutions to address this problem can be categorized as static ones such as ''discharge by noon'' procedures and dynamic ones that are activated at the time ED crowding occurs, such as placing boarding patients in inpatient unit (IU) hallways, encouraging IU staff to schedule discharges to match historical patterns of expected admissions, and activation of inpatient resources based on the level of ED crowding.[8–10]

In contrast to most hospitals, manufacturing settings improve flow by starting some production early based on predicted demand, rather than waiting for all orders to be placed.[11] As illustrated in Figure 1, patient flow might similarly be modeled by estimating the likely number of patients who will be admitted at a point in the near future and sharing this information with IU staff who may then mobilize resources before crowding becomes an issue.

In current practice, bed requests and preparation to receive the patient often are delayed until admission is certain. ED crowding metrics have been used to trigger activation of additional resources. However, ED crowding metrics are an imperfect predictor of future demand because they are a composite of patients boarding in the ED and total ED census, uninformed by the patient characteristics. This has led to interest in the development of better models to predict patient admission.

Most studies that predict admission have focused either on the entire ED population[12,13] or on specific

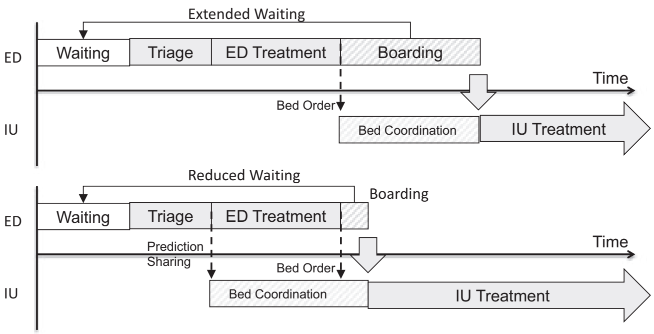

Figure 1. Overall bed management context of predicting and reacting to ED-to-IU admission estimates. (A) Current paradigm based on requesting a bed after or soon before ED treatment finishes, resulting in boarding; (B) proposed approach using admission prediction to begin bed coordination while patients are still receiving ED care. IU = inpatient unit.

How likely is it that the patient will need admission to the hospital?

- ❑ Definitely Yes (95-100%)
- ❑ Highly Likely (75-94%)
- ❑ Likely (50-74%)
- ❑ Unlikely (25-49%)
- ❑ Highly Unlikely (5-24%)
- ❑ Definitely No (0-4%)

Figure 2. Expert opinion triage questionnaire.

categories of patients[14–17] and treat admissions as binary in the sense of estimating ''yes'' or ''no'' at the patient level. This approach may be less useful when the goal is to predict aggregate demand. Pooling of patient admission probabilities across all ED patients should theoretically provide more precise predictions of near-future aggregate demand for inpatient beds. The primary objective of this study was to describe and evaluate three simple methods for generating admission predictions based on patient characteristics available at the time of patient triage. The secondary objective of this study was to introduce a new method for using predictive information by aggregating the individual patient predictions into a summative measure of near-future IU bed demand, rather than sharing single patient predictions.

METHODS

Study Design

Three methods to predict IU admission at the time of ED triage were developed and tested: expert opinion, naïve Bayes conditional probability, and generalized linear regression with a logit link function (logit-linear regression). Retrospective patient visit data were collected to form two data sets. Statistical models were created using a development data set. To avoid overestimation of model performance due to overfitting, the models were assessed using a separate validation data set, and final logit-linear regression and naïve Bayes models were identified. A third test data set was developed during a study of triage nurse expert opinion predictions. The performance of the final statistical models was then assessed on this test data set, which allowed for direct comparison of expert opinion and the two statistical models.

All portions of the study were approved by the institutional review board of the Boston VA Healthcare System. All analysis was performed using The MATLAB (R2011b-7.13.0.564, MathWorks, Inc., Natick, MA) and Microsoft Office Excel 2007 (Microsoft Corp., Redmond, WA).

Study Setting and Population

Boston VA Healthcare System is a federal tertiary care referral hospital devoted to the care of the United States of America's veteran population. It both serves the local community and acts as a referral hospital for the six other VA Medical Centers in the Boston region. The hospital is affiliated with two medical schools and has house staff from affiliated programs. The hospital has a 13-bed ED with an annual volume in 2010 of 12,672 visits; there are six inpatient wards and four special care units comprising approximately 170 beds. The hospital receives a capitation-based budget, but can also receive compensation from private insurers. The ED receives local ambulances carrying patients from the surrounding communities.

Study Protocol

The expert opinion portion of the study was conducted from September 22, 2010, to November 26, 2010, between the hours of 7:00 a.m. and 5:00 p.m. when ''out-of-room'' triage was operational. Due to lower patient volume between 11:00 a.m. and 5:00 p.m., patients are sent directly to ED beds and are triaged at the bedside, consequently bypassing the expert opinion process that took place at the triage stations. During the study period, 767 of 1,160 patients were triaged in the triage stations. Triage nurses classified each patient's likelihood of admission, using their expert judgment, into one of six categories (Figure 2).

Nursing staff were treated as an institutional review board–defined vulnerable population, and their predictions were not shared with any other ED staff or supervisors. Triage nurses were blinded to the specific purpose of the study, but were aware that it was being conducted to improve ED patient flow.

Structured interviews conducted with each of the triage nurses identified six possible patient characteristics available at the time of triage for possible inclusion in the predictive model. These were: 1) patient age—continuous; 2) primary complaint—free text entered by triage nurse; 3) ED provider—provider assigned to the patient; 4) designation—fast-track bed or standard ED bed; 5) arrival mode—stretcher, wheelchair, or ambulatory; and 6) urgency using the Emergency Severity Index (ESI) Levels 1 through 5.

For the development of the statistical models, retrospective triage data on all patients who entered the ED were collected from January 1, 2010, to May 6, 2010, totaling 4,187 patient visits. Using this development data set, analysis of variance was performed for each of the selected factors identified from the expert opinion study, and they were found to be significantly associated with hospital admission. A validation data set for model assessment and selection of the optimal mode was composed of ED visits between May 7, 2010, to May 31, 2010, and September 1, 2010, to September 21, 2010, totaling 1,614 patient visits. Table 1 summarizes basic patient characteristics for the patients included in the development data set, validation data set, and test data set.

Our statistical approaches make use of event probabilities and conditional probabilities, which require categorical data. Age was categorized into decades. Primary complaint was coded using a previously established system slightly modified to remove the free-text options, resulting in 62 complaint categories.[18] All other factors were already categorical. Table 2 lists examples of categories from each factor and their corresponding empirical probabilities estimated from the data, where P(X) means the unconditional probability of event X (used as the independent variable values in the logit-linear regression models) and P($X|Y$) is the conditional probability of event X given that event Y has occurred (used in the naïve Bayes models). For example, reading from Table 2, fifth row, historically 10.69% of all patients (admitted to an IU or not) arrive by wheelchair, whereas 15.49% of all admitted patients arrive by wheelchair and 49.11% of those patients arriving by wheelchair were admitted.

A naïve Bayesian model and a logit-linear regression model were then created for each of the 63 possible combinations of the six identified factors. For instance, a naïve Bayes model and a logit-linear regression model were created for the case where just patient age is used as a predictive factor; then patient age and primary complaint are used; then patient age, primary complaint, and mode of arrival are used, etc. These models are then applied to the validation data set to evaluate their performance. A final model that has a balanced performance in each measure was then selected and applied to the test data set, enabling the comparison of predictions for each of the three prediction methods.

Data Analysis

Each of the logit-linear regression and naïve Bayes models were constructed using the development data set of 4,187 historical patient points and evaluated for predictive ability using the 1,614 patient points that were included in the validation data set. To illustrate how the naïve Bayesian method works,[19,20] given three hypothetical factors "F1," "F2," and "F3," the admission probability for any particular patient is estimated as

$$\mathrm{P[Admit|F1,F2,F3]} = \frac{\mathrm{P[F1|Admit]} * \mathrm{P[F2|Admit]} * \mathrm{P[F3|Admit]} * \mathrm{P[Admit]}}{\mathrm{P[F1]} * \mathrm{P[F2]} * \mathrm{P[F3]}}.$$

If the model is a combination of patient age and complaint, the equation would only use those two factors; if the model is the combination of all six of the identified triage factors, the equation would use all six factors. The data for each factor are calculated using the development data set and a sample is displayed in Table 2. The naïve Bayes models were calculated using Microsoft Excel.

The logit-linear regression method that was employed uses the conventional log-odds link function and is calculated as

$$\begin{aligned}\mathrm{Log(P[Admit]/1\text{-}P[Admit])} = \beta_0 &+ \beta_1 * \mathrm{P[Admit|F1]} \\ &+ \beta_2 * \mathrm{P[Admit|F2]} \\ &+ \beta_3 * \mathrm{P[Admit|F3]}.\end{aligned}$$

The admission probability then is estimated via the inverse logit as

$$\mathrm{P[Admit|F1,F2,F3]} = \frac{e^{\beta_0+\beta_1*\mathrm{P[Admit|F1]}+\beta_2*\mathrm{P[Admit|F2]}+\beta_3*\mathrm{P[Admit|F3]}}}{1+e^{\beta_0+\beta_1*\mathrm{P[Admit|F1]}+\beta_2*\mathrm{P[Admit|F2]}+\beta_3*\mathrm{P[Admit|F3]}}},$$

where the size of the β coefficients represents the amount of influence each factor has on admission probability. Both methods can be calculated in standard spreadsheet or statistical software. The logit-linear regression models were calculated using the statistical package built into Matlab.

As described in the introduction, other published models for making admission predictions in the ED seek to assign a yes/no value to the patient.[12–17] This use of predictions can indeed drive the process suggested in Figure 1, by simply placing the admission order sooner. One common method to evaluate a prediction model that has the goal of suggesting a yes or no prediction is the receiver operating characteristic (ROC) plot's area under the curve (AUC). This value was calculated for each model and allows the user to calibrate the model to reduce false orders.

A qualitative method was also applied for evaluating model accuracy by categorizing patients into probability groups and judging whether the model accurately categorizes patients. For instance, if 20% to 30% of patients assigned an admission probability in the 20% to 30% range are actually admitted, then the model is seen as accurate in that range.

Binary prediction of admission increases estimation error by forcing the computed probability of admission from a fractional value to 1 or 0. While this may be useful strategy for early communication of likely IU admission for an individual patient, it increases estimation error when the predictions are summed across a group of patients to provide an estimate of aggregate near-future IU bed need. Instead, an ED can maintain an aggregate measure of future bed need based on the summation of raw probabilities.

This "running bed need" can be calculated using any method that generates an admission probability, such as those applied in this study. The resultant probabilities are totaled across all patients currently in the ED as shown in Figure 3 to produce a total momentary predicted bed need. For example, given n ED patients each with IU admission probabilities of p_1, p_2, and so on, the estimated total number of admissions E(T) to expect is E(T) = p_1 + p_2 + … p_n. Since the actual number may be

Table 1
Basic Patient Characteristics Between Development, Validation, and Test Data Sets

Characteristic	Development Data Set (24 Hours)	Validation Data Set (24 Hours)	Test Data Set (7 a.m.–5 p.m.)
Urgency			
1	7	1	1
2	56	26	7
3	2,441	892	585
4	1,347	388	318
5	336	302	249
Arrival mode			
Ambulatory	2,844	1,139	860
Stretcher	895	308	156
Wheelchair	448	167	144
Age, yr			
10–19	1	0	0
20–29	248	103	37
30–39	196	70	55
40–49	311	123	66
50–59	697	309	215
60–69	1,052	403	315
70–79	770	293	209
80–89	779	278	228
>90	133	35	35
Sex			
Female	200	78	53
Male	3,987	1,536	1,107

higher or lower, the standard deviation (SD) of total admissions, σ(T) can be estimated as

$$\sigma(\mathrm{T}) = \sqrt{(\mathrm{p}_1(1-\mathrm{p}_1) + \mathrm{p}_2(1-\mathrm{p}_2) + \ldots \mathrm{p}_n(1-\mathrm{p}_n))}$$

and, for more advanced applications, this can be used to generate confidence bounds on the number of predictions. Using these calculations, at any moment of a day, bed demand information can be compared with hospital-wide availability and appropriate actions taken.

Two methods are employed to evaluate how accurately each model generates the running bed need. The first method is to use visual inspection. Over the course of the day, the running bed need and the cumulative admissions are plotted side by side and it can be seen whether the two are well correlated. The peak value of the bed need for each day can then be compared to the peak number of admissions to see how well informed the IU staff were when they received this value.

Another mathematical way to assess model performance at generating the running bed need is to simply add all of the predictions for each day and to compare these predictions to the actual number of patients who were admitted each day by generating an R^2 correlation value. Noting that an R^2 correlation does not reflect errors in magnitude, this can be combined with a study of model residuals to achieve a better understanding of how well the model aggregates predictions. None of the methods described above are perfect evaluators on their own; however, in combination they provide a good sense of how well the model performs.

RESULTS

The 63 naïve Bayes and 63 logit-linear regression models created with the development data set were applied to the validation data set. With AUC, R^2, residual analysis, and goodness of fit into prediction categories, final models were selected for application to the test data set. Although multiple models performed well in some evaluative measures, a few performed consistently well in all. Consequently the final models chosen are not the only options, but provide a basis for comparing methodologies and a sense of model potential. It is likely that the unique traits of a hospital exploring this methodology will influence the weighting of factors that emerge as better predictors and the chosen model for that specific hospital.

When applied to the validation data set, the logit-linear regression model that performed consistently high in all analyses (and highest in some) consisted of patient age, primary complaint, bed type designation, and mode of arrival. In contrast, the naïve Bayes methodology incurred many tradeoffs, and the final model was chosen for consistent high performance in all categories, although it was the best in none. When applied to the test data set, these final models had AUCs of 0.841 and 0.887 for the naïve Bayes model and the logit-linear regression model, respectively. In contrast, the worst performing models were those that just used the ED provider as the predictive factor, with an AUC of 0.5 for both the naïve Bayes and logit-linear regression versions of the model.

Figure 4 compares how well the triage nurse predictions, the final naïve Bayes model, and the final logit-linear regression model assign patients into admission probability categories, using the test data set. For the latter two cases probabilities, which were continuously assigned by the models, were grouped into the same ranges used in the expert opinion for comparison; e.g., all patients assigned a probability between 0 and 4% were put in the ‘‘definitely no’’ category. As shown, the logit-linear regression results best fit the midpoint of each category in all but the ‘‘definitely no’’ tail (where naïve Bayes appears better), whereas expert opinion significantly underestimates admission in all categories.

Table 3 summarizes the significant factors for the best-performing logit-linear regression model. Note that all listed primary factors (patient age, primary complaint, bed type designation, arrival mode) are highly significant statistically ($p < 0.0001$ in all cases).

Figure 5 compares continuous actual versus predicted bed census (as described by Figure 3) for 15 days using expert predictions (top), logit-linear regression, and naïve Bayes (bottom). These data were generated by breaking up the test data set into hourly ED census. For each hour, the model predictions of admission probability for each patient were added together, with the probability of boarding patients taken to be a 1. For expert opinion, admission probability was taken as the midpoint for each category (i.e., 84.5% for patients in the highly likely category) and based only on patients who physically went through triage, as opposed to all patients in the ED, leading to reduced numbers in the chart. As shown, the logit-linear regression method

Table 2
Factors Tested for Admission Prediction Ability and the Empirical Probabilities of Occurrence

Factor/Code	Probability of Code	Probability of Code Given Admit	Probability of Admit Given Code
Designation	P[Designation]	P[Designation|Admit]	P[Admit|Designation]
ED	0.6237	0.9888	0.5383
Fast track	0.3763	0.0112	0.0101
Arrival mode	P[Mode]	P[Mode|Admit]	P[Admit|Mode]
Ambulatory	0.6793	0.4035	0.2014
Stretcher	0.2138	0.4415	0.7006
Wheelchair	0.1069	0.1549	0.4911
Urgency level	P[Urgency]	P[Urgency|Admit]	P[Admit|Urgency]
1	0.0017	0.0042	0.8571
2	0.0134	0.0211	0.5357
3	0.5830	0.9415	0.5477
4	0.3217	0.0267	0.0282
5	0.0802	0.0063	0.0268
Patient age, yr	P[Age]	P[Age|Admit]	P[Admit|Age]
10–19	0.0002	0.0000	0.0000
20–29	0.0592	0.0070	0.0403
30–39	0.0468	0.0134	0.0969
40–49	0.0743	0.0479	0.2186
50–59	0.1665	0.1606	0.3271
60–69	0.2513	0.2690	0.3631
70–79	0.1839	0.2085	0.3844
80–89	0.1861	0.2458	0.4480
90–99	0.0318	0.0479	0.5113
Provider	P[Provider]	P[Provider|Admit]	P[Admit|Provider]
1	0.0262	0.0126	0.1636
2	0.0160	0.0134	0.2836
3	0.0105	0.0112	0.3636
4	0.0086	0.0112	0.4444
5	0.1150	0.1019	0.3008
Primary complaint	P[Complaint]	P[Complaint|Admit]	P[Admit|Complaint]
Abdominal pain	0.0480	0.0685	0.4850
Abdominal problems	0.0504	0.0749	0.5048
Abnormal labs	0.0134	0.0275	0.6964
Cardiac arrest	0.0065	0.0148	0.7778
Cardiovascular complaint	0.0310	0.0516	0.5659
Chest pain	0.0480	0.0862	0.6100
Cold/flu	0.0595	0.0106	0.0605
Fainting/syncope	0.0074	0.0155	0.7097
Fall	0.0250	0.0311	0.4231
Fever	0.0158	0.0297	0.6364
Joint problems	0.0353	0.0056	0.0544
Kidney and liver failure	0.0151	0.0219	0.4921
Laceration	0.0096	0.0035	0.1250
Medication refill	0.0247	0.0000	0.0000
Psychiatric/social problems	0.0429	0.0523	0.4134
Respiratory problems	0.0909	0.1801	0.6728
Skin complaint/trauma	0.0420	0.0162	0.1314
Total probability of admit	0.3395		

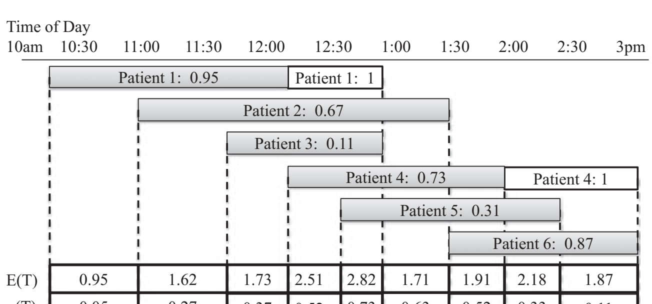

Figure 3. Conceptual illustration of real-time bed demand forecast (running expected number of admissions and SD).

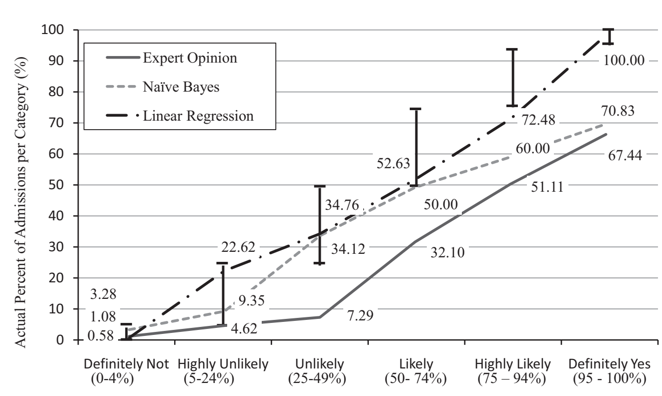

Figure 4. Categorized predictions of patient admissions versus percentage of patients admitted from each category.

appears to match actual admit volumes most accurately, with all three methods providing several hours advance notice. Over all 2 months of data used in the validation data set, the difference between predicted peak bed demand and actual demand for the expert opinion, naïve Bayes, and logit-linear regression methods on average were 0.82, 0.69, and –0.26, respectively (with SDs of 0.93, 1.81, and 1.59). These predicted peaks occurred on average 3.0, 3.7, and 3.52 hours before the actual peaks, respectively (with SDs 1.96, 2.20, and 1.96 hours).

Figure 6 compares actual and predicted total daily admissions, for the test data set, using expert opinion, naïve Bayes, and logit-linear regression, respectively. The R^2 value for the logit-linear regression is the greatest at 0.5826, followed by 0.5775 for the naïve Bayes model and 0.5243 for expert opinion. None of the methods perform well at predicting small admission volumes (since ideal fits would pass close to the origin as demonstrated by the horizontal line in each figure). As mentioned earlier, R^2 is a measure of how well the prediction trend follows the actual trend. On its own, R^2 does not prove the accuracy of a model; it is therefore valuable to analyze the residuals of the models. Figure 7 illustrates the residuals (predicted minus actual) for each model. The residuals expose the tendency of each method to overpredict on days of low admissions. Logit-linear regression appears to perform the best across most of the midrange, although underpredicting high admissions, where naïve Bayes and expert opinion appear slightly better.

Table 3
Model Parameters for Best Fitting Logit-linear Regression

Factor or Interaction	Coefficient (β)	Significance (p-value)
Constant	–7.02	1.4×10^{-56}
Designation (fast track or not)	5.48	1.0×10^{-21}
Primary complaint	2.89	5.3×10^{-24}
Patient age	3.39	1.9×10^{-05}
Mode of arrival	2.69	2.8×10^{-21}

DISCUSSION

Delayed patient ED discharge to an IU remains a major contributor to ED crowding. Most prior studies have focused on predicting individual patient admission or have focused on methods to predict longer-term admission trends. Common approaches, for example, focus on resource planning and staffing for future days,[21–24] predict short-term ED visit surges,[25,26] or use ED crowding indexes to predict ED congestion in the near future.[27–29] While forecasting can help set a baseline staff level, these forecasts are not based on same-day demand and therefore do not sufficiently inform real-time bed management and encourage behavior based on immediate, direct incentives. Alternatively, while predictors of short-term ED demand surges or measures of crowding may inform hospital staff and increase the pace of work and sense of urgency, these measures do not necessarily translate to high IU demand and admissions and therefore can mislead IU staff who choose their actions based on these measures.

In response to this, one suggestion to improve ED-to-IU flow is to predict admission demand when patients arrive to the ED.[30] In contrast to ED crowding measures, admission predictions are a more direct measure of incoming IU demand and can be used to more accurately inform the actions of IU staff. We have described a method to aggregate individual patient admissions predictions into a summative measure of near-future IU bed demand which may be useful for informing hospital-wide decisions on a daily, real-time basis.

The three prediction methods discussed in this paper are fairly easy to implement, with logit-linear regression being the most accurate in our test setting, followed by the naïve Bayes approach. ROC curve results suggest that these models could be used as part of the

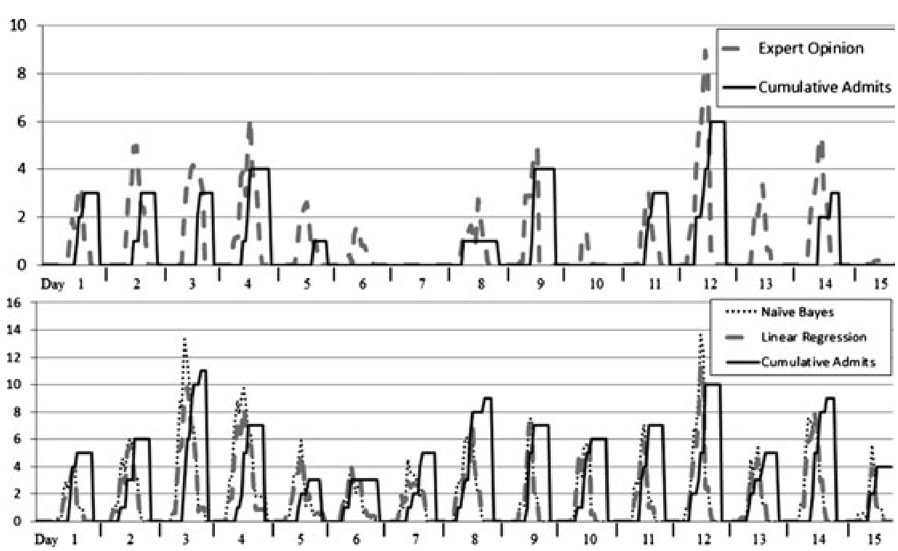

Figure 5. Real-time, statistically predicted–expected and actual number of cumulative admissions.

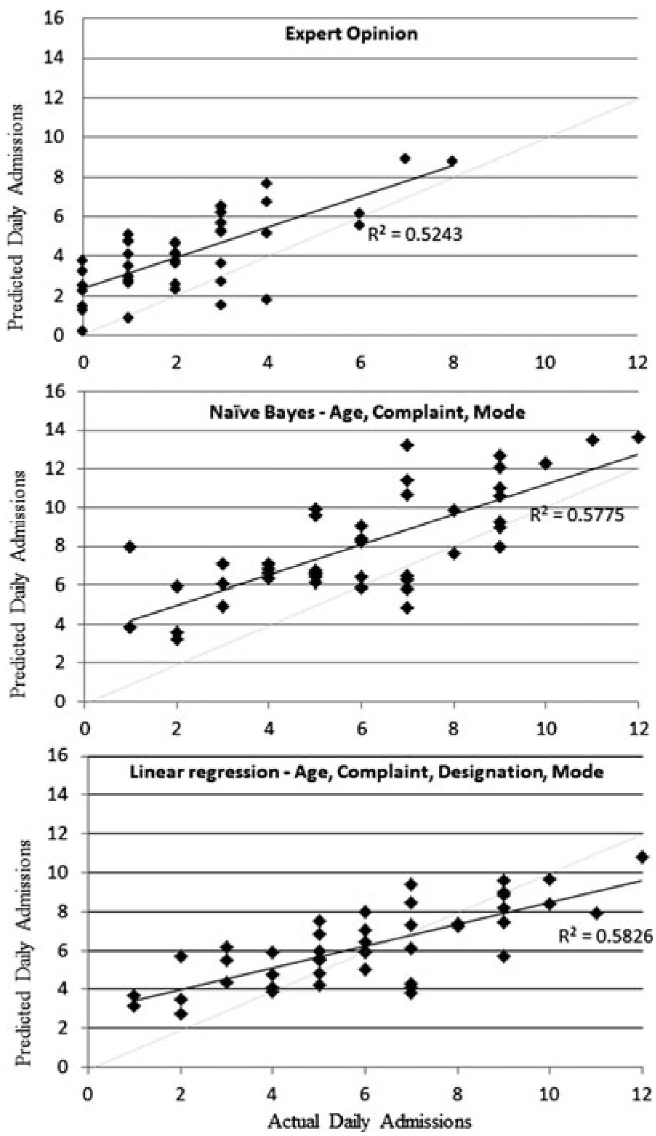

Figure 6. Correlation between predicted admissions and actual admissions based on expert opinion (*top*), naïve Bayes (*middle*), and logit-linear regression (*bottom*) approaches.

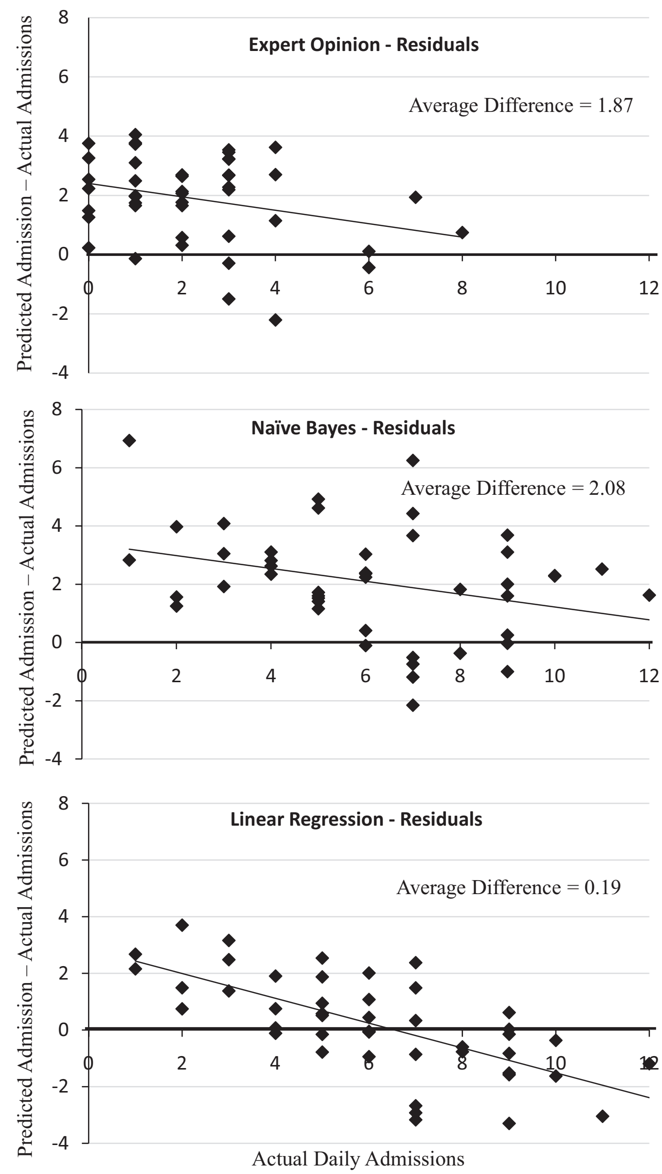

Figure 7. Prediction residuals (predicted minus actual) based on expert opinion (*top*), naïve Bayes (*middle*), and logit-linear regression (*bottom*) approaches.

current work flow, where orders for admission are made for specific patients using predictions rather than waiting for the final provider's order. However, all three methods that have been explored in this article also enable risk pooling of individual admission probabilities and thus may be more accurate at the aggregate level than methods that dichotomously classify each patient as "admit" or "not admit." (More complex approaches—e.g., Bayesian belief networks, neural networks, others—also tend to fall in this latter category.) For example, three ED patients each with a 45% IU admission probability might each be classified as "no admit" by such a method, spurring no action, whereas expected admissions under our approach is 0.45 + 0.45 + 0.45 = 1.35 with a SD of 0.86, suggesting the IU probably should open at least one bed and perhaps as many as three (using the mean plus 2 SDs).

The results in Figure 5 also suggest that predicted admission information can allow bed managers to start planning for peak demand significantly earlier than what currently occurs. Other benefits may result from sharing these data hospital-wide, such as allowing medical staff to better prioritize clinical activities, discharge ready patients in timelier manners, or manage bed preparations and room assignments for specific kinds of patients.

A practical question concerns how many beds to prepare relative to the expected demand, SD, and likely range. That is, if a need for 8.7 beds is predicted with a 95% interval ranging from 5.8 to 11.6 beds, it is not clear if a bed manager should plan for 6, 9, 12, or some other number of beds. This decision might be based on the relative costs of being over- versus underprepared. This decision also may evolve as a day progresses and knowledge is gained as to which early ED patients in fact were admitted. Admission likelihood estimates also could be updated during a patient's ED visit, such as based on test results, physician evaluations, and changes in physiologic status. Additionally, it could be useful to predict each ED patient's length of stay to better estimate IU bed demand timing (e.g., estimated ED arrival-to-ED discharge time) over the course of each day.

The final models that were chosen in this study may give rise to questions of face validity, given that many would consider patient urgency/triage level as the likely candidate for best predictor. Although models that used this factor did perform well, the authors believe that they may not be the best performers because age, primary complaint, and mode of arrival (which were in both of the final models) strongly influence ESI level and therefore could be acting as surrogate variables that may then include other less predictive parts such as predicted resource usage. Another possible explanation is that the study site may not assign ESI levels the same way as other hospitals where ESI data would be more predictive.

LIMITATIONS

While the methods described above are simple and effective, a few limitations exist. The simplicity of the models allows for a reduction in the data requirements necessary to achieve useful results. This makes the models and methods easily implemented by hospital staff with limited knowledge and software. However, this simplicity may also lead to reduced performance compared to more complicated models.[19]

The prediction models in this study were developed from data at one site and, in the above results, have not been demonstrated to generalize to other EDs. Furthermore, the ED where the model was developed receives a low patient volume and resides in a small tertiary care VA hospital, providing care to a specialized population. The authors believe that the set of factors that leads to an admission at a small hospital should be similar in larger hospitals. We believe the underlying method and triage characteristics identified can be applied to other settings, even if the calculated probability values for each factor may differ. This is a clear direction for future work that has already begun.

Predictive models only remain accurate if the underlying behavior of the system being modeled remains stationary. Therefore, models may need to be recalibrated when there are substantial shifts in admission patterns. For instance, such change may occur due to introduction of more effective treatment methods, treatments that shift care from the inpatient to the outpatient setting, changes in insurance practices, or payment structures. Similarly, the methodology for applying ESI in another site may lead to it becoming a more (or less) predictive factor as described in the discussion.

From an implementation perspective, both probability methods require initial effort to develop a coded data set, including coding primary complaints, and to calculate the probabilities and coefficients used in the logit-linear regression and naïve Bayes methods. Additionally, while we adapted a coding scheme from another study for convenience, it is unclear whether this scheme is best for prediction purposes. Any coding method may also suffer from intercoder reliability; the coding in this study was all performed by the same investigator. When implementing the proposed methodology in an ED setting, multiple people would be entering codes, which may reduce or improve the functionality of the chosen models. How this implementation affects model performance, and whether implementing predictions does indeed improve patient flow, are other important directions for future research.

CONCLUSIONS

We described and evaluated models for using data available at the time of triage to predict ED-to-inpatient unit admissions using expert opinion and two simple statistical models. We have also introduced a method for combining these predictions into a summative measure of near-term ED demand for inpatient unit beds. The logit-linear regression model performed the best, with an area under the receiver operating characteristic curve of 0.887, an R^2 of 0.58, and a daily average estimation error for the summative model of 0.19 beds. This method was based on four readily available inputs (patient age, primary complaint, bed type designation, and arrival mode). Recent studies have suggested that ED flow can be improved by anticipating inpatient unit bed demand. Our summative measure provides a reliable estimate of near-future inpatient unit bed demand that replaces traditional ED crowding measures for influencing inpatient unit staff behavior and decisions. This is in contrast to a yes/no predictor that seeks to preempt provider bed orders in current work flow paradigms. Future work is needed to assess if this predictive method yields similar results in other settings and also to demonstrate that it performs similarly when deployed.

The authors thank the staff of the Boston VA Healthcare System Emergency Department for their continued input, support, and participation; the staff of the New England Veterans Engineering Resource Center and VA Office of Systems Redesign for their continued support; and finally, MIT Professor Stephen Graves for his input.

References

1. Asplin BR, Magid DJ, Rhodes KV, Solberg LI, Lurie N, Camargo CA Jr. A conceptual model of emergency department crowding. Ann Emerg Med. 2003; 42:173–80.
2. Falvo T, Grove L, Stachura R, et al. The opportunity loss of boarding admitted patients in the emergency department. Acad Emerg Med. 2007; 14:332–7.
3. Hoot NR, Aronsky D. Systematic review of emergency department crowding: causes, effects, and solutions. Ann Emerg Med. 2008; 52:126–36.
4. Olshaker J, Rathlev N. Emergency department overcrowding and ambulance diversion: the impact and potential solutions of extended boarding of admitted patients in the emergency department. J Emerg Med. 2006; 30:351–6.
5. US General Accounting Office. Hospital Emergency Departments: Crowded Conditions Vary Among Hospitals and Communities. GAO-03-460. Available

at: http://www.gao.gov/new.items/d03460.pdf. Accessed Jun 30, 2012.
6. US GAO. Hospital Emergency Departments: Crowding Continues to Occur, and Some Patients Wait Longer Than Recommended Time Frames. GAO-09-347. Available at: http://www.gao.gov/new.items/d09347.pdf. Accessed Jun 30, 2012.
7. Williams M. Hospitals and clinical facilities, processes and design for patient flow. In: Hall RW. Patient Flow: Reducing Delay in Healthcare Delivery. Los Angeles, CA: Springer, 2006, pp 45–77.
8. Viccellio A, Santora C, Singer AJ, Thode HC Jr, Henry MC. The association between transfer of emergency department boarders to inpatient hallways and mortality: a 4-year experience. Ann Emerg Med. 2009; 54:487–91.
9. Rubino L, Stahl L, Chan M. Innovative approach to the aims for improvement: emergency department patient throughput in an impacted urban setting. J Ambul Care Manage. 2007; 30:327–37.
10. Vicellio P, Schneider S, Asplin B, et al. Emergency Department Crowding: High-impact Solutions. American College of Emergency Physicians. Available at: http://www.acep.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=50026&libID=50056. Accessed Jun 30, 2012.
11. Simchi-Levi D, Kaminsky P, Simchi-Levi E. Designing and Managing the Supply Chain. 2nd ed. New York, NY: McGraw-Hill/Irwin; 2002.
12. Li J, Guo L, Handly N. Hospital admission prediction using pre-hospital variables. BIBMIEEE Proc., Los Alamitos, CA 2009, pp 283–6.
13. Sun Y, Heng BH, Tay SY, Seow E. Predicting hospital admissions at emergency department triage using routine administrative data. Acad Emerg Med. 2011; 18:844–50.
14. Arslanian-Engoren C. Do emergency nurses' triage decisions predict differences in admission or discharge diagnoses for acute coronary syndromes? J Cardiovasc Nurs. 2004; 19:280–6.
15. Walsh P, Rothenberg SJ, O'Doherty S, Hoey H, Healy R. A validated clinical model to predict the need for admission and length of stay in children with acute bronchiolitis. Eur J Emerg Med. 2004; 11:265–72.
16. Levine SD, Colwell CB, Pons PT, Gravitz C, Haukoos JS, McVaney KE. How well do paramedics predict admission to the hospital? A prospective study. J Emerg Med. 2006; 31:1–5.
17. Clesham K, Mason S, Gray J, Walters S, Cooke V. Can emergency medical service staff predict the disposition of patients they are transporting? Emerg Med J. 2008; 25:691–4.
18. Aronsky D, Kendall D, Merkley K, James BC, Haug PJ. A comprehensive set of coded chief complaints for the emergency department. Acad Emerg Med. 2001; 8:980–9.
19. Witten IH, Frank E. Data Mining: Practical Machine Learning Tools and Techniques. Burlington, MA: Morgan Kaufmann, 2005.
20. Shmueli G, Patel NR, Bruce PC. Data Mining for Business Intelligence: Concepts, Techniques, and Applications in Microsoft Office Excel with XLMiner. Hoboken, NJ: Wiley-Interscience, 2007.
21. Tandberg D, Qualls C. Time series forecasts of emergency department patient volume, length of stay, and acuity. Ann Emerg Med. 1994; 23: 299–306.
22. Jones SS, Thomas A, Evans RS, Welch SJ, Haug PJ, Snow GL. Forecasting daily patient volumes in the emergency department. Acad Emerg Med. 2008; 15:159–70.
23. Jones SA, Joy MP, Pearson J. Forecasting demand of emergency care. Health Care Manage Sci. 2002; 5:297–305.
24. Abraham G, Byrnes GB, Bain CA. Short-term forecasting of emergency inpatient flow. IEEE Trans Inf Technol Biomed. 2009; 13:380–8.
25. Hoot NR, Leblanc LJ, Jones I, et al. Forecasting emergency department crowding: a prospective, real-time evaluation. J Am Med Inform Assoc. 2009; 16:338–45.
26. Schweigler LM, Desmond JS, McCarthy ML, Bukowski KJ, Ionides EL, Younger JG. Forecasting models of emergency department crowding. Acad Emerg Med. 2009; 16:301–8.
27. Bernstein SL, Verghese V, Leung W, Lunney AT, Perez I. Development and validation of a new index to measure emergency department crowding. Acad Emerg Med. 2003; 10:938–42.
28. Weiss SJ, Derlet R, Arndahl J, et al. Estimating the degree of emergency department overcrowding in academic medical centers: results of the national ED overcrowding study (NEDOCS). Acad Emerg Med. 2004; 11:38–50.
29. Epstein SK, Tian L. Development of an emergency department work score to predict ambulance diversion. Acad Emerg Med. 2006; 13:421–6.
30. Yen K, Gorelick MH. Strategies to improve flow in the pediatric emergency department. Pediatr Emerg Care. 2007; 23:745–9.