Health Care Management Science (2020) 23:339–359
https://doi.org/10.1007/s10729-019-09496-y

Prediction of emergency department patient disposition decision for proactive resource allocation for admission

Seung-Yup Lee[1] · Ratna Babu Chinnam[2] · Evrim Dalkiran[2] · Seth Krupp[3] · Michael Nauss[3]

Received: 26 March 2019 / Accepted: 7 August 2019 / Published online: 23 August 2019

Abstract

We investigate the capability of information from electronic health records of an emergency department (ED) to predict patient disposition decisions for reducing "boarding" delays through the proactive initiation of admission processes (e.g., inpatient bed requests, transport, etc.). We model the process of ED disposition decision prediction as a hierarchical multiclass classification while dealing with the progressive accrual of clinical information throughout the ED caregiving process. Multinomial logistic regression as well as machine learning models are built for carrying out the predictions. Utilizing results from just the first set of ED laboratory tests along with other prior information gathered for each patient (2.5 h ahead of the actual disposition decision on average), our model predicts disposition decisions with positive predictive values of 55.4%, 45.1%, 56.9%, and 47.5%, while controlling false positive rates (1.4%, 1.0%, 4.3%, and 1.4%), with AUC values of 0.97, 0.95, 0.89, and 0.84 for the four admission (minor) classes, i.e., intensive care unit (3.6% of the testing samples), telemetry unit (2.2%), general practice unit (11.9%), and observation unit (6.6%) classes, respectively. Moreover, patients destined to intensive care unit present a more drastic increment in prediction quality at triage than others. Disposition decision classification models can provide more actionable information than a binary admission vs. discharge prediction model for the proactive initiation of admission processes for ED patients. Observing the distinct trajectories of information accrual and prediction quality evolvement for ED patients destined to different types of units, proactive coordination strategies should be tailored accordingly for each destination unit.

Keywords Emergency department · Patient flow · Disposition decision prediction · Proactive coordination

MSC 62 J12 · 62H30 · 62P30

✉ Seung-Yup Lee
seungyup.lee@haskayne.ucalgary.ca

Ratna Babu Chinnam
ratna.chinnam@wayne.edu

Evrim Dalkiran
ey5796@wayne.edu

Seth Krupp
skrupp1@hfhs.org

Michael Nauss
mnauss1@hfhs.org

[1] Haskayne School of Business, University of Calgary, 2500 University Dr. NW, Calgary, AB T2N 1N4, Canada

[2] Department of Industrial & Systems Engineering, Wayne State University, 4815 Fourth St, Detroit, MI 48202, USA

[3] Department of Emergency Medicine, Henry Ford Hospital, 2799 W. Grand Blvd, Detroit, MI 48202, USA

1 Introduction and motivation

Overcrowding has long been identified as a critical issue in emergency department (ED) management [1–4]. In response, various approaches have been suggested including the adoption of fast-track units, advanced patient triage strategy, and the implementation of Six Sigma and "lean" programs to alleviate crowding in the ED [5–11]. However, prolonged patient "boarding" delays (i.e., delays experienced by patients admitted into the hospital by being held up in the ED due to admission, bed coordination, and transport delays), is known to be one of the most significant factors contributing to ED crowding [2, 12]. In particular, Pines et al. have investigated hospitals in 15 different countries and identified that the common main cause for ED crowding is the boarding of admitted ED patients [13, 14]. There is also a growing body of literature reporting the gravity of ED patient boarding in different countries and attesting to its negative clinical, operational, and financial impacts on healthcare management [15–19].

To improve patient flow out of the ED, different approaches have been introduced in the field of healthcare management science and operations research/management. Saghafian et al. [20] studied a patient streaming strategy that separates ED patients into two different streams according to anticipated admission decisions and discussed that the strategy could help alleviate the ED patient boarding issue. Shi et al. [21] investigated the inpatient unit (IU) discharge process and suggested that ED boarding delays can be reduced by altering inpatient discharge times. Osorio and Bierlaire [22] analyzed finite capacity queues to identify sources of bed blocking and suggest operational implications. Among others, as a potential remedy to considerably reduce boarding delay, the ideas of "early task initiation" (e.g., proactively checking admission criteria and seeking admission approval from hospitals) or "proactive resource preparation" for admission (e.g., preparing an inpatient bed for an impending admission) have received increasing academic and industrial attention [23–26]. The underlying premise is that modern-day electronic health record systems can support the realization of real-time hospital admission prediction models while the patient is still undergoing ED treatment. The rationale behind this proactive strategy for patient flow coordination is that bed request and preparation to admit a patient are often delayed until admission is certain [23]. When there are a large number of patient discharges at IUs combined with high demand for clean IU beds for admissions, bed management becomes a prioritization problem where environmental service staff (responsible for cleaning and turning around beds) should be dispatched according to admissions coming into each IU. This is symptomatic of most EDs for IU discharges tend to surge around midday (after morning rounds by providers) and environmental services struggle to keep up with the rate of patient discharge and are forced to clean beds based on admission priority lists rather than turning around beds as they become vacated. The patterns in Fig. 1 illustrate this situation routinely occurring within the ED-to-IU workflow of a leading level-1 trauma healthcare facility in the United States Midwest, where this study is conducted (HEM/ONC/BMP stands for hematology/oncology/blood and marrow transplant unit.). The ED suffers from severe crowding (indicated by the dashed line with the secondary y-axis) influenced by the increasing levels of patient boarding (indicated by the solid line with the primary y-axis) in the afternoon and evening (Fig. 1a). Even though there are unoccupied inpatient beds during the same period (Fig. 1b) generated by the high rate of IU patient discharges in the afternoon (Fig. 1c), less effective (i.e., reactive) bed management leads to excessive boarding delays.[1] It is under these circumstances that ED patient admission predictions combined with proactive coordination can limit the trajectory of boarding delay by judiciously allocating resources and proactively initiating tasks for predicted admissions [27–30].

While there has been some progress with ED patient admission prediction modeling research to enable proactive coordination [23, 26, 31–37], the models are lacking in their granularity to allow operationalization of prediction outcomes in real-world settings. Just predicting that a patient will be "admitted" will not necessarily allow full proactive coordination of resources across the ED-to-IU workflow for streamlined patient admission and flow since in most cases the allocation of necessary resources, e.g., inpatient beds, requires information on "which IU the patient is likely to be admitted". Generally, inpatient care can be categorized into three main types based on the intensity of required care [38]: general care (least intensive care), telemetry/stepdown care (intermediate care), and intensive care. They in turn define the three main types of IUs, i.e., general practice unit (GPU), telemetry unit (TU; also known as stepdown unit), and intensive care unit (ICU). While GPU and ICU constitute traditional and well-established inpatient care units in hospitals, TU has been increasingly adopted in hospitals for providing an intermediate level of care for patients with requirements between that of GPU and ICU. The comparable settings can be readily found in the literature describing other healthcare institutes, where we observe that the inpatient care structure comprised of GPU, TU, and ICU is becoming common in many different countries [39–41].

The study hospital has the aforementioned three main types of IUs for regular IU admissions. While a disposition decision can further specify the most proper specialty unit within a main IU for a patient, if all beds in that specialty unit are in use, the patient is generally transferred to the second most proper specialty IU and so on (also termed "overflow") and the disposition decision is updated accordingly. It is worth noting that the features of inpatient beds and accessories are generally common across a main unit, and this in turn constrains overflow to happen within the same main unit. In addition to the overflow policy, most hospitals physically separate ICU, TU, and GPU, pooling resources (e.g., beds and nursing services) within each main unit, exclusively. Therefore, given routine patient overflow occurring within a main IU and the resource pooling strategy at the study hospital, we seek to predict disposition decisions at the level of the main IUs. In addition, the ED physicians at the study hospital were only making disposition decisions at the main IU level about half the time (without further specifying any specialty) further justifying the prediction approach.

In addition to the IUs, the study hospital also operates an observation unit (OU) for providing extended care to patients who finished care processes in the ED. OUs are increasingly used as a short-stay (< 23 h) clinical decision

[1] In Figure 1b, we conservatively exclude the case of having only one unoccupied bed to account for any possibility that a bed is temporarily unavailable due to, for example, infection concerns from a fellow roomed patient and so on.

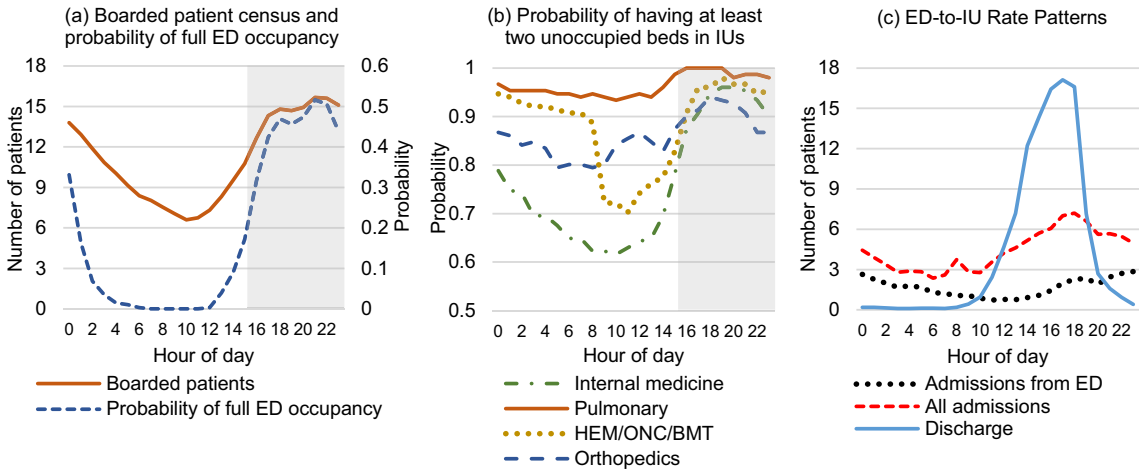

Fig. 1 ED and IUs patient flow graphs. Note: The shaded area indicates the period within a day when the ED suffers from severe crowding. During this period, there are unoccupied inpatient beds for most weekdays that could help alleviate the ED congestion issue when related processes are proactively coordinated

unit for patients who require further observation and treatment. Even though the OU is not a part of regular IUs in most hospitals, it plays a significant role to control demand to IUs, and a significant number of ED patients can be sent to OU (around 6% in the study hospital).

The main goal of this research is to predict disposition decisions at the level of the main IUs and the OU for facilitating the unit-specific proactive coordination of admission processes to reduce ED patient boarding. Furthermore, our prediction modeling strategy incorporates the progressive nature of ED care processes, where more clinical information is revealed and accumulated for the patient as he/she goes through more ED processes (e.g., monitoring and testing) and treatment. This is the first study to conceptualize and evaluate the ED disposition decision prediction problem in the context of proactive coordination.

2 Methods

2.1 Study design

In this study, we define three distinct levels of classification schemes for modeling disposition decisions, considering both practical and academic relevance (Fig. 2b). At the first level (denoted C1), the outcome of ED patient disposition decision classification is the binary admission decision (i.e., admission vs. discharge). The OU class patients are included in the admission class.

The C1 classification scheme has been adopted in most of the ED disposition decision prediction modeling research to date [23, 26, 31–37]. Due to its relatively simple structure, it could produce the most accurate results. At the second level (C2), the admission class at the C1 classification scheme is further segmented into two subclasses, i.e., IU vs. OU admission. While the IU class is regarded as official admission, OU patients may not be considered as "admitted" depending on the hospital even though OU patients could have features that are clinically similar to IU patients. Rather, the OU treatment is often regarded as "extended ED care", and the mechanism of patient transfers to OUs is different from regular IUs in many hospitals. Finally, at the most granular level (C3), the IU class at C2 is further categorized into three main IU classes, i.e., ICU, TU, and GPU. We believe that the C3 classification scheme is where the most significant operational benefit can be derived by enabling the unit-specific proactive coordination of admission processes across the ED-to-IU workflow. To provide further justification for the need to model each classification scheme, we summarize the possible clinical and operations management applications and use cases for prediction results stemming from each classification scheme in Table 1. In doing so, we considered general resource allocation practices in hospitals to provide a common and general description. Note that the table lists the incremental utility of prediction results at each classification scheme, i.e., utility of classification at C3 also includes the utility of C2 classification; C2 includes the utility of C1.

As discussed earlier, each IU class at C3 can be further subdivided into specialty subunits. For instance, in the study hospital, the GPU has 12 distinct subunits based on specialty of care, including general internal medicine, nephrology, obstetrics/gynecology, neurology, pulmonary, and so on. Besides the overflow policy, about half the time, ED physicians only make disposition decisions at the main IU level without further specifying any specialty. Therefore, in this study, we focus on predicting the disposition decision at the level of the five classes at C3, considering its higher practical relevance. As expected, as we increase the granularity of prediction (i.e., from C1 through C3), the prediction problem becomes more challenging. Considering how the classes at C3 (the ICU, TU,

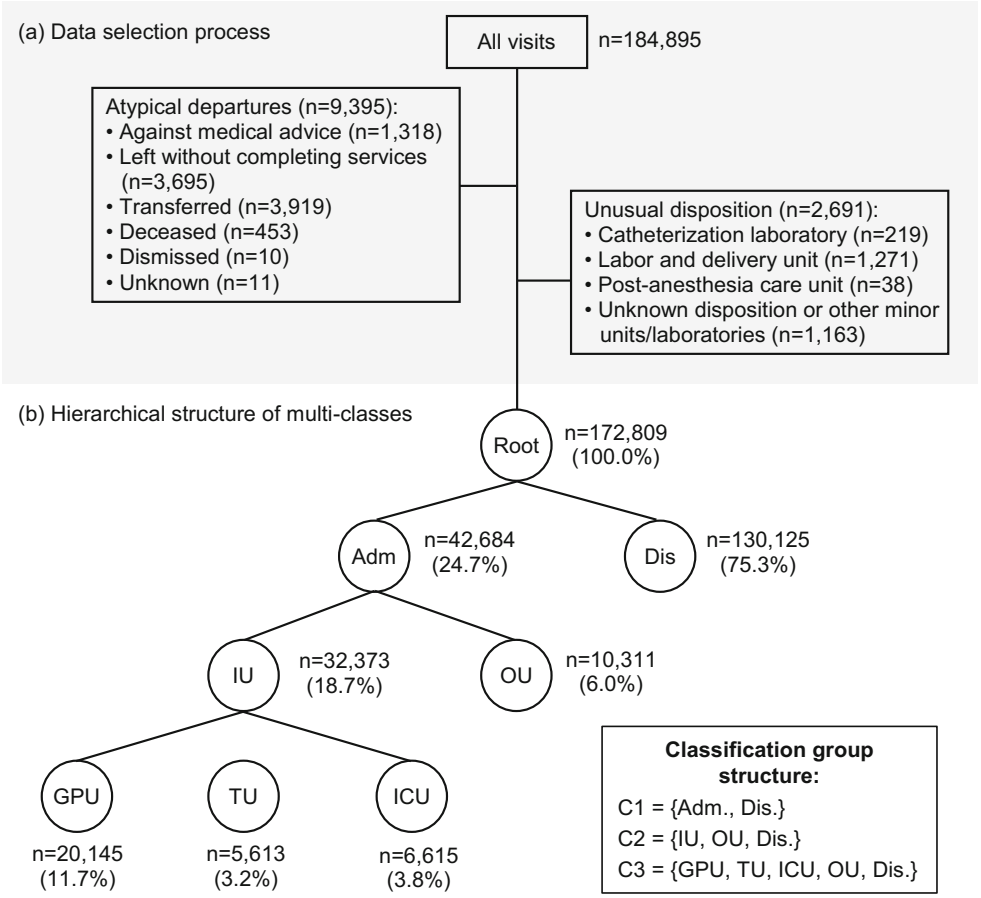

Fig. 2 **a** The data selection process of the study and **b** the hierarchical structure of ED disposition decision. Note: Adm., admission; Dis., discharge

GPU, OU, and discharge classes) are defined, the classification model should be able to discriminate the "clinical care intensity" of ED patients, which makes the prediction task challenging. Moreover, since disposition decisions and clinical care decisions (e.g., ordering laboratory/imaging tests and other clinical interventions) will not be completely consistent across different physicians, it is important to check how effectively data-driven machine-based prediction models can match the actual disposition decisions by exploiting clinical, demographic, and operational data.

2.2 Study setting and population

The study is based on electronic health record data collected at the ED of an academic urban level-1 trauma center, which delivers comprehensive care services including the fields of cardiology, cardiovascular surgery, neurology, neurosurgery, orthopedics and sports medicine, organ transplants, and treatments for prostate, breast and lung cancers. The data items were collected for the period from May 2014 to April 2016 and covers 184,895 patient visits. After accounting for abnormal departures

Table 1 Anticipated application and use cases for prediction results of each classification scheme

Classification scheme	Prediction classes	General description of possible utility of prediction results
C1	Admission, discharge	- Mostly within the ED (i.e., differentiate ways to provide care to the two different patient groups/types)
		- Coordination of resources that are common and shared for ED discharge
C2	IU, OU, discharge	- Coordination of services that are required regardless of the type of admitted IUs, including admission approval
		- Coordination of resources at the OU
C3	ICU, TU, GPU, OU, discharge	- Enhanced treatment for high risk patient groups (e.g., patients likely to be admitted to ICU) in the ED
		- Coordination of resources that require unit-specific coordination including prioritization in bed preparation and patient transporter dispatch and the early initiation of inpatient nursing services

(including patients leaving without completing services, transfers to other facilities, as well as patients who went to units that are not regarded as regular IUs, e.g., catheterization laboratory and perioperative unit), 172,809 patients remain in the dataset, which corresponds to 93.5% of the total visits (Fig. 2a). We provide general statistics around the study hospital and ED in Table 2 (based on information from the study data period). Because our dataset includes no patient-identifying information (PII data was removed by information technology staff before sharing the dataset), the study is exempted by the institutional review board of the study hospital.

2.3 Study protocol and methodology

We first introduce the proposed classification strategy briefly with few associated concepts from data science. The ED disposition prediction inherently presents a hierarchical, mandatory-leaf classification structure, where no two parent nodes share a common child node, and the most proper class is always found at the lowest level (Fig. 2b). The main approaches to tackle hierarchical classification problems can be classified into two main categories: "big-bang" and "top-down" [42]. While the top-down approach starts its classification task from the parent node and uses the obtained prediction outcomes for classification at its child nodes, the big-bang approach classifies the most proper class for the full problem with a single model. Given the hierarchical structure of ED disposition and the class membership imbalance at the C1 classification scheme (most patients, 75.3%, are negative [discharged] in the study hospital ED), the top-down approach would incur serious challenges at the downstream levels (C2 and C3). This is because false positive predictions occurring in the dominant negative class (discharge class in our case) at C1 propagate down the hierarchy and greatly affect the predictions for minor classes at C3. Especially, when the problem is of a mandatory-leaf node structure, false positive cases will spread throughout the hierarchical classification tree. Therefore, we choose to adopt the big-bang approach for predicting disposition decisions. Moreover, since each prediction class level could bring about their own operational benefits (as shown in Table 1), we built classification models at each level. Hence, we model and analyze the mandatory-leaf node, tree-structured ED disposition classification problem with the big-bang classifier per level approach.

Table 2 General statistics on the study hospital and ED during the data period

Item	Value
Number of beds in the hospital	877
Number of beds in ED rooms	77
Number of extra beds that ED hallways can accommodate during full occupancy	31
Annual ED visits	92,448
Proportion of ED visits admitted to IUs	17.5%
Annual discharges from IUs	35,383

In addition, we also consider the temporal aspects of ED disposition prediction. The diagnosis uncertainty decreases through the ED caregiving processes [26, 36]. Specifically, we identify four different ED caregiving epochs: "patient arrival at ED" (denoted T1-Arrival), "triage complete" (T2-Triage), "first provider encounter" (T3-Initial Assessment), and "first set of laboratory results returned" (T4-Initial Lab Results), which is graphically represented in Fig. 3. At T1-Arrival, ED patients arrive with some basic information such as arrival time/mode, prior ED visit history, health history (International Classification of Diseases, Tenth Revision, ICD-10 codes, from previous ED visits), and demographics. Then, patients go through triage processes, where the patient's vital signs and chief complaints are recorded (T2-Triage). T2-Triage is when most admission decision prediction models have been built in the literature [31–35]. At T3-Initial Assessment, ED care providers examine patients, issue orders for laboratory/imaging tests and decide on other clinical interventions. We incorporate only the first set of laboratory/imaging test items ordered and other clinical interventions provided within 30 min of the patient's first encounter with care providers, assuming that these items are dependent on examination at T3-Initial Assessment (upon examining the results from these tests, care providers can order additional tests and provide more clinical interventions downstream within the ED caregiving cycle, which are outside the scope of data employed for the T3-Initial Assessment setting). T4-Initial Lab Results indicates the time when the results for the first set of laboratory orders are fully reported. With the motivation of this study being "proactive" coordination within the ED-to-IU workflow, there is no point in making accurate disposition predictions using information that arrives too late in the ED caregiving cycle when the final disposition decision is already available or imminent. Unlike the laboratory tests, the results for radiology tests are not available electronically in a form that can be readily coded and are not included in this study. Considering the progressive nature of information accumulation throughout the ED caregiving process, the prediction model at T2-Triage incorporates predictors gained until epoch T2-Triage, the prediction model at T3-Initial Assessment incorporates predictors gained until epoch T3-Initial Assessment, and the prediction model at T4-Initial Lab Results incorporates predictors gained until epoch T4-Initial Lab Results.

Table 3 describes the time spans between these different ED care epochs in the study hospital over a two-year period (May 2014 to April 2016). Of note is the median "door to disposition decision" time of about four and a half hours for admitted ED patients and the median "disposition decision to admitted ED patient departure" (i.e., boarding) of about three hours for admitted patients.

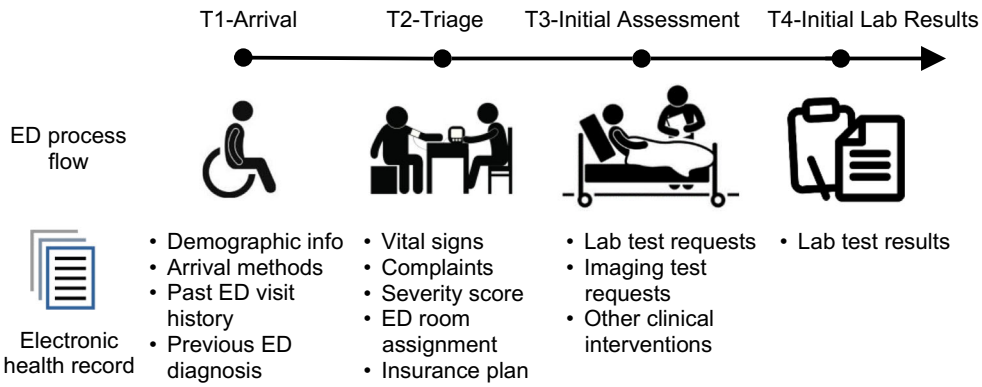

Fig. 3 Graphical representation of ED caregiving epochs and resulting data items considered in this study

While various classification modeling techniques that include multinomial logistic regression, multilayer perceptron neural networks, and support vector machines are applied, we do not see any notable difference in prediction performance that is worthy of detailed discussion in this study. Also, to enable rapid model calibration for individual hospitals and operationalization of the proposed early task initiation scheme, "simpler" and "explainable" models, such as multinomial logistic regression, should be favored over complex "black box" models (such as neural networks). Hence, we present in detail the results gained by a well-established approach, multinomial logistic regression in the Results section, and provide modeling hyperparameters tested in other machine learning models as well as their results (for the C3 classification scheme at T4-Initial Lab Results) in the Appendix (Table 11 and Table 12, respectively). Multinomial logistic regression is an extension of simple logistic regression for estimating the association between a set of predictors and a nominal outcome that has multiple categories (classes). In this study, we model logistic models that generalize the simple logistic regression by setting a baseline category among c categories (similar to the simple logistic regression that introduces a dummy category) with p number of predictors [43]. Therefore, we solve $c-1$ logit equations having $(c-1)\times p$ parameters as follows:

$$\ln\left(\frac{\pi_{ij}}{\pi_{ij'}}\right) = \alpha_j + \beta_j X, \qquad \forall j \neq j' \tag{1}$$

where j is the category index, j' is the baseline category, α_j is the intercept for category j, β_j is a vector having p coefficients for predictors, and π_{ij} is the probability that the membership of ith observation is j. The probability can be calculated by rewriting Eq. 1 as follows:

$$\pi_{ij} = \frac{\exp(\alpha_j + \beta_j X)}{1 + \sum_{k \neq j'} \exp(\alpha_j + \beta_j X)}, \qquad \forall j \neq j'. \tag{2}$$

The model finds parameter β_j by maximizing the likelihood of the data. The entire dataset is split into two parts: first 85% of the patient visits for training the models (146,888 visits) and the rest for testing (25,921 visits). We only report prediction results obtained from the testing dataset in the Results section.

2.4 Predictor variable transformation and selection

To provide details about the predictors employed by the prediction models, Table 4 presents the information on how each set of predictors is transformed and entered the models (including the feature categorization rules from the Centers for Medicare and Medicaid Services [44] and the National Library of Medicine at the National Institute of Health [46]). All the feature selection and categorization methods are applied to the training dataset to extract the feature transformation rules, then the rules are applied to the testing data for preprocessing. In particular, the chi-square test is a fundamental and well-established approach that can measure the extent of dependence (or independence) among variables having multiple class levels [47]. The χ^2 statistic of variable X is defined as follows:

$$\chi^2(X) = \sum_i \sum_j \frac{(O_{ij} - E_{ij})^2}{E_{ij}}, \tag{3}$$

Table 3 Length of time spent by patients at different ED caregiving intervals at target hospital

ED caregiving interval	Q1/median/Q3 (in minutes)
T1-Arrival to T2-Triage	11.3/17.6/29.0
T2-Triage to T3-Initial Assessment	19.5/47.5/92.8
T3-Initial Assessment to T4-Initial Lab Results	35.0/55.0/90.0
T4-Initial Lab Results to disposition decision (for admitted patients)	95.7/146.2/195.3
Admission approval to departure (for admitted patients)	102.8/182.5/297.7

Table 4 Summary of feature transformation

Predictor (ED caregiving epoch)	Feature type and its transformation	Detail
Age (T1-Arrival)	From numeric to categorical, according to the age categorization rule applied for a personal healthcare spending study conducted by the Centers for Medicare and Medicaid Services [44].	0–18, 19–44, 45–64, 65–84, and 85 and over.
Gender (T1)	Categorical (no transformation).	
Arrival time (T1)	From time to categorical, bi-hourly.	12 AM–2 AM, 2 AM–4 AM, and so on.
Arrival method (T1)	From 61 categories to 15 categories, by aggregating similar methods.	For example, fire department paramedic services provided by all different fire departments were merged into a single "fire department paramedic services" category. The categories include car, medical flight, public transportation, fire department paramedic services, Emergency Medical Services, walk in, and so on.
Last month ED visit history (T1)	Binary, whether a patient has visited the ED within last 30 days or not.	
Super utilizer indicator (T1)	Binary, whether a patient has visited the ED five or more times last year.	
Previous ED diagnosis (T1)	From ICD-10 code to 23 classes, by extracting the first letter code (type of injury or disease). Created "none" category for ED visits having no previous ED visits.	A~T, Z, and "none" categories.
Emergency severity score (T2-Triage)	Categorical, with 5 levels.	The Emergency Severity Index triage algorithm [45] is applied.
Chief complaint (T2)	From 758 categories to 168 categories, by Chi-square feature selection. Created "unspecified" category for ED visits having no complaint record.	20 example categories include edema, dizziness, dialysis, abdominal pain, fever, difficulty in breathing, nausea and vomiting, shortness of breath, migraine, blurred vision, COPD, fatigue, hematuria, pruritus, tachycardia, psychiatric evaluation, nasal congestion, hypotension, weakness, and unresponsive.
Insurance plan (T2)	From 285 categories to 14 categories, by Chi-square feature selection.	Each category consists of the names of insurance provider and benefit plan.
Vital signs (T2)	From numeric to categorical, mainly based on the normal vital sign ranges introduced by the National Library of Medicine at the National Institute of Health [46]. Created "not measured" category for each vital sign item to handle ED visits that have no data.	Temperature (°C): not measured, < 95.1, 95.1~97.7, 97.8~99.1, 99.2~100, and > 100. Pulse (rpm): not measured, <60, 60~100, and > 100. Respirations (bpm): not measured, <12, 12~18, 19~25, and > 25. Pulse oximetry (%): not measured, <95 and 95~100. Systolic blood pressure (mmHg): not measured, <90, 90~120, and > 120. Diastolic blood pressure (mmHg): not measured, <60, 60~80, and > 80. Pain score: not measured, 0~2, 3~4, 5~6, 7~8, and 9~10. Glasgow coma scale: not measured, <10, 10~11, 12~13, 14, and 15.
Assigned ED care area (T2)	Categorical, indicating one of the 6 compartmental primary care areas that a patient is assigned based on triage information.	
Laboratory test item (T3-Initial Assessment)	Binary (whether ordered or not) for each of 150 types of laboratory test items.	
Imaging test item (T3)	Binary (whether ordered or not) for each of 152 types of imaging test items.	
Other clinical intervention item (T3)	Binary (whether provided or not) for each of 108 types of clinical intervention items.	
Laboratory test result (T4-Initial Lab Results)	From numerical to pre-set categorical, for the laboratory test items for which the health information technology system provides result flags that automatically categorize numerical laboratory values into pre-set categories. Created "not ordered" category for each laboratory test result item to handle ED visits that have no data. For the laboratory test items that do not have result flags, which are minor tests, binary (whether ordered or not) variables are used as in T3-Initial Assessment.	For example, magnesium test values (mg/dL) are categorized into 5 groups: low panic ≤0.9 < low ≤1.8 < normal ≤2.3 < high ≤4.0 < high panic.

where O_{ij} is the observed frequency count for the ith level of the categorical variable X for class j, and E_{ij} is the expected frequency count for the ith level of the categorical variable X for class j. If variables entail a high χ^2 value, it infers that the

variables deviate significantly from the independence assumption and regarded relevant to each other.

A chi-square statistic-based filter method is designed and applied as a feature selection mechanism in this study. In particular, to handle the predictors having excessively many categories, we conducted feature selection in two stages, i.e., within a predictor and between predictors. First, within-predictor feature selection is performed when a predictor has more than 50 categories (including patient arrival mode, chief complaint, and insurance plan predictors) that would unnecessarily increase computational complexity during the between-predictors feature selection process (the second stage). Then, in the second stage, supervised feature selection is performed among predictors obtained through the first stage selection, by building prediction models and fitting on the training dataset. This second stage gives the proper threshold of a chi-square statistic that optimally filters in and out features based on prediction results gained with the training dataset. We used prediction accuracy as the performance measure to be maximized during the second stage feature selection.

The list of all the laboratory/imaging tests and other clinical intervention predictors is presented in detail in Table 13 in the Appendix.

3 Data analysis

Table 5 summarizes univariate statistics for demographics, emergency severity score, and chief complaints (ten representative complaints). Table 6 reports bivariate statistics (admission vs. discharge) for categorical variables including chief complaint, emergency severity score, and clinical intervention items. We report odds ratios based on whether a clinical test/intervention is given or a chief complaint is recorded for a patient or not. Table 7 presents bivariate statistics for numerical values from laboratory test results and vital signs. We recognize that in very large samples, *p*-values from *t*-tests can misinform regarding the practical significance of a variable [48]. Therefore, we computed a well-established

Table 5 Univariate analysis for demographics, emergency severity score, and chief complaints. Note: Q1, 25% quantile; Q3, 75% quantile

Variable	Median/mean/count/proportion	(Q1, Q3)	Notes
Age (median), years	45.0	(27.0, 60.0)	
Patient arrivals (median count of arrivals per hour, proportion of arrivals for the interval by day)			
12 AM~4 AM	5.0, 8.6%	(3.0, 7.0)	
4 AM~8 AM	4.0, 6.8%	(3.0, 5.0)	
8 AM~12 PM	13.0, 21.1%	(10.0, 16.0)	
12 PM~4 PM	15.0, 24.5%	(12.0, 17.3)	
4 PM~8 PM	14.0, 22.3%	(11.0, 16.0)	
8 PM~12 AM	10.0, 16.7%	(8.0, 12.0)	
Gender (proportion)			
Female	54.9%		
Male	45.1%		
Insurance plan			285 different plans
Patient arrival mode			61 different modes
High utilizer			
Visit of a patient with any prior ED visit within last 30 days	23.1%		0.4 prior visits/patient on average within last 30 days (across all patients)
Proportion of super utilizers among total patients (≥5 visits/year)	1.3%		1163 super utilizers among 90,308 patients
Proportion of super utilizers' visits among total visits (≥5 visits/year)	11.0%		19,091 super utilizers' visits among 172,809 visits
Emergency severity score (proportion of patients categorized into each severity score)			
Unspecified	0.5%		
Severity score 1 – resuscitation	2.0%		
Severity score 2 – emergent	36.3%		
Severity score 3 – urgent	54.4%		
Severity score 4 – less urgent	6.3%		
Severity score 5 – non-urgent	0.5%		
Chief complaint (proportion of patients presenting each complaint for the 10 most frequent complaints)			
Abdominal pain	10.0%		
Shortness of breath	5.5%		
Chest pain	5.5%		
Back pain	3.6%		
Headache	3.3%		
Emesis	2.4%		
Fall	2.3%		
Cough	2.3%		
Leg pain	2.1%		
Dizziness	1.9%		

Table 6 Bivariate analysis for categorical variables gathered since triage (representative items)

Variable	Discharge % of discharge samples with each variable	Admission % of admission samples with each variable	Odds ratio (95% CI)
Chief complaint (10 most frequent complaints)			
Abdominal pain	10.2%	9.3%	0.90 (0.86–0.93)
Shortness of breath	3.2%	12.6%	4.4 (4.3–4.6)
Chest pain	4.0%	10.0%	2.7 (2.6–2.8)
Back pain	4.2%	1.7%	0.38 (0.35–0.41)
Headache	3.8%	1.8%	0.47 (0.44–0.51)
Emesis	2.3%	2.6%	1.1 (1.1–1.2)
Fall	2.1%	2.8%	1.4 (1.4–1.5)
Cough	2.6%	1.5%	0.56 (0.51–0.61)
Leg pain	2.2%	1.6%	0.69 (0.63–0.75)
Dizziness	1.8%	2.2%	1.2 (1.1–1.3)
Emergency severity score			
Unspecified	0.5%	0.7%	3.3 (2.9–3.8)
Severity score 1 – resuscitation	0.7%	5.7%	18.2 (17.0–19.5)
Severity score 2 – emergent	26.6%	66.0%	5.7 (5.6–5.9)
Severity score 3 – urgent	63.3%	27.4%	Reference
Severity score 4 – less urgent	8.4%	0.2%	0.06 (0.05–0.07)
Severity score 5 – non-urgent	0.6%	< 0.1%	0.04 (0.02–0.09)
Laboratory test order made after first encounter with doctor (5 example laboratory tests)			
Troponin I (ng/mL)	8.6%	44.3%	8.3 (8.1–8.6)
BNP (pg/mL)	3.3%	25.8%	10.3 (9.9–10.7)
PT/INR/PTT (sec)	10.0%	43.4%	6.9 (6.7–7.1)
Lactate whole blood (mmol/L)	4.9%	30.2%	8.4 (8.1–8.7)
CBC with differential (K/uL)	32.4%	74.2%	6.0 (5.8–6.1)
Imaging test order made after first encounter with doctor (5 example imaging tests)			
Chest x-ray	3.5%	26.4%	9.9 (9.5–10.3)
CT head scan without contrast	3.3%	11.1%	3.7 (3.5–3.9)
CT abdomen and pelvis scan with contrast	0.86%	2.0%	2.3 (2.1–2.5)
Acute abdominal series	1.8%	3.2%	1.8 (1.7–1.9)
CT pulmonary embolism scan	0.17%	0.79%	4.8 (4.1–5.7)
Other clinical interventions after first encounter with doctor (5 example items)			
Insert saline lock	2.9%	5.7%	2.0 (1.9–2.1)
Orthostatic vital signs	2.3%	3.7%	1.6 (1.5–1.7)
Cardiac monitoring	0.9%	4.9%	5.7 (5.3–6.1)
Consult to neurology	0.4%	2.2%	5.7 (5.1–6.3)
Consult to acute care surgery	0.4%	2.0%	4.6 (4.2–5.1)

BNP, brain natriuretic peptide; bpm, breaths per minute; CBC, complete blood count; CI, confidence interval; CT, computed tomography; PT/INR/PTT, prothrombin time/international normalized ratio/partial thromboplastin time; rpm, rate per minute

alternative, Hedge's *g*, to measure the effect size recorded for the continuous variables. The rule-of-thumb thresholds for interpreting the effect size through the Hedge's *g* statistic are as follows: 0.20, small; 0.50, medium; 0.80, large [49]. According to the thresholds, respirations, pulse oximetry, BNP (brain natriuretic peptide), PT/INR/PTT (prothrombin time/international normalized ratio/partial thromboplastin time), lactate blood, and CBC (complete blood count) with differential variables present the small to medium level of effect between the two classes. Due to space constraints and the size of the feature set, we do not present data analysis results for all the variables for all the classification schemes (i.e., C1, C2, and C3). Rather, we selectively present the multivariate analysis statistics for some key features for the 5-class scheme (the C3 classification scheme). The results are reported in detail in the Results section.

4 Results

4.1 Preliminary data analysis results

We provide the univariate and bivariate analysis results obtained from the whole dataset in Table 5 through Table 7.

4.2 C1 classification scheme

As shown in Table 8, right upon arrival to the ED (at door, T1-Arrival), without any clinical information, we can predict admission decision of ED patients with 77.9% (95% confidence interval, CI, of 77.7–78.2) accuracy. While incorporating more information allows the model to enhance its performance (from T1-Arrival through T4-Initial Lab Results), the biggest improvement is made at triage (T2-Triage) where 56.3% of admitted patients are correctly predicted with less than 7% of false positives. There

Table 7 Bivariate analysis for continuous variables gathered since triage (six and five representative items for triage vital signs and laboratory test results, respectively)

Variable	Discharge		Admission		
	Mean (median)	(Q1, Q3)	Mean (median)	(Q1, Q3)	Hedge's *g* statistic (absolute value)
Triage vital sign					
Temperature (°C)	98.3 (98.2)	(97.9, 98.6)	98.4 (98.2)	(97.9, 98.6)	< 0.01
Pulse (rpm)	86.4 (85)	(74, 96)	89.5 (88)	(75, 103)	0.16
Systolic blood pressure (mmHg)	133.4 (132)	(119, 147)	137.1 (136)	(119, 154)	0.16
Diastolic blood pressure (mmHg)	78.5 (78)	(69, 87)	79.1 (78)	(67, 88)	0.05
Respirations (bpm)	18.6 (18)	(16, 20)	19.7 (18)	(18, 20)	0.22
Pulse oximetry (%)	98.2 (99)	(97, 100)	97.2 (98)	(96, 99)	0.30
Results from laboratory tests ordered at first encounter with doctor (5 example laboratory tests)					
Troponin I (ng/mL)	0.04 (0.04)	(0.04, 0.04)	0.19 (0.04)	(0.04, 0.05)	0.09
BNP (pg/mL)	185.3 (35)	(12.0, 110.0)	568.9 (128)	(33.0, 587.0)	0.41
PT/INR/PTT (sec)	15.2 (13.8)	(13.2, 14.6)	16.8 (14.4)	(13.5, 16.1)	0.20
Lactate blood (mmol/L)	1.8 (1.4)	(1.1, 2.1)	2.3 (1.7)	(1.2, 2.5)	0.28
CBC with differential (K/uL)	8.0 (7.5)	(5.8, 9.6)	9.5 (8.3)	(6.2, 11.3)	0.28

are additional information gains with the first set of laboratory/imaging test orders and other clinical intervention items (T3-Initial Assessment) and laboratory test results (T4-Initial Lab Results) that lead to further improvements in prediction quality. Unlike the admission prediction study conducted at the ED in an Israeli hospital [26], the results of laboratory tests (at T4-Initial Lab Results) seem more informative than the decisions to order specific tests and provide other clinical interventions (at T3-Initial Assessment) to enhance the predictions in our study. For reference, we also compare our prediction performance for the binary classification at triage (C1 at T2-Triage) with other admission prediction models made around triage as well as triage nurse predictions (Table 14) [31, 37, 50, 51].

4.3 C2 classification scheme

Similar to the C1 classification scheme, the C2 class scheme gains the steepest increase in prediction ability at triage (Table 9). It is noticeable that the OU class, being an intermediate class between the IU class and the discharge class, does not seem to have clear clinical distinction compared to the other classes. The laboratory test result items (T4-Initial Lab Results) prove their predictive power at the C2 level. In particular, the precision of classification for the OU class exceeds 50% utilizing laboratory test results, doubling the sensitivity at T2-Triage.

4.4 C3 classification scheme

Predictions on the C3 classification scheme provide the most actionable information among the three classification schemes for the proactive coordination of admission processes. We present prediction and analysis results for the C3 classification scheme from multiple angles to open a rich discussion.

4.4.1 Prediction quality evolvement trajectories

The C3 class classification scheme would provide the most informative results for proactive resource coordination. We choose to report prediction results graphically to clearly depict prediction quality progression for each class (Fig. 4) as well as with a table that reports numerical values of classification sensitivity and precision levels at C3 (Table 15 in the Appendix). Figure 4 reports the sensitivity (Fig. 4a) and precision (Fig. 4b) results for each class. The C3 classification scheme is an

Table 8 Comparing prediction model performance at different caregiving stages for C1 classification scheme

C1 class Accuracy (95% CI)		T1-Arrival 77.9% (77.7–78.2)	T2-Triage 84.6% (84.4–84.8)	T3-Initial Assessment 85.5% (85.2–85.7)	T4-Initial Lab Results 87.1% (86.8–87.4)
Admission (24.3%)	Sensitivity	30.2%	56.3%	58.7%	63.9%
	PPV	60.0%	74.8%	76.6%	79.2%
Discharge (75.7%)	Specificity	93.4%	93.8%	94.2%	94.6%
	NPV	80.4%	86.8%	87.5%	89.1%

PPV, positive predictive value; NPV, negative predictive value. The percentage value beside each class name reports the true prevalence of the class

Table 9 Comparing prediction model performance at different caregiving stages for C2 classification scheme

C2 class Accuracy (95% CI)		T1-Arrival 76.6% (76.3–76.8)	T2-Triage 82.3% (82.1–82.6)	T3-Initial Assessment 83.2% (82.9–83.4)	T4-Initial Lab Results 84.9% (84.6–85.2)
IU (17.7%)	Sensitivity	23.5%	56.4%	58.5%	64.6%
	Precision	49.1%	66.8%	69.9%	73.1%
OU (6.6%)	Sensitivity	0%	6.6%	9.4%	15.3%
	Precision	(no prediction)	44.3%	44.6%	55.3%
Discharge (75.7%)	Sensitivity	96.0%	95.3%	95.7%	95.8%
	Precision	79.1%	85.6%	86.2%	87.8%

Note: The term 'precision' is used for C2 and C3 schemes rather than 'PPV' of C1 binary classification since these schemes do not entail binary classification

imbalanced multiclass classification problem with a small fraction of the patients belonging to minor classes (especially TU and ICU with ~2% and ~4% in the testing sample, respectively), compared to the major class (i.e., discharge with 75.7%). At T1-Arrival, the prediction model fails to detect the minor classes, suggesting that there is insufficient information to identify patients belonging to the minor classes. At T2-Triage, prediction quality varies depending on the class. The highest levels of sensitivity and prediction are obtained for the ICU class. It indicates that ICU patients possess the most distinct features at the triage stage, and even though the number of ICU patients is small (only 3.8%), around 40% of ICU patients can be detected with about 50% precision. Triage information also carries considerable importance for GPU class prediction.

We also observe that predictions of the TU and OU classes are markedly less accurate compared to the ICU and GPU classes at T2-Triage, implying that unlike ICU and GPU patients, TU and OU patients do not have clear clinical trajectory at triage and remain as in-between states, i.e., the TU between the ICU and GPU, and the OU between the GPU and discharge. However, we can recognize that the physicians' decisions to order certain laboratory/imaging tests and provide other necessary clinical interventions prove useful for predicting the TU class (T3-Initial Assessment). While clinical care intensity and needs of TU patients are difficult to estimate at triage, physicians would start to assess the clinical severity and require care services (especially constant cardiac monitoring) for patients based on triage information and try to differentiate TU patients by ordering laboratory/imaging tests and providing other clinical interventions.

The results in Fig. 4 show that information gained from laboratory test results is generally larger than that from laboratory test order items. The precision levels of prediction are greater than 45.0% in all 5 classes at T4-Initial Lab Results, while sensitivity levels vary. The ICU and GPU classes mark the highest sensitivity levels with more than 42.0% (apart from discharged patients with 96.4%), while 36.4% of TU classes are detected. With the comparatively high-performance levels for the ICU class at T2-Triage, the additional gains of prediction ability for the ICU class at T3-Initial Assessment and T4-Initial Lab Results are not drastic compared to other minor units. The distinct progression behaviours of different classes at C3 clearly indicate that any proactive coordination strategy that utilizes ED disposition decision prediction should consider the different levels of prediction quality obtained at different care epochs for each IU.

Setting each of the four admission classes as a positive class at the C3 classification scheme, the false positive rates at T4-Initial Lab Results are only 1.4%, 1.0%, 4.3%, and 1.4%

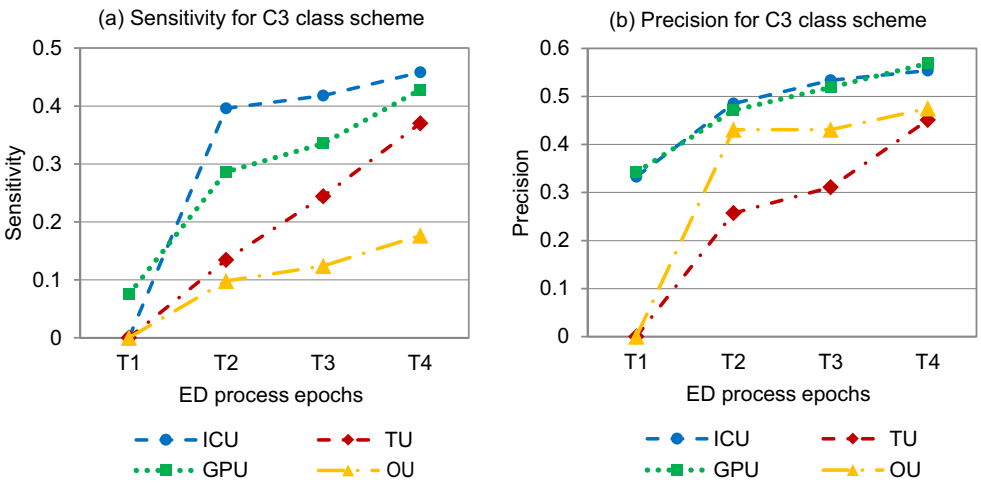

Fig. 4 Characterization of progressive nature of disposition decision predictions. Note: Each line represents each class in the C3 classification scheme except the discharge class. The two figures compare prediction model performances (sensitivity and precision, respectively) at the different caregiving epochs

for the ICU, TU, GPU, and OU classes, respectively. Also, the area under curve (AUC) values are 0.97, 0.95, 0.89, and 0.84 for the four classes, respectively. The AUC value for the discharge class is 0.92. We provide the receiver operating characteristic curves for the five classes at the C3 classification scheme at T4-Initial Lab Results (Fig. 5).

Paired *t*-tests were run for the five paired observations (five AUC values for the five classes) through the four caregiving epochs to report formal statistical test results for incremental improvement in disposition decision prediction. Fig. 6 provides the forest plots reporting the weights (i.e., the size of the black squares calculated as the inverse of the standard error values) and 95% CI of the mean difference between paired observations (the AUC values for the five classes). The *p*-values for T1-Arrival to T2-Triage, T2-Triage to T3-Initial Assessment, and T3-Initial Assessment to T4-Initial Lab Results are 0.002, 0.02, and 0.007, respectively. The set of test results conforms to our observations from Fig. 4 and Table 15.

4.4.2 Prediction threshold analysis at C3

It is important to explore how the prediction results can be further exploited to enable the effective operationalization of prediction information. Especially, being a probabilistic classifier, multinomial logistic regression outputs an estimated probability value for each of the classes for a patient as well as a predicted class. The estimated probability values are metrics to measure the level of confidence of membership at each disposition class. By imposing a probability confidence threshold in making a prediction, the model does not make any prediction unless one of the classes has a higher probability value than the threshold. Therefore, as we set higher and higher thresholds,

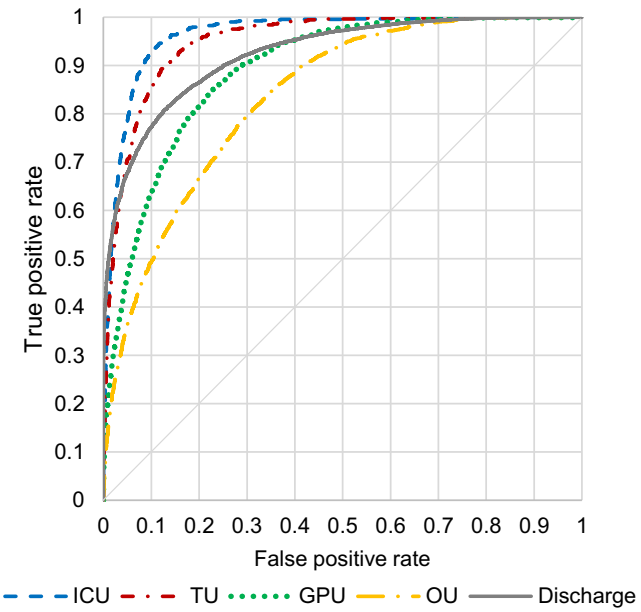

Fig. 5 ROC curves for the five classes at the C3 classification scheme at T4-Initial Lab Results

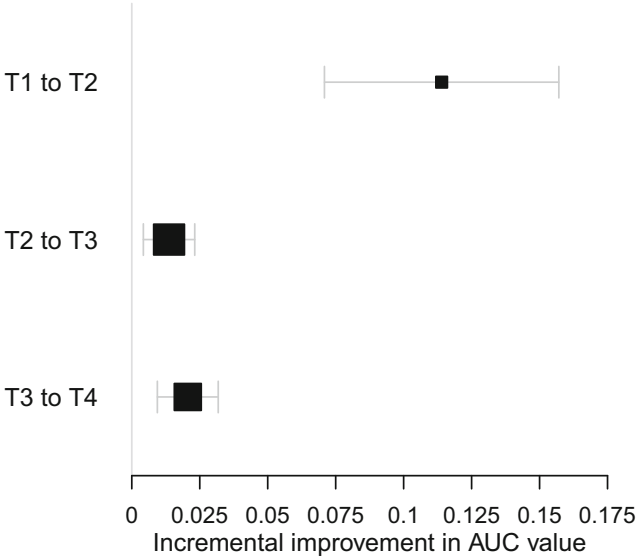

Fig. 6 Forest plots of the mean difference in the AUC values for the five classes at different intervals

predictions would become more and more reliable with increasing precision levels. Figure 7a, b show behaviours of the sensitivity and precision of prediction respectively for each class among the patients who have higher probability values than the different threshold levels. Figure 7c displays the fraction of patients remaining in the analysis with the different threshold values. For instance, if we impose 60% threshold to the ICU class, the sensitivity of ICU class prediction among the fraction of the patients increases from 45.8% to 62.1%, and the precision would increase from 55.4% to 68.8% while 82.6% of the patients remain (are included) in the analysis. It is noticeable that although the higher threshold values tend to bring higher sensitivity values for the ICU and TU class patients, it does not affect the GPU and OU class patients in the same way. This implies that the predictions of the ICU and TU classes have higher confidence (with higher probability values) than those of the GPU and OU classes because of their clinical distinctiveness.

4.4.3 Feature analysis for C3 classification scheme at T4-Initial Lab Results

We report the most statistically significant features (top 30 variables) at T4-Initial Lab Results for the C3 scheme, based on the chi-square test (Table 10). The study hospital ED has four compartmental care areas. Triage nurses determine the most appropriate care area (e.g., the main ED and fast track) for an ED patient based on the triage information of the patient. This feature presents high importance for disposition decision prediction at C3. It is also notable that 14 laboratory test result features are included in the top 30 list at T4-Initial Lab Results (the "LAB" items in Table 10). The laboratory test results provide detailed information on patients' conditions that can help predictions at the most granular level (C3). In Table 16, we

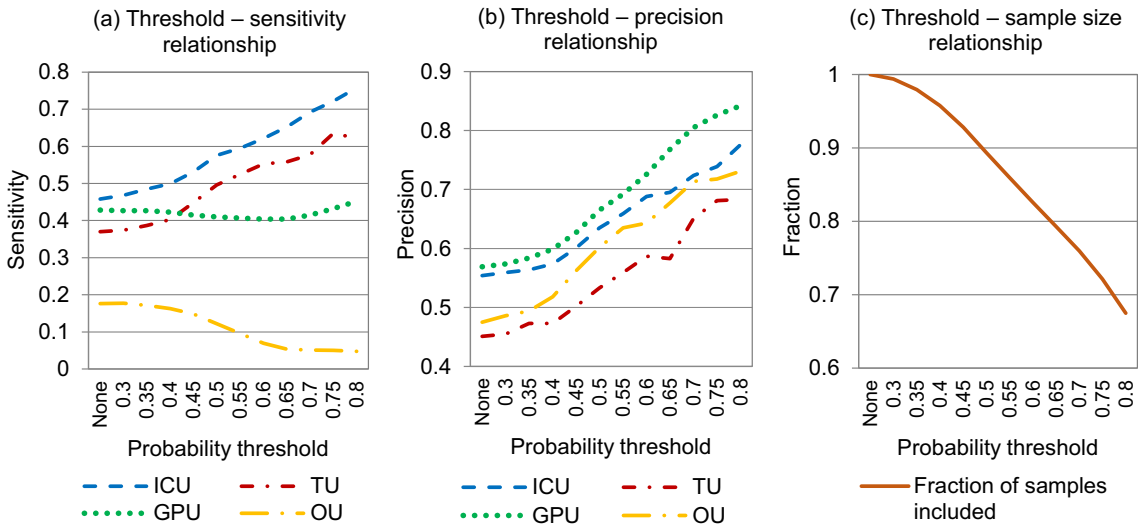

Fig. 7 Performance analysis conducted according to different threshold probability values for making predictions. Note: Figure (c) depicts the decreasing pattern of sample size as more samples are filtered out by the higher threshold values

compare the distributions of informative feature values across the five classes in the C3 classification scheme. The laboratory test result items as well as the primary care area and emergency severity score variables are selected to be presented in the table.

Table 10 Statistically important features for C3 classification scheme prediction at T4-Initial Lab Result

Order of importance	C3 classification scheme predictor
1	Primary care area
2	LAB: troponin I
3	Emergency severity score
4	LAB: PT/INR/PTT
5	LAB: BNP
6	LAB: Lactate whole blood
7	LAB: CBC with differential
8	IMG: Chest x-ray
9	LAB: Basic metabolic profile
10	Age
11	INT: ECG 12-lead
12	Insurance plan
13	LAB: Magnesium
14	LAB: Blood gas venous
15	LAB: Blood culture
16	VT: Temperature
17	LAB: Blood gas arterial
18	Arrival method
19	LAB: Liver profile
20	VT: Pulse oximetry
21	LAB: Phosphorus
22	IMG: CT head without contrast
23	CC: Chest pain
24	INT: Cardiac monitoring
25	LAB: POC glucose
26	CC: Stroke rule out
27	CC: Altered mental status
28	IMG: CT angiography head neck with contrast
29	LAB: Lipase
30	VT: Pulse

CC, chief complaint; IMG, radiology (imaging) test; INT, other intervention; LAB, laboratory test; POC: point-of-care; VT, vital sign

A probability value is highlighted in bold when the probability at a minor class exceeds 30%, or the probability of the discharge class surpasses 95%. For instance, it is noticeable that patients with "high panic" results in the troponin I test are likely to be admitted to the TU with 51% probability.

5 Discussion

We attempted to frame ED disposition prediction as an analytics problem, seeking proactive resource allocation and task initiation for potential ED admissions to reduce patient boarding. To the best of our knowledge, this work is the first attempt to define the ED disposition decision prediction as a hierarchical multiclass classification problem, categorizing the admission patients into more detailed classes so that the outcomes of the prediction can become useful for unit-specific proactive coordination across the ED-to-IU workflow. This study shows that in the study hospital, as patient information is accumulated throughout the ED caregiving process, the prediction power gradually increases with unique patterns depending on the target class. It implies that a proper proactive resource allocation and task initiation strategy should vary across the different classes (i.e., destination units) depending on their own uncertainty reduction behaviors.

To discuss the use of prediction information in more detail, we refer back to Table 1, where the general application ideas are introduced along with promising use cases. In our case study hospital, it is encouraging to see that reasonable predictive capability for the four admission classes is obtained with enough lead time for proactive coordination (i.e., around two and a half hours prior to the actual disposition decisions for the admitted patients). This finding shows that there is good potential for realizing the proactive coordination across the ED-to-IU workflow as well as the improved care delivery for patients within an ED, which are

briefly listed in Table 1. One application that operationalizes the prediction information is to send an advance bed request signal for a patient to the predicted destination unit when a prediction probability value exceeds a pre-set threshold probability. In this way, a proper inpatient bed can be identified and prepared in advance while the patient is still going through his/her remaining ED caregiving processes. This approach can significantly reduce delay in bed allocation for an admitted patient. In addition, considering the distinct trajectories of prediction capability for patients destined to different IUs, advance bed request signals would be sent to, for example, ICU more often than to TU at triage. A bed manager would wait until more information is revealed and collected (beyond triage) for patients that are likely to be admitted to TU to avoid prediction errors. A data processing/analytics platform is currently under development to be implemented at the study hospital for early inpatient bed preparation and allocation. A similar approach can be applied for providing a better coordinated transport service to admitted patients. In large hospitals, inefficient dispatch of transporters could incur serious delays, and the routing of transporters can be very complex. Therefore, it is important to assign transporters, considering currently waiting admitted patients as well as patients to be admitted soon. By acquiring advance inpatient bed request signals, a transporter dispatcher can have a better sense for performing planned transporter dispatches to enhance operational efficiency. Lastly, the acquisition of advance information on disposition decisions can improve resource allocation in other indirect ways. For instance, hospitals can better balance workload over multiple IUs by proactively mobilizing care providers, responding to projected demand. The advance information can also allow time to prepare spaces in areas that are easily congested in hospitals (such as discharge holding areas).

This study also provides insights into the relationship between physicians' disposition decisions and clinical laboratory test result values through large scale data analysis.

6 Limitations

One of the limitations of this work is that we cannot guarantee whether the models have exploited the collected information to the greatest extent. For instance, while we categorized the numerical variables such as vital signs and laboratory test results into finite levels, more advanced data driven methodologies, such as deep learning approaches, would be able to better extract features by thoroughly exploiting interrelationships between the numerical variables. These approaches might compromise the repeatability and consistency of prediction, but the methods are rapidly becoming mature with technical advances. Another limitation of our work is that there is other information in the patient's electronic health record that could be included in the models to further improve the predictive power of the models. In particular, physician "notes" record clinically significant information in various formats including text. Since the notes contain refined information that comes from the interpretation of the clinical examination of patients, the inclusion of features derived from physician notes can probably improve the prediction quality for disposition decisions. In addition, considering that radiology and other imaging tests are one of the most heavily utilized diagnostic items in the ED, future works should incorporate imaging data and/or results summaries from clinicians, when available in a timely fashion, into disposition prediction models to improve accuracy. Besides the clinical information, socio-demographic information (such as place of residence and income) can also contribute to improving disposition decision prediction quality. While this work has primarily focused on exploiting clinical information readily available in most EHR systems for building prediction models, the impact of the inclusion of socio-demographic information is worth investigating in future works. Finally, even though the study is conducted at a large level-1 trauma center using a relatively large dataset from a two-year period, it would be important to replicate the study at other hospitals. For example, while we could thoroughly incorporate all the data items into the models due to the well-developed information system adopted by the study hospital, this may not be feasible in other hospitals. However, the accrual of clinical information to power the proposed patient disposition prediction models would be comparable across most EDs. We captured the general caregiving steps in the ED for analytical discussion on the predictive capability of ED patient information in predicting disposition decisions.

7 Conclusion

This work builds on a growing body of academic and industrial literature on the usefulness and feasibility of proactive coordination in healthcare that can be empowered by predictive analytics. A proper multiclass prediction framework that considers the actual patient flow, resource requirements, and resource management practices around the ED-to-IU workflow can help with proactive early task initiation and resource allocation for admitted ED patients to reduce patient boarding. Since reactive processes are prevalent in current ED-to-IU workflow operations across most hospitals (including hospital admission approval, administrative procedures, bed management, transporter assignment and so on) the prediction results could contribute in different forms and ways depending on the specific ED-to-IU operations and practices at different hospitals.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Financial support None.

Appendix

Table 11 Model hyperparameters selected and used in classification models

Prediction model	Hyperparameters
Multinomial logistic regression	- No hyperparameters needed to set.
Neural network	- Hidden layer structure: a single-hidden-layer neural network.
	- Number of nodes in the hidden layer: tested 10, 20, 50, and 100 for each training.
	- Fitting method: least-squares.
	- Weight decay (the regularization parameter): tested 0.001, 0.01, 0.1, 0, and 1 for each training.
Support vector machine	- Kernel type: radial basis function.
	- Cost of constraint violation (the constant of the regularization term in the support vector Lagrange formulation): tested 0.01, 0.1, 1, 10, and 100 for each training.
	- Tolerance of termination criterion: 0.001
	- Epsilon value in the insensitive loss function of support vectors: 0.1

Table 12 Comparison of prediction results gained by different classification techniques at C3 classification scheme

C3 class at T4 Accuracy (95% CI)		Multinomial logistic regression 81.6% (81.2–82.0)	Support vector machine 81.4% (80.9–81.8)	Multilayer perceptron neural network 81.2% (80.9–81.4)
ICU (3.6%)	Sensitivity	45.8%	48.7%	45.4%
	Precision	55.4%	52.5%	56.2%
TU (2.2%)	Sensitivity	37.0%	36.5%	38.0%
	Precision	45.1%	42.1%	42.5%
GPU (11.9%)	Sensitivity	42.8%	40.3%	41.9%
	Precision	56.9%	57.9%	56.2%
OU (6.6%)	Sensitivity	17.7%	15.5%	16.0%
	Precision	47.5%	46.2%	46.0%
Discharge (75.7%)	Sensitivity	96.2%	96.4%	96.2%
	Precision	86.9%	86.6%	86.8%

Table 13 List of laboratory test, imaging test, and other clinical intervention predictors used to build prediction model at T3-Initial Assessment (including only items) and T4-Initial Lab Results (including available result information of laboratory tests). Note: For the laboratory test results (at T4-Initial Lab Results) that have categories pre-set by the health information technology system in the study hospital, the pre-set categories were entered into the prediction model. While the table lists all the test items included in the prediction modeling, some test items that have almost identical test names were merged as a single item

Laboratory test items (with or without results)	Imaging test items (without results)	Other clinical interventions (without results)
Only test item: Blood culture, urinalysis, fluid culture non-CSF (anaerobic/aerobic) with gram stain, cell count and differential fluid, albumin fluid, glucose fluid, lactate dehydrogenase fluid, protein fluid, urine culture, strep A screen rapid, strep A culture throat, miscellaneous test A, stool culture, sodium urine, creatinine urine, hepatitis B surface antigen, respiratory culture, hemoglobin total, wound/-abscess/ drainage culture (aerobic), fungal culture blood, potassium urine, chloride urine, urea nitrogen urine, antinuclear antibodies screen and titer, cell count and differential CSF, CSF culture, Rhogam evaluation, anaerobic culture, *Clostridium difficile* polymerase chain reaction, emergency group O red blood cells, streptococcus culture, vaginal wet mount, herpes simplex virus culture, crystals fluid, gonorrhoeae culture, respiratory syncytial virus antigen, carcinoembryonic antigen peritoneal fluid, vital culture, rapid plasma and reagin quantitative. **Binary results (such as positive/ negative or normal/abnormal):** Liver profile, salicylate level, beta-hydroxybutyrate, osmolality urine, B-type natriuretic peptide, urinalysis microscopic, pregnancy urine meter, HIV antibody 1 and 2 rapid, *Neisseria gonorrhoeae* rRNA, chlamydia trachomatis rRNA, C-reactive protein, hemoglobin A1c, blood alcohol (serum ethanol) quantitative, D-dimer, bilirubin direct, AST/SGOT, ALT/SGPT, uric acid, infectious mononucleosis antibody screen, influenza A and B antigens, amylase, hepatitis screen acute, D-dimer quantitative, leudocytes stool, Clostridium difficile toxin, exposure panel HIV 1 and 2, legionella antigen, HIV-1 and HIV-2 antibodies, treponemal antibody TPPA, exposure panel hepatitis B surface antigen, exposure panel hepatitis C antibody, T lymphocyte helper/suppressor, CKMB, prealbumin, carboxyhemoglobin, hepatitis screen, rubella antibody IgG, sickle cell test, and HIV antigen/antibody. **Three result categories (categorized according to numerical levels or discrete outcomes):** Troponin I, prothrombin time/INR, PTT, calcium ionized, acetaminophen level, lactate whole blood, blood gas venous, BHCG serum, PT/INR/PTT, TSH, LDH total, ferritin, albumin, microscopic ova and parasites, erythrocyte sedimentation rate, rapid plasma reagin, CPK total, lactic acid plasma, lipid profile, ammonia, T4 free, coagulation screen, FK506 level, thyroid screen, bilirubin total, haptoglobin, reticulocyte, pregnancy urine qualitative, confirm ABO/RH type, phenytoin level free, hemoglobin and hematocrit,	**Only test item:** CT head without contrast, CT head orbit without contrast, CT head cervical spine without contrast, CT head brain without contrast CT maxillofacial without contrast, CT head sinuses without contrast, knee left 2 views, knee left 3 views, knee left 4 or more views, knee right 2 views, knee right 3 views, knee right 4 or more views, chest 1 view, chest 2 views, lower extremity venous imaging, abdomen 2 views, US OB < 14 weeks abdomen single fetus, US OB < 14 weeks abdomen and transvaginal single fetus, US OB limited, US OB transvaginal, US OB > 14 weeks abdomen single fetus, US OB follow-up single fetus, CT 3D reconstruction of body, forearm right 2 views, forearm left 2 views, US pelvis non-OB and US transvaginal non-OB Doppler, US pelvis complete non-OB and US transvaginal non-OB, CT lumbar spine without contrast, abdomen 1 view, VP shunt chest 2 views abdomen, CT pulmonary embolism, echocardiogram, CT abdomen pelvis without contrast, foot left 2 views, foot left 3 views, foot right 2 views, foot right 3 views, lumbar sacral spine 2 or 3 views, femur right 2 views, CT cervical spine without contrast, US abdomen complete, MRI brain with and without contrast, hip left 2 views lateral and anteroposterior pelvis, CT angiography abdomen pelvis with contrast, CT angiography abdomen pelvis with runoff, CT renal stone, CT chest without contrast, CT chest abdomen pelvis with contrast with or without 2D construction, CT chest with and without abdomen pelvis with contrast, CT chest abdomen pelvis with contrast with or without 2D construction, CT chest abdomen pelvis without contrast with or without 2D construction, hand right 3 views, hand left 3 views, CT thoracic spine without contrast, CT thoracic lumbar spine reformat with contrast, CT thoracic lumbar spine without contrast, CT thoracic aortic dissection with and without contrast, acute abdominal series, cervical spine 2–3 views, CT face with contrast, CT face without contrast, shoulder left 2 views, shoulder left 3 views, shoulder right 2 views, shoulder right 3 views, hip left 1 view, hip left 1 view with pelvis, hip right 1 view, hip right 2 views lateral and anteroposterior pelvis, femur left 2 views, elbow left 3 views, CT angiography head with contrast, CT angiography neck with contrast, CT angiography head neck with contrast, wrist left 3 views, wrist right 3 views, wrist right scaphoid 1 view, humerus right 2 views, soft tissue neck, nuclear medicine ventilation-perfusion lung scan with chest x-ray if needed, elbow right 2 views, CT	**Only intervention item:** Nursing communication, ECG 12-lead, check temperature, pelvic exam chaperone, anticoagulation, pulse oximetry, apply heat to affected area, consultations, initiate low intensity heparin protocol, incision and drainage, insert saline lock, laceration repair, central line, transfuse red blood cell, low extremity arterial testing, ventilator, pharmacologic venous thromboembolism prophylaxis, applying sling, straight catheter for urine collection, vital signs, insert arterial line, insert Foley catheter, visual acuity screening, bladder scan, blood transfusion reactions, orthostatic vital signs, check pulse oximetry while ambulating, pelvic examination setup, open reduction and internal fixation of proximal humerus, orthopedic injury treatment, cardiac monitoring, check blood glucose, check blood pressure, check respiratory rate, adult non-invasive ventilation, oxygen device 2 l per minute (nasal cannula), peak flow measurement pre/post, strep A screen, elevate extremity, apply ace wrap, ice pack, pre-procedural sedation, post-void residual urine test, intubation, nasogastric tube insertion/maintenance, blood pressure check on specific side/area, soap suds enema, oral hydration, cervical collar hard adjustable, incentive spirometry nursing, feed patient, crutches, undress patient, ice to affected area, focused assessment with sonography in trauma, bed rest, nursing dysphagia screen, emergency department nasal cannula oxygen, pharmacy to dose vancomycin, measuring fluid intake and output, wound care, initiate high intensity heparin protocol, nerve block, straight catheter, apply soft cervical collar, urinary catheter, foreign body removal, abdominal binder, apply warming blanket, initiate stroke protocol, hot pack or cold pack, weigh patient, peak expiratory flow rate, initiate sepsis protocol, eye irrigation, Glasgow coma scale, and monitor end tidal CO_2.

Table 13 (continued)

Laboratory test items (with or without results)	Imaging test items (without results)	Other clinical interventions (without results)
cortisol, iron and TIBC, triglyceride, osmolality, B12 and folate serum, vitamin B-12, glucose CSF, protein CSF, levetiracetam level, HIV-1 RNA quantitative, reflexed test order: FT4, carbamazepine level, and hemoglobin evaluation. **Four result categories (categorized according to numerical levels or discrete outcomes):** Lipase, CBC, emergency toxicology screen, blood gas arterial, drug screen urine, phenytoin level total, fibrinogen, digoxin level, lithium level, valproic acid level, and phenobarbital level. **Five result categories (categorized according to numerical levels or discrete outcomes):** Basic metabolic profile, magnesium, phosphorus, CBC and differential, potassium, comprehensive metabolic profile, and electrolytes whole blood. **Six result categories and more (categorized according to numerical levels or discrete outcomes):** ABO group and RH type, antibody screen, glucose, and vaginosis screen.	cervical spine without contrast, MRI lumbar spine with and without contrast, ankle left 3 views, angle right 2 views, ankle right 3 views, ankle right gravity stress 1 view, CT 3D reconstruction neuro on independent workstation, elbow left 2 views, MRI cervical spine with and without contrast, MRI cervical spine without contrast, CT pulmonary embolism dissection chest with contrast, CT soft tissue neck with contrast, tibia and fibula right 2 views, humerus left 2 views, sacrum and coccyx 2 views, feeding tube check, tibia and fibula left 2 views, MRA head without contrast, US abdomen limited, US kidneys complete, ribs left unilateral 2 views, ribs right unilateral 2 views, US soft tissue groin, elbow right 3 views, CT soft tissue neck without contrast, MRI lumbar spine without contrast, US scrotum Doppler with color and single-photon emission computed tomography analysis, fingers left 2 views, finger right 2 views, MRA neck without contrast, US kidneys transplant (includes Doppler), US pelvis complete, pelvis 1 or 2 views, CT pelvis with contrast, CT pelvis without contrast, CT lumbar spine without contrast, thoracic spine 3 views, CT pulmonary abdomen pelvis with contrast, US breasts performed in emergency department, CT lumbar spine reformat with contrast, CT lumbar spine reformat without contrast, toes left 2 views, toes right 2 views, US appendix, US abdomen kidney complete, thoracic spine 2 views, pelvis 3 views or more, clavicle left, clavicle right, gastrostomy tube change with fluoroscopy including contrast injection, US soft tissue abdominal wall, CT orbit with contrast, CT orbit without contrast, hip right 1 view with pelvis, orghopantogram (panorex jaw), mandible 4 views, ribs left unilateral 2 views with posteroanterior chest, US aorta, US scrotum with or without color flow, nasal bones 3 views, calcaneus left 2 views, and calcaneus right 2 views.	

2D, 2-dimensional; 3D, 3-dimensional; ALT/SGPT, alanine aminotransferase/serum glutamic pyruvic transaminase; AST/SGOT, aspartate aminotransferase/serum glutamic-oxaloacetic transaminase; CKMB, creatine kinase-muscle/brain; CO_2, carbon dioxide; CPK, creatine phosphokinase; CSF, cerebrospinal fluid; CT, computerized tomography; ECG, Electrocardiogram; FK506, tacrolimus; FT4, free thyroxine; HIV, human immunodeficiency virus; LDH, lactate dehydrogenase; MRA, magnetic resonance angiography; MRI, magnetic resonance imaging; OB, obstetric; PTT, partial thromboplastin time; (r)RNA, (ribosomal) ribonucleic acid; T4, thyroxine; TIBC, total iron binding capacity; TPPA, Treponema pallidum particle agglutination; TSH, thyroid stimulating hormone; US, ultrasound

Table 14 Comparison of prediction performances from selected admission prediction works done around triage, which is the most commonly selected caregiving epoch in EDs for admission prediction

Study	Method	Sensitivity	Specificity	PPV	NPV
Present study	Logistic regression	56.3%	93.8%	74.8%	86.8%
Sun et al. [31]	Logistic regression	33.4%	96.8%	81.6%	71.8%
Walsh et al. [37]	Artificial neural network	78.0%	82.0%	68.0%	89.0%
Stover-Baker et al. [50]	Expert (triage nurse)	75.6%	84.5%	62.2%	91.1%
Arslanian-Engoren [51]	Expert (triage nurse)	57.0%	59.0%	68.0%	56.0%

We select studies that have reported all the four performance measures for comparison. While it is difficult to directly compare different studies due to various factors (e.g., the proportion of admitted patients and studied patient groups) affecting the performance levels, high PPV and NPV values are essential for the reliable prediction of admissions and discharges

Table 15 Disposition decision prediction results comparing prediction model performance at different caregiving epochs for C3 classification scheme

C3 class		T1-Arrival	T2-Triage	T3-Initial Assessment	T4-Initial Lab Results
Accuracy (95% CI)		75.5% (75.2–75.8)	78.8% (78.5–79.1)	79.6% (79.3–79.9)	81.6% (81.2–82.0)
ICU (3.6%)	Sensitivity	< 1.0%	39.6%	41.8%	45.8%
	Precision	33.3%	48.5%	53.4%	55.4%
TU (2.2%)	Sensitivity	0%	13.4%	24.4%	37.0%
	Precision	(no prediction)	25.7%	31.1%	45.1%
GPU (11.9%)	Sensitivity	7.5%	28.6%	33.5%	42.8%
	Precision	34.2%	47.1%	51.9%	56.9%
OU (6.6%)	Sensitivity	0%	9.8%	12.4%	17.7%
	Precision	(no prediction)	43.1%	43.0%	47.5%
Discharge (75.7%)	Sensitivity	98.9%	96.7%	96.3%	96.2%
	Precision	76.6%	83.9%	84.9%	86.9%

Table 16 Distribution of important feature values over the C3 classification scheme classes

Feature values	Number of cases (172,809 in total)	ICU (0.04)	TU (0.03)	GPU (0.12)	OU (0.06)	Discharge (0.75)
(a) Primary care area						
Area 1	28,898	0.21	0.15	0.21	0.14	0.29
Area 2	47,008	< 0.01	0.01	0.14	0.07	0.77
Area 3	42,453	< 0.01	0.01	0.13	0.06	0.79
Area 4	53,086	< 0.01	< 0.01	0.01	0.01	**0.98**
(b) LAB: Troponin I						
High panic	1545	**0.33**	**0.51**	0.13	0.01	0.03
High	4610	0.20	0.25	**0.30**	0.11	0.14
Normal	24,037	0.09	0.08	0.21	0.18	0.44
(c) Emergency severity score						
Severity score 1	3118	**0.43**	0.07	0.17	0.06	0.28
Severity score 2	60,373	0.07	0.07	0.20	0.11	0.55
Severity score 3	95,455	0.01	0.01	0.07	0.04	0.88
Severity score 4	11,995	< 0.01	< 0.01	< 0.01	< 0.01	**0.99**
Severity score 5	880	0	< 0.01	<0.01	0	**0.99**
(d) LAB: PT/INR/PTT						
High	11,984	0.17	0.14	**0.31**	0.09	0.29
Normal	19,293	0.10	0.07	0.20	0.14	0.49
Low	221	0.15	0.07	0.22	0.12	0.44
(e) LAB: Lactate whole blood						
High	7458	0.26	0.09	0.31	0.08	0.26
Normal	11,817	0.11	0.08	**0.33**	0.10	0.38
Low	7	**0.57**	0.14	0.29	0	0
(f) LAB: CBC with differential						
High panic	123	**0.40**	0.05	**0.43**	0.03	0.09
High	16,052	0.13	0.06	0.29	0.08	0.44

Table 16 (continued)

Feature values	Number of cases (172,809 in total)	ICU (0.04)	TU (0.03)	GPU (0.12)	OU (0.06)	Discharge (0.75)
Normal	54,955	0.05	0.06	0.17	0.10	0.61
Low	2751	0.06	0.06	0.25	0.09	0.54
(g) Basic metabolic profile						
High panic	100	**0.58**	0.08	**0.31**	0.01	0.02
High	862	0.19	0.08	0.28	0.06	0.39
Normal	63,736	0.06	0.06	0.17	0.10	0.60
Low	9224	0.11	0.07	**0.31**	0.08	0.42
Low panic	274	**0.40**	0.10	**0.42**	0.03	0.04
(h) LAB: Magnesium						
High panic	15	**0.79**	0	0.21	0	0
High	417	0.23	0.10	**0.35**	0.07	0.24
Normal	1572	0.08	0.09	0.22	0.12	0.49
Low	641	0.10	0.09	0.27	0.11	0.43
Low panic	12	0.18	0.18	**0.35**	0.15	0.14
(i) LAB: Blood gas venous						
High	3958	0.14	0.08	**0.32**	0.10	0.36
Normal	3001	0.13	0.09	**0.35**	0.10	0.34
Low	8807	0.12	0.10	**0.36**	0.10	0.32
(j) LAB: Blood culture						
Measured	5705	0.21	0.08	**0.47**	0.06	0.18
(k) Blood gas arterial						
High	1993	**0.39**	0.05	0.22	0.05	0.28
Normal	867	0.25	0.06	0.22	0.05	0.42
Low	649	**0.32**	0.09	0.27	0.06	0.27
Low panic	504	**0.40**	0.09	**0.32**	0.04	0.16
(l) LAB: Liver profile						
High	6469	0.15	0.07	**0.30**	0.08	0.40
Normal	12,634	0.07	0.05	0.22	0.09	0.58
(m) LAB: Phosphorus						
High	1562	0.25	0.08	**0.37**	0.05	0.25
Normal	6801	0.10	0.09	0.26	0.11	0.44
Low	814	0.11	0.08	**0.34**	0.09	0.37
Low panic	46	0.20	0.09	**0.41**	0.09	0.22
(n) LAB: POC Glucose						
Abnormal	290	**0.43**	0.13	0.19	0.05	0.20
High panic	1005	0.15	0.07	0.25	0.07	0.46
High	228	0.09	0.05	0.16	0.07	0.62
Normal	12,679	0.08	0.06	0.19	0.10	0.56
Low	14	0.07	0	0.29	0.07	0.57
Low panic	46	0.28	0.09	0.20	0.11	0.33
(o) LAB: Lipase						
High panic	530	0.12	0.06	**0.54**	0.06	0.22
High	750	0.11	0.05	**0.31**	0.09	0.43
Normal	14,058	0.07	0.05	0.20	0.09	0.60
Low	1	0	**1.00**	0	0	0

References

1. Liu S, Hobgood C, Brice JH (2003) Impact of critical bed status on emergency department patient flow and overcrowding. Acad Emerg Med 10:382–385
2. Hoot NR, Aronsky D (2008) Systematic review of emergency department crowding: causes, effects, and solutions. Ann Emerg Med 52:126–136
3. Pines JM, Batt RJ, Hilton JA, Terwiesch C (2011) The financial consequences of lost demand and reducing boarding in hospital emergency departments. Ann Emerg Med 58:331–340
4. Carter EJ, Pouch SM, Larson EL (2014) The relationship between emergency department crowding and patient outcomes: a systematic review. J Nurs Scholarsh 46:106–115
5. King DL, Ben-Tovim DI, Bassham J (2006) Redesigning emergency department patient flows: application of lean thinking to health care. Emerg Med Australas 18:391–397
6. Kelly AM, Bryant M, Cox L, Jolley D (2007) Improving emergency department efficiency by patient streaming to outcomes-based teams. Aust Health Rev 31:16–21
7. Ben-Tovim DI, Bassham JE, Bennett DM, Dougherty ML, Martin MA, O'Neill SJ, Sincock JL, Szwarcbord MG (2008) Redesigning care at the Flinders medical Centre: clinical process redesign using "lean thinking". Med J Aust 188:S27cS31
8. Rodi SW, Grau MV, Orsini CM (2006) Evaluation of a fast track unit: alignment of resources and demand results in improved satisfaction and decreased length of stay for emergency department patients. Qual Manag Health Care 15:163–170
9. Considine J, Kropman M, Kelly E, Winter C (2008) Effect of emergency department fast track on emergency department length of stay: a case-control study. Emerg Med J 25:815–819
10. Dickson EW, Singh S, Cheung DS, Wyatt CC, Nugent AS (2009) Application of lean manufacturing techniques in the emergency department. J Emerg Med 37:177–182
11. Holden RJ (2011) Lean thinking in emergency departments: a critical review. Ann Emerg Med 57:265–278
12. 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
13. Pines JM, Hilton JA, Weber EJ, Alkemade AJ, Al Shabanah H, Anderson PD, Bernhard M, Bertini A, Gries A, Ferrandiz S, Kumar VA (2011) International perspectives on emergency department crowding. Acad Emerg Med 18:1358–1370
14. Pines JM, Bernstein SL (2015) Solving the worldwide emergency department crowding problem–what can we learn from an Israeli ED? Isr J Health Policy 4:52
15. Van Der Linden MC, Khursheed M, Hooda K, Pines JM, Van Der Linden N (2007) Two emergency departments, 6000 km apart: differences in patient flow and staff perceptions about crowding. Int Emerg Nurs 35:30–36
16. Rabin E, Kocher K, McClelland M, Pines J, Hwang U, Rathlev N, Asplin B, Trueger NS, Weber E (2012) Solutions to emergency department 'boarding' and crowding are underused and may need to be legislated. Health Affair 31:1757–1766
17. Kim MJ, Park JM, Je SM, You JS, Park YS, Chung HS, Chung SP, Lee HS (2012) Effects of a short text message reminder system on emergency department length of stay. Int J Med Inform 81:296–302
18. Zhou JC, Pan KH, Zhou DY, Zheng SW, Zhu JQ, Xu QP, Wang CL (2012) High hospital occupancy is associated with increased risk for patients boarding in the emergency department. Am J Med 125: 416–4e1
19. Centers for Medicare and Medicaid Services (2018) Hospital Compare datasets. https://data.medicare.gov/data/hospital-compare/. Accessed December 11, 2018
20. Saghafian S, Hopp WJ, Van Oyen MP, Desmond JS, Kronick SL (2012) Patient streaming as a mechanism for improving responsiveness in emergency departments. Oper Res 60:1080–1097
21. Shi P, Chou MC, Dai JG, Ding D, Sim J (2015) Models and insights for hospital inpatient operations: time-dependent ED boarding time. Manag Sci 62:1–28
22. Osorio C, Bierlaire M (2009) An analytic finite capacity queueing network model capturing the propagation of congestion and blocking. Eur J Oper Res 196:996–1007
23. Peck JS, Benneyan JC, Nightingale DJ, Gaehde SA (2012) Predicting emergency department inpatient admissions to improve same-day patient flow. Acad Emerg Med 19:E1045–E1054
24. Peck JS, Benneyan JC, Nightingale DJ, Gaehde SA (2014) Characterizing the value of predictive analytics in facilitating hospital patient flow. IIE Trans Healthc Syst Eng 4:135–143
25. Qiu S, Chinnam RB, Murat A, Batarse B, Neemuchwala H, Jordan W (2015) A cost sensitive inpatient bed reservation approach to reduce emergency department boarding times. Health Care Manag Sci 18:67–85
26. Barak-Corren Y, Israelit SH, Reis BY (2017) Progressive prediction of hospitalisation in the emergency department: uncovering hidden patterns to improve patient flow. Emerg Med J 34:308–314
27. Bayley MD, Schwartz JS, Shofer FS, Weiner M, Sites FD, Traber KB, Hollander JE (2005) The financial burden of emergency department congestion and hospital crowding for chest pain patients awaiting admission. Ann Emerg Med 45:110–117
28. Yancer DA, Foshee D, Cole H, Beauchamp R, de la Pena W, Keefe T, Smith W, Zimmerman K, Lavine M, Toops B (2006) Managing capacity to reduce emergency department overcrowding and ambulance diversions. Joint Comm J Qual Patient Saf 32:239–245
29. Thomas BG, Bollapragada S, Akbay K, Toledano D, Katlic P, Dulgeroglu O, Yang D (2013) Automated bed assignments in a complex and dynamic hospital environment. Interfaces 43:435–448
30. Tortorella F, Ukanowicz D, Douglas-Ntagha P, Ray R, Triller M (2013) Improving bed turnover time with a bed management system. J Nurs Admin 43:37–43
31. Sun Y, Heng BH, Tay SY, Seow E (2011) Predicting hospital admissions at emergency department triage using routine administrative data. Acad Emerg Med 18:844–850
32. Boyle J, Jessup M, Crilly J, Green D, Lind J, Wallis M, Miller P, Fitzgerald G (2012) Predicting emergency department admissions. Emerg Med J 29:358–365
33. Cameron A, Rodgers K, Ireland A, Jamdar R, McKay GA (2015) A simple tool to predict admission at the time of triage. Emerg Med J 32:174–179
34. Dinh MM, Russell SB, Bein KJ, Rogers K, Muscatello D, Paoloni R, Hayman J, Chalkley DR, Ivers R (2016) The Sydney triage to admission risk tool (START) to predict emergency department disposition: a derivation and internal validation study using retrospective state-wide data from New South Wales, Australia. BMC Emerg Med 16:46
35. Handly N, Thompson DA, Venkat A (2013) Derivation and validation of a hospital admission prediction model adding coded chief complaint to demographic, emergency department operational and patient acuity data available at emergency department triage using neural net methodology. Ann Emerg Med 62:S138
36. Kim SW, Li JY, Hakendorf P, Teubner DJ, Ben-Tovim DI, Thompson CH (2014) Predicting admission of patients by their presentation to the emergency department. Emerg Med Australas 26:361–367
37. Walsh P, Cunningham P, Rothenberg SJ, O'Doherty S, Hoey H, Healy R (2004) An artificial neural network ensemble to predict disposition and length of stay in children presenting with bronchiolitis. Eur J Emerg Med 11:259–264

38. Gordon PC, Wanklin JM, Harvey NH, Hatcher GH (1966) An approach to patient care classification. Can Med Assoc J 95: 1228–1236
39. Bhutta ZA, Khan I, Salat S, Raza F, Ara H (2004) Reducing length of stay in hospital for very low birthweight infants by involving mothers in a stepdown unit: an experience from Karachi (Pakistan). BMJ 329:1151–1155
40. Marra AR, D'Arco C, de Arruda Bravim B, Martino MD, Correa L, Silva CV, Lamblet LC, Junior MS, de Lima G, Guastelli LR, Barbosa L (2008) Controlled trial measuring the effect of a feedback intervention on hand hygiene compliance in a step-down unit. Infect Cont Hosp Epidemiol 29:730–735
41. van Sluisveld N, Hesselink G, van der Hoeven JG, Westert G, Wollersheim H, Zegers M (2015) Improving clinical handover between intensive care unit and general ward professionals at intensive care unit discharge. Intensive Care Med 41:589–604
42. Silla CN Jr, Freitas AA (2011) A survey of hierarchical classification across different application domains. Data Min Knowl Disc 22:31–72
43. Vittinghoff E, Glidden DV, Shiboski SC, McCulloch CE (2011) Regression methods in biostatistics: linear, logistic, survival, and repeated measures models. Springer Science & Business Media
44. Centers for Medicare and Medicaid Services (2017) National Health Expenditure Data. https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/index.html. Accessed December 1, 2018
45. Gilboy N, Tanabe T, Travers D, Rosenau AM (2011) Emergency severity index (ESI): a triage tool for emergency department care, version 4. Agency for Healthcare Research and Quality Publication No, Rockville, pp 12–0014
46. National Library of Medicine (2015) Vital signs. https://medlineplus.gov/vitalsigns.html. Accessed December 15, 2018
47. Forman G (2003) An extensive empirical study of feature selection metrics for text classification. J Mach Learn Res 3:1289–1305
48. Lin M, Lucas HC Jr, Shmueli G (2013) Research commentary—too big to fail: large samples and the p-value problem. Inf Syst Res 24(4):906–917
49. Cohen J (2013) Statistical power analysis for the behavioral sciences. Routledge
50. Stover-Baker B, Stahlman B, Pollack M (2012) Triage nurse prediction of hospital admission. J Emerg Nurs 38:306–310
51. Arslanian-Engoren C (2004) Do emergency nurses' triage decisions predict differences in admission or discharge diagnoses for acute coronary syndromes? J Cardiovasc Nurs 19:280–286