2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Novel Dynamic Prediction of Daily Patient Discharge in Acute and Critical Care

Arad Lajevardi-Khosh*, Ali Jalali, Kuldeep Singh Rajput and Nandakumar Selvaraj

Abstract— **Determining when a patient can be discharged from a care setting is critical to optimize the utilization and delivery of timely care. Furthermore, timely discharge can lead to better clinical outcomes by effectively mitigating the prolonged length of stay in a care environment. This paper presents a novel algorithm for the prediction of likelihood of patient discharge within the next 24 or 48 hours from acute or critical care environments on a daily basis. Continuous patient monitoring and health data obtained from acute hospital at home environment (n=303 patients) and a critical care unit environment (n=9,520 patients) are retrospectively used to train, validate and test numerous machine learning models for dynamic daily predictions of patients discharge. In the acute hospital at home environment, the area under the receiver operating characteristic (AUROC) curve performance of a top XGBoost model was 0.816 ± 0.025 and 0.758 ± 0.029 for daily discharge prediction within 24 hours and 48 hours respectively. Similar independent prediction models from the critical care environment resulted in relatively a lower AUROC for likewise predicting daily patient discharge. Overall, the results demonstrate the efficacy and utility of our novel algorithm for dynamic predictions of daily patient discharge in both acute– and critical care healthcare settings.**

I. Introduction

Prolonged length of stay is shown to be correlated to severe adverse events, increased cost, and poor clinical outcomes [1], [2]. Hence, determining whether or not a patient should be discharged from a care environment is an important decision for patients, clinicians, health care systems, and insurance providers. From the medical care team's perspective, discharging a patient will free up resources and allow for optimal hospital utilization and care delivery for other more critical patients [3]. In addition to clinicians and nurses time, optimal discharge time can also free up physical resources such as medical equipment [4]. From a patient's perspective, optimal discharge time will allow them to return to normal daily activities which is often their primary goal. Finally, health care providers may benefit from optimal discharge time by allowing for an improved estimate of medical care costs.

Currently, clinicians utilize a combination of objective and subjective criteria to determine if a patient should or should not be discharged. Some objective factors that clinicians often consider when making a patient discharge decision are the stability of a patient's weight, vital signs, or normality of biological lab measurements [5]. More subjective factors include hospital management style or day of the week [6].

Authors are with Biofourmis Inc, Boston, MA 02110, USA.
(*correspondence e-mail: arad.l@biofourmis.com)

Traditionally, patient discharge prediction has been approached with a focus on a patient's condition at the beginning of care. In other words, many previous studies have predicted the length of stay in a care environment based upon data available within the first 24-48 hours [7], [8]. Although this approach can provide an initial estimate for resource utilization and treatment approaches, its accuracy is limited by an inability to capture patient's improvements or deteriorations during the hospital stay. Utilizing new patient health data as it becomes available will lead to an improved ability to optimally predict patient discharge from a patient's dynamically changing health status which could help clinicians provide a better care for their patients and hospitals to manage their resources more optimally.

In this work, we propose a novel machine learning method to dynamically predict if a patient will be discharged from a care environment within the next 24 or 48 hours on a daily basis. Towards this goal, we retrospectively develop and validate independent machine learning classifier models to predict patient discharge probabilities utilizing the available patient health data in an emerging hospital at home (HH) acute setting and also a traditional critical care hospital setting.

II. Methods

A. Data Sets

1) Hospital at Home Cohort:

The patient data used in this study were collected from a HH program in which acute care services usually associated with the traditional inpatient hospital are provided in a patient's own home [9]. Briefly, all patients had their vitals continually monitored via a VitalPatch (VitalConnect, San Jose, CA) capable of ambulatory monitoring of vital signs remotely. The patients received at least 1 daily visit from an attending general internist and 2 daily visits from a home health registered nurse. From the total study patients, we have included patients who had a length of stay of at least 24 hours and had least 24 hours worth of objective heart rate and respiratory rate vitals data. This led to the inclusion of 303 patients in the final cohort for analysis. Clinical and demographic characteristics of the study cohort are shown in Tab. I.

2) eICU Cohort: The eICU Collaborative Research Database (The Philips eICU Research Institute) is a multi-center database comprised of health data from over 200,000 ICU admissions in the United States between 2014-2015 [10]. The dataset consists of over 190,000 unique patients, admitted to one of 335 hospital units spanning over 208

TABLE I: Patient Cohort Characteristics from Hospital at Home and eICU datasets

Characteristic	Hospital at Home	eICU (Hospital admission as start time)	eICU (ICU discharge as start time)
Number of Patients	303	9,520	6,435
Length of Stay (days)	4.22 ± 2.86	9.28 ± 7.28	5.25 ± 4.80
Number of Patient Day	1,298	75,204	27,420
Male/Female	39.9/60.1%	50.4/49.6%	49.8/50.2%
Age (years)	71.8 ± 16.9	65.2 ± 17.4	65.8 ± 17.2

hospital locations. The database is de-identified, includes vital signs measurements, laboratory measurements, traditional risk assessment measurements such as APACHE, as well as care and treatment plans. Many of the patients in this dataset have multiple ICU visits due to their high severity of illness.

The current analysis included a more homogeneous patient cohort of age 18 years or older, who had a single ICU visit as a transfer from any of the hospital floor units and survived. This eICU patient cohort was split into two subgroups. The first cohort had time events of hospital admission through hospital discharge and the second cohort had time events of ICU discharge through subsequent recovery at the hospital until hospital discharge. Therefore, two different starting points were used within the eICU dataset to create two unique cohorts for our analysis. Hospital admission (n=9,520 patients) was used as the start of one observation window and discharge from the ICU (n = 6,989 patients) was used as a the start of a second observation window. Both observation windows end when the patient is discharged from the hospital. Clinical and demographic characteristics of the cohort are highlighted in Table 1.

B. Data Processing

The physiological and actigraphy measurements associated with daytime versus the nighttime are known to be unique, reflecting the health status deeply [11]. Hence, numerous time windows have been explored for feature extraction considering both the daytime and nighttime. In the current study, we used the daytime, defined as 06:00 to 22:59, for daily feature generation. This daytime window maximizes the availability of patient health measurements and predictive power for dynamic patient discharge.

C. Features Used

A variety of healthcare data were acquired during patient care in acute and critical care settings. Additionally, the frequency of measurement amongst the different types of data varied. There was strong overlap between the vitals measurements and laboratory values used from the HH and eICU datasets. However, the HH dataset contained numerous quality of life questionnaire data which were not present in the eICU dataset.

1) HH Features: The demographic data used in modelling included answers to quality-of-life surveys and diagnosis at admission. Vitals features included statistical moments and extreme values of heart rate, respiratory rate, temperature and heart rate variability. Additionally, vitals data such as blood pressure and oxygen saturation (SpO2) were also manually measured at least two times a day. Laboratory values such as sodium, potassium, and creatinine (among others) used in the modelling were acquired as needed based upon the discretion of the care team. Finally, other relevant health data included daily use of intravenous medication, daily use of diuretics, and the ability to walk one flight of stairs.

2) eICU Features: The demographic data used in modelling included patient age, weight, and the APACHE diagnosis at admission. The vitals features used were limited to the mean, standard deviation, maximum and minimum values of heart rate, respiratory rate, and temperature. Additionally, systolic blood pressure, diastolic blood pressure, and Spo2 values were also used. The laboratory values used were matched to the same features used from HH. Healthcare utilization as well quality of life questionnaire data similar to the features found in the HH dataset were not found and thus not included in the final modelling.

D. Model Evaluation

The task of daily patient discharge prediction was approached from a dynamic classification perspective. The daily binarized time to discharge annotations were determined as either the positive or negative class based upon actual discharge of the patient within 24 and 48 hours.

Two models were developed to predict discharge within 24- and 48-hours thresholds. The input to the models included 69 and 34 features for the HH and eICU datasets, respectively, as described in the previous sections and the output was the discharge probability within 24 or 48 hours. Accordingly, if the calculated output for time to discharge was less than the threshold, that data point was assigned to the positive class and the remaining data points were assigned to the negative class.

In this study, we used Extreme gradient boosting (XGBoost) for the prediction task. The XGBoost model has been successfully used by previous researchers to model healthcare data [12] . XGBoost models are gradient boosted decision tree ensembles. In gradient boosting, the learning procedure consecutively fits new weak learners i.e. decision trees to the input data to provide a more accurate estimate of the outcome which helps overcome the limited performance associated with using a single decision tree.

We used grid search methodology, an extensive search over the model hyperparameters, to find the best parameters of the machine learning models in each case. We optimized the maximum tree depth, criterion for splitting the trees, coefficients of regularization and maximum number of features used to build each tree. After optimization, early stopping was used to prevent overfitting to the training data. Patients were split into 75% training and 25% testing groups. There was no overlap in patients between the two groups. Patient splitting was repeated 10 times randomly to ensure robustness of the model performance and generate mean and standard deviation of performance metrics. The same modelling approach was applied towards the eICU dataset.

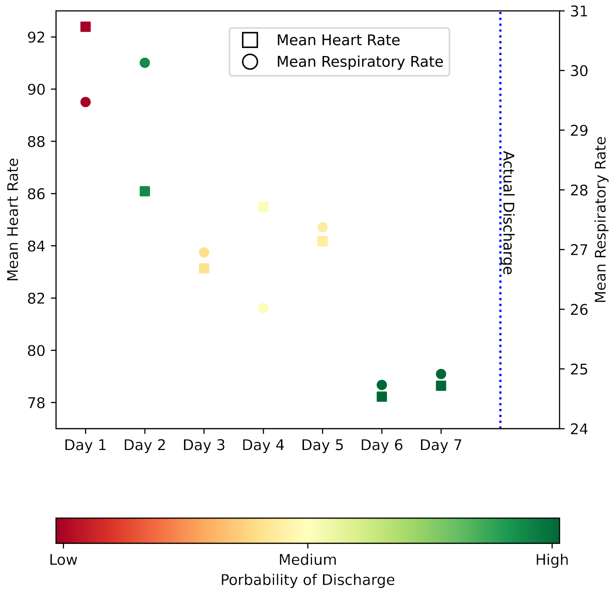
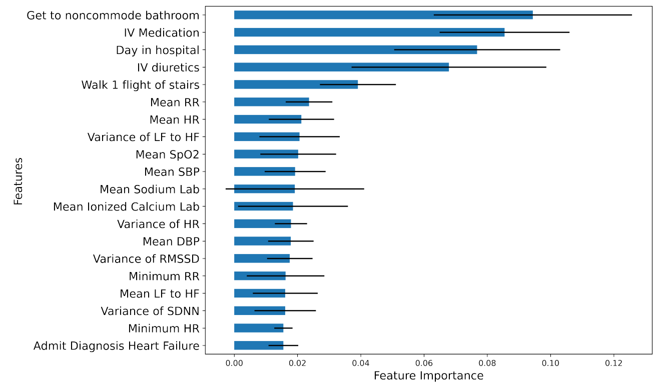
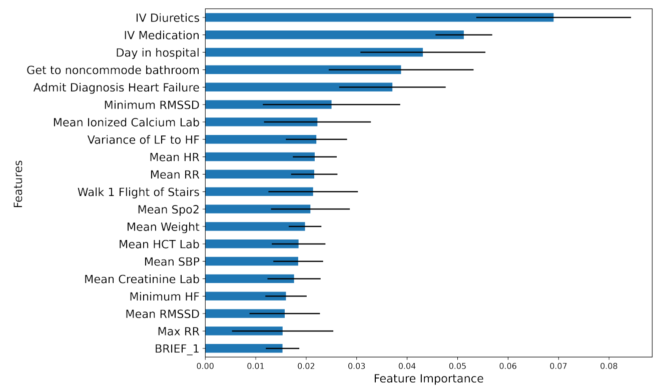


Fig. 1: The daily probability of discharge for a patient with a length of stay of 8 days. Low and high probabilities correspond to 0 and 1, respectively. In this particular example, the probability of discharge is low when the daily mean heart rate and respiratory rate are relatively abnormal. As these vitals approach more normal ranges, the probability of discharge increases.



(a)



(b)

Fig. 2: Top 20 feature importances for discharge prediction within (a) 24 hours and (b) 48 hours. The bars represent the mean values from the 10 random splits and the error bars represent the standard deviation.

III. Results

A. Example of Daily Patient Discharge

An example of daily patient discharge probabilities along with the measured heart rate and respiratory rate measurements is shown in Fig. 1. Following this patients journey throughout the duration of care, we can observed that this particular patient was discharged on the 8th day of care. Starting with day 1, this patient has a relatively higher daily average heart rate of 92.4 beats/minute, but also with a quite abnormal value of respiratory rate of nearly 30 breaths/minute. Correspondingly the probability of discharge for this patient on day 1 is determined to be low. As the hospital stay extends, the daily mean heart rate and daily mean respiratory rate values both show decreasing daily trends and approach the range of normal values. At the same time, the daily patient discharge probabilities are determined to be increasing, and the highest probability of discharge occurs within 24 hours of the actual patient discharge as demonstrated in this representative patient of Fig. 1.

B. Feature Importance

For each of the 10 random splits, feature importance scores were generated and the average feature importance is shown in Fig. 2. We calculated the feature importance using the weight method which is the number of times the feature appeared in a tree. The 20 highest ranking features had significant overlap between the 24-hour (Fig. 2a) and 48-hour (Fig. 2b) models.

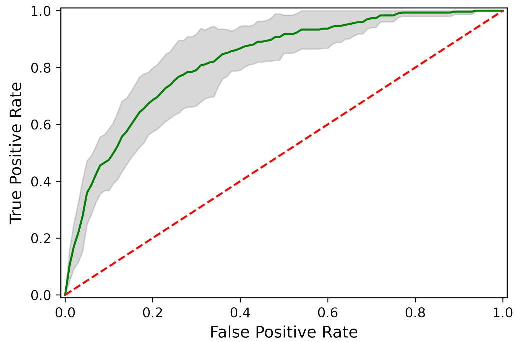
C. Hospital at Home Model Performance

The receiver operating characteristic (ROC) curve and the precision recall curve for discharge prediction within 24 hours and 48 hours are shown in Fig. 3 and Fig. 4, respectively. The results show that the area under ROC (AUROC) and average precision score from the 10 random patient splits are 0.816 ± 0.025 and 0.560 ± 0.050 respectively for the prediction of patient discharge within 24 hours. For a 48-hour prediction window we achieve an AUROC of 0.758 ± 0.029 while the average precision score is 0.715 ± 0.043. Overall, shifting from a 24-hour to 48-hour prediction window resulted in a decrease in the AUROC score while the average precision score increased.

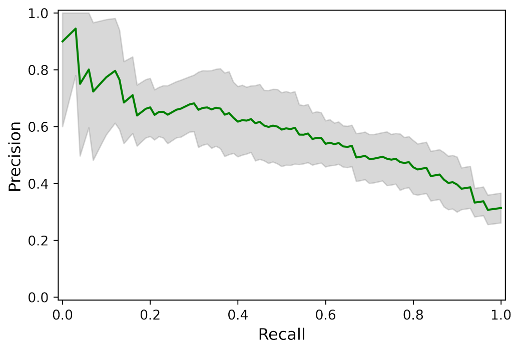
To further evaluate the model performance, the output probability decision threshold was varied in order to meet a variety of performance constraints. Tab. II indicates model performance measures when the f1 score is maximized, the highest specificity with a sensitivity of at least 80%, and the highest sensitivity with a specificity of at least 80%. Maximizing the mean of the precision and sensitivity results in an f1 score of 0.748, a precision of 0.499, a sensitivity of 0.740, and a specificity of 0.764 for the prediction of discharge within 24 hours. Maximizing the same criteria lead to an f1 score of 0.710, a precision of 0.687, a sensitivity of 0.711, and a specificity of 0.712 for the prediction of discharge within 48 hours.

D. eICU Model Performance

Model performance on the traditional intensive care environment of the eICU dataset is highlighted in Tab III. For

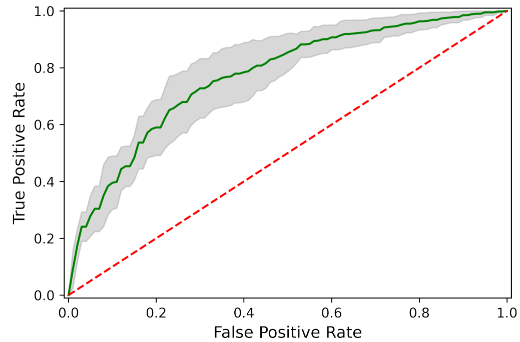


(a)

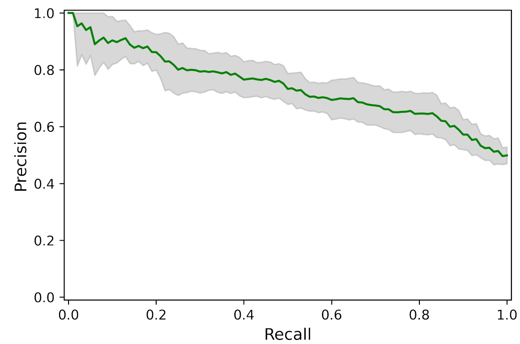


(b)

Fig. 3: The (a) ROC and (b) precision recall curves for discharge within 24 hours. The dashed red line represents a no skill classifier. The green line and shaded grey region represent the average and standard deviation, respectively, of the 10 random splits.



(a)



(b)

Fig. 4: The (a) ROC and (b) precision recall curves for discharge within 48 hours. The dashed red line represents a no skill classifier. The green line and shaded grey region represent the average and standard deviation, respectively, of the 10 random splits.

daily discharge predictions using hospital admission as the start time, the AUROC scores for discharge within 24 and 48 hours from the 10 random patient splits was $0.684 \pm .003$ and 0.668 ± 0.003 respectively. Moreover, the average precision score for discharge within 24 and 48 hours was $0.175 \pm .006$ and $0.323 \pm .008$. When daily discharge predictions were made using ICU discharge as the starting point, the AUROC scores for discharge within 24 and 48 hours was $0.618 \pm .007$ and 0.617 ± 0.007 respectively. Additionally, the average precision score for discharge within 24 and 48 hours was $0.287 \pm .006$ and $0.474 \pm .008$ respectively.

IV. Discussion

In this work, we demonstrate a novel machine learning modeling approach to predict patient discharge, from an acute, at home environment both within 24 and 48 hours and also in a traditional intensive care environment. While previous studies have focused on predicting how long a patient will receive care at the beginning of treatment [13], [14], we present a practical approach in which discharge predictions are made on a daily basis based upon newly available data. This dynamic approach accounts for changes in patient's health status and will help in an improved ability to determine when a patient can be discharged since patient improvements or deteriorations are accounted for on a daily basis.

Our results are comparable to the limited number of studies which have attempted daily discharge modelling. Barnes et al., predicted discharge at 2pm and end of day based upon electronic health record data available at 7am from the same day using different types of modelling approaches [6]. They reported sensitivities ranging from 0.60 to 0.72 and specificities ranging from 0.52 to 0.68 based upon the type of model and time of day. More recent studies have predicted patient discharge within 24 hours and have reported AUROC values of 0.84 [15] and 0.85 [5]. Although the performance of our algorithm is similar, it is important to note that previous studies used hundreds to thousands of more patients and patient days, a shorter prediction window, and were focused on a different care setting. The ability of our algorithm to match and in some cases exceed their performance may be due to the high quality and diverse health data used in our study including but not limited to continuous vitals measurements, laboratory measurements, and daily quality of life questions.

In order to determine the applicability of the daily discharge approach developed in this study, the same modeling efforts were applied to the eICU dataset. Daily patient discharge predictions within 24 hours had higher AUROC scores as compared to discharge predictions within 48 hours. More interestingly, the average precision recall score was higher for discharge predictions which began after patients

TABLE II: Hospital at home model performance at different decision thresholds. Decision thresholds were varied to attain the maximum f1 score, highest senstivity with at least 80% specificity, and highest specificity with at least 80% sensitivity.

Prediction Window	Condition	F1	Precision	Sensitivity	Specificity
24 hour	Maximize F1	0.748 ± 0.029	0.499 ± 0.051	0.740 ± 0.068	0.764 ± 0.056
	Top Specificity w/ at least 80% Sensitivity	0.726 ± 0.037	0.431 ± 0.044	0.811 ± 0.017	0.661 ± 0.066
	Top Sensitivity with at least 80% Specificity	0.715 ± 0.052	0.511 ± 0.023	0.647 ± 0.085	0.808 ± 0.019
48 hour	Maximize F1	0.710 ± 0.030	0.687 ± 0.029	0.711 ± 0.055	0.712 ± 0.037
	Top Specificity w/ at least 80% Sensitivity	0.661 ± 0.052	0.624 ± 0.048	0.808 ± 0.004	0.564 ± 0.078
	Top Sensitivity with at least 80% Specificity	0.657 ± 0.050	0.716 ± 0.030	0.558 ± 0.073	0.806 ± 0.004

TABLE III: Model Performance on eICU Dataset

Cohort	Prediction Window	AUROC	Average Precision Recall Score
Hospital admission as starting time	Within 24 hours	0.684 ± 0.003	0.175 ± 0.006
	Within 48 hours	0.668 ± 0.003	0.323 ± 0.008
ICU discharge as starting time	Within 24 hours	0.618 ± 0.007	0.287 ± 0.006
	Within 48 hours	0.617 ± 0.007	0.474 ± 0.008

were discharged from the ICU as compared to discharge predictions which began at the beginning of hospital admission. One possible reason for this difference may be that the post-ICU period of time is representative of more stable physiological patterns whereas the pre-ICU and actual ICU period of time are characterized by abnormal physiological instabilities. As a result, the model may be able to better learn physiological patterns during the post-ICU period leading to an improved ability to correctly predict discharges and thus a higher average precision recall score.

Overall, classification performance in terms of the AUROC and the average precision score was relatively decreased for the eICU dataset, despite having many more patient days available for training and testing of the classifier. The performance difference may be attributed to the fundamental inherent differences between the post-acute patients in the HH dataset and the intensive care patients in the eICU dataset. One key feature of the HH dataset is the availability of high quality continuous vitals data and a larger number of other high importance health features collected as highlighted in the feature importance ranking. In contrast, the frequency of the eICU vitals data is relatively low and there are time periods in which eICU patients have no new measurements of health data. These data availability limitations prevent the model from learning new relationships from the data and thus reducing the performance of the eICU model. Additionally, although we attempted to develop inclusion/exclusion criteria to select an eICU cohort that was similar to the HH dataset, the eICU cohort are still more severely ill than the HH cohort. This is evidenced by the fact that the eICU cohort spent approximately twice as much time in the care environment as compared to the HH cohort.

We also demonstrated that various desired performance metrics could be attained by fine tuning the decision threshold. This is an important consideration since it gives clinicians the freedom to set their desired threshold for patient discharge based on clinical priorities. For example, a clinical team may opt to achieve a higher sensitivity by setting a low discharge threshold to enable as many discharges as possible at the cost of some potentially early discharges. This trade off may be acceptable in situations where the disease states are mild or the patient population is relatively healthy. Alternatively, if resource utilization is a more relevant constraint, a higher discharge threshold could be used to improve specificity in order to better plan for the use of medical equipment or clinician care time.

Our study is not without limitations. First, the patient cohort is unique compared to other studies which tend to focus on care within a hospital or ICU environment. Thus, our results may be limited to our cohort and the area in which the study took place. However, despite the differences between the HH and traditional hospital environments, it was important to perform this investigation since there has been a recent paradigm shift towards healthcare in a home environment. Another limitation is that there is no gold standard for patient discharge. The discharge decision can be subjective based on a clinicians years of experience and visual observation of the patient. Hence, there may be situations were the model decision could have been correct but in disagreement with the clinical teams decision. Future, randomized clinical trial studies with an emphasis on post discharge outcomes are needed to determine the long term effects associated with optimal discharge times.

V. Conclusion

This study demonstrated a novel machine learning algorithm to predict daily patient discharges in both acute and intensive care settings. Such a tool could be potentially useful to aid clinicians or healthcare systems in making patient discharge decisions and ensure timely care by furthermore enhancing the hospital utilization.

References

[1] Stacy Ackroyd-Stolarz, J. Read Guernsey, N. J. MacKinnon, and G. Kovacs. The association between a prolonged stay in the emergency department and adverse events in older patients admitted to hospital: A retrospective cohort study. *BMJ Quality and Safety*, 20(7):564–569, jul 2011.

[2] Riccardo Caccialanza, Catherine Klersy, Emanuele Cereda, Barbara Cameletti, Alberto Bonoldi, Chiara Bonardi, Maurizia Marinelli, and Paolo Dionigi. Nutritional parameters associated with prolonged hospital stay among ambulatory adult patients. *CMAJ*, 182(17):1843–1849, nov 2010.

[3] A Azari, VP Janeja, A Mohseni International Journal of Knowledge, and undefined 2012. Healthcare data mining: predicting hospital length of stay (PHLOS). *igi-global.com*.

[4] Frederick C. Ryckman, Paul A. Yelton, Amy M. Anneken, Pamela E. Kiessling, Pamela J. Schoettker, and Uma R. Kotagal. Redesigning intensive care unit flow using variability management to improve access and safety. *Joint Commission Journal on Quality and Patient Safety*, 35(11):535–543, nov 2009.

[5] Kyan C. Safavi, Taghi Khaniyev, Martin Copenhaver, Mark Seelen, Ana Cecilia Zenteno Langle, Jonathan Zanger, Bethany Daily, Retsef Levi, and Peter Dunn. Development and Validation of a Machine Learning Model to Aid Discharge Processes for Inpatient Surgical Care. *JAMA network open*, 2(12):e1917221, 2019.

[6] S Barnes, E Hamrock, M Toerper Journal of the ..., and undefined 2016. Real-time prediction of inpatient length of stay for discharge prioritization. *academic.oup.com*.

[7] Thanos Gentimis, Ala Jamil Alnaser, Alex Durante, Ala ' J Alnaser, Kyle Cook, and Robert Steele. Predicting Hospital Length of Stay Using Neural Networks on MIMIC III Data.

[8] Fengyi Tang, Cao Xiao, Fei Wang, Jiayu Zhou, and Shaw Ln. Predictive modeling in urgent care: a comparative study of machine learning approaches.

[9] David M. Levine, Kei Ouchi, Bonnie Blanchfield, Agustina Saenz, Kimberly Burke, Mary Paz, Keren Diamond, Charles T. Pu, and Jeffrey L. Schnipper. Hospital-level care at home for acutely ill adults a randomized controlled trial. *Annals of Internal Medicine*, 172(2):77–85, 2020.

[10] TJ Pollard, AEW Johnson, JD Raffa, LA Celi, RG Mark Scientific Data, and undefined 2018. The eICU Collaborative Research Database, a freely available multi-center database for critical care research. *nature.com*.

[11] Shaun Davidson, Mauricio Villarroel, Mirae Harford, Eoin Finnegan, Joao Jorge, Duncan Young, Peter Watkinson, and Lionel Tarassenko. Vital-sign circadian rhythms in patients prior to discharge from an ICU: A retrospective observational analysis of routinely recorded physiological data. *Critical Care*, 24(1):205, 2020.

[12] Ali Jalali, Hannah Lonsdale, Nhue Do, Jacquelin Peck, Monesha Gupta, Shelby Kutty, Sharon R. Ghazarian, Jeffrey P. Jacobs, Mohamed Rehman, and Luis M. Ahumada. Deep Learning for Improved Risk Prediction in Surgical Outcomes. *Scientific Reports*, 10(1):1–13, dec 2020.

[13] Pei-Fang Tsai, Po-Chia Chen, Yen-You Chen, Hao-Yuan Song, Hsiu-Mei Lin, Fu-Man Lin, and Qiou-Pieng Huang. Length of Hospital Stay Prediction at the Admission Stage for Cardiology Patients Using Artificial Neural Network. 2016.

[14] A Morton, E Marzban, G Giannoulis 2014 13th ..., and undefined 2014. A comparison of supervised machine learning techniques for predicting short-term in-hospital length of stay among diabetic patients. *ieeexplore.ieee.org*.

[15] Anand Avati, Stephen Pfohl, Chris Lin, Thao Nguyen, Meng Zhang, Philip Hwang, Jessica Wetstone, Kenneth Jung, Andrew Ng, and Nigam H. Shah. Predicting Inpatient Discharge Prioritization With Electronic Health Records. Technical report, 2018.