Bridging the Gap: Forecasting ICU Demand and Readmission Risk Using MIMIC Data

BA878 Machine Learning and Data Infrastructure in Healthcare

Team 6: Kayla Huang, Wei An Huang, Zhao Peng, Jyun Ru Huang

Professor Ned McCague

Table of Contents

Abstract

Hospitals often face challenges predicting which emergency department (ED) patients will need intensive care and which ICU patients may require readmission soon after discharge. To address these gaps, we developed two machine learning models using data from MIMIC-IV-ED v2.2 and MIMIC-IV v3.1. The first model predicts ICU admission within 12 hours of ED arrival using early triage information, demographic variables, and home medication history. The second model predicts ICU readmission within 72 hours based on clinical data from the patient's first ICU stay.

We used XGBoost for both tasks and applied feature engineering to transform triage records, medication lists, and ICU event data into structured inputs. The 12-hour ICU admission model performed strongly, achieving an AUC of 0.96 and identifying most patients who needed early intensive care. The readmission model achieved an AUC of 0.72, reflecting the complexity of post-ICU deterioration but still providing useful screening information.

Finally, we combined both models into a proof-of-concept ICU capacity forecasting framework that links predicted ICU inflow, estimated bed occupancy, and short-term readmission risk. This demonstrates how predictive analytics can support short-term ICU planning. Future work should test this approach in other hospital settings and explore real-time implementation.

Introduction

Emergency department (ED) crowding and intensive care unit (ICU) strain stay as the challenges in hospital operations. Patients who initially appear stable in the ED may deteriorate within hours and require urgent ICU transfer. Moreover, some patients who have already been discharged from the ICU may experience readmission to critical care shortly after returning to a general ward. These transitions : from ED to ICU and from ICU to ward, are clinically consequential. Delays in necessary ICU escalation or premature ICU discharge can significantly increase patient mortality and prolong hospital stays. At the same time, they intensify the burden on already limited critical-care resources [1].

ED crowding tends to escalate when clinical deterioration isn't recognized early. Research indicates that early clinical information collected in the ED, such as vital signs and laboratory results can provide meaningful signals. In terms of empirical evidence, previous studies have shown that dynamic ED data, including serial measurements over the first few hours, can help predict which patients are more likely to require ICU care. For instance, a logistic regression model using vital signs and laboratory data achieved a cross-validated area under the receiver operating characteristic curve (AUROC) of approximately 0.70 for predicting ICU transfer within 24 hours [1]. In addition, evaluations of early warning scores such as NEWS2 have reported AUROC values close to 0.90 to find patients who may need ICU care [2]. Recent machine learning models that integrate a wider range of ED features including laboratory results and clinical orders, have achieved AUROC values of 0.88 for forecasting ICU admission within one day [3]. These findings suggest that timely and structured ED data can help identify high-risk patients.

Also, there is increasing interest in prediction of ICU readmission after discharge. Traditional scoring systems, such as the Stability and Workload Index for Transfer (SWIFT), have shown limited accuracy in identifying patients who are likely to be readmitted to the ICU. In contrast, recent machine learning approaches that use detailed electronic health record (EHR) data have strongly improved predictive performance. One study that combined structured ICU discharge data with radiology note features achieved an AUROC of approximately 0.78 for predicting 72-hour ICU readmissions. The performance had a better result compared to the SWIFT score, which showed an AUROC of only 0.62 in the same patient cohort [4]. Furthermore, a 2023 systematic review found that incorporating longitudinal patient data and time-series features can greatly enhance the predictive accuracy of ICU readmission models when compared to traditional models [5]. These results suggest that having both detailed clinical data and timing information is key to identifying patients at risk after ICU discharge.

Although some advances have been made, most studies still focus on just one side of the problem: either predicting who will need ICU care from the ED, or who might come back to the ICU after discharge. This split approach makes it hard for hospitals to fully optimize the short-term ICU demand. In practice, both ends of the ICU flow are important: recognizing early deterioration in the ED and timely discharge decisions in the ICU. Although each transition has been studied on its own, there is still no integrated approach that connects them.

In response to these challenges, our team's study aimed to develop and validate two complementary machine learning models. The first model predicts ICU admission within 12 hours using early ED data. The second model predicts ICU readmission within 72 hours based on clinical data available prior to ICU discharge. Together, these models are intended to support hospital decision-making by providing a preliminary unified framework for real-time ICU capacity forecasting.

Method

1. Data Sources

Our team integrates data from two publicly available databases on PhysioNet: MIMIC-IV-ED v2.2 and MIMIC-IV v3.1. These datasets collectively contain detailed information on emergency department (ED) encounters, hospital admissions, and intensive care unit (ICU) stays, allowing us to reconstruct each patient's care trajectory from ED arrival to ICU discharge and potential readmission.

MIMIC-IV-ED v2.2

We use four component tables from the MIMIC-IV-ED database: ed_stays, diagnosis, medrecon, and pyxis. These tables capture ED patient demographics, visit characteristics, medication reconciliation records, medications dispensed in the ED, and diagnosis information recorded during the encounter. Together, they provide the initial clinical and operational context for each patient before hospital or ICU admission.

MIMIC-IV v3.1

From MIMIC-IV v3.1, we use both the Hospital (HOSP) module and the ICU module. The Hospital module includes the admissions, transfers, and patients tables. These tables contain hospital-level information such as admission type, timing of encounters, intra-hospital movements, and demographic characteristics. The ICU module includes icustays, caregiver, d_items, datetimeevents, and inputevents. These tables provide high-resolution ICU data, including ICU stay timing, medication administration, treatment activities, and item-level metadata.

Data Integration

To define the analytic cohort, we combine the ED, HOSP, and ICU modules using the shared identifiers subject_id, hadm_id, and stay_id. This integration allows us to trace the full clinical course for each patient, starting from the ED presentation, continuing through the hospital admission and the initial ICU stay, and ending with the assessment of whether an ICU readmission occurs within 72 hours after ICU discharge.

It is important to note that the ED and ICU modules use entirely different numbering systems for their stay_id columns. Therefore, we could not merge these two datasets directly using stay_id and instead relied on hadm_id as the common key to link records across departments. For the ED data construction, we first linked the relevant tables using the ED specific stay_id. Since a single hospital admission usually corresponds to one primary ED visit, we grouped the data by hadm_id and selected the first chronological observation to represent the patient's arrival. We applied a similar logic to the ICU data processing. We merged the ICU tables using the ICU specific stay_id and then aggregated the results by hadm_id, again retaining only the first ICU stay to ensure each hospital admission was represented by a single sequence of events.

A complete summary of the datasets used in this study is provided in Supplementary Table S1 and S2, which includes descriptions of each data source.

Table S1. Description of MIMIC-IV-ED v2.2 Data Sources

Dataset	Descriptions

1. ed_stays	Encounter-level ED stay information, including ED arrival and discharge timestamps, patient demographics (e.g., gender, race), arrival mode, discharge disposition, and identifiers linking ED visits to subsequent hospital and ICU encounters.
2. diagnosis	ED-assigned diagnosis codes and descriptive text documenting clinical assessments made during the ED visit.
3. medrecon	Medication reconciliation records containing outpatient/home medications reported at ED intake, reflecting pre-admission medication exposure.
4. pyxis	Medication dispensing records from automated dispensing machines in the ED, including medication names, quantities dispensed, and timestamps of administration.
5. triage	Triage captures the earliest clinical and administrative information collected when a patient arrives in the ED. It includes vital signs, pain scores, ESI acuity levels, demographics, arrival details, and initial clinical assessments.

Table S2. Description of Data Sources- MIMIC-IV-v3.1

Dataset	**Descriptions**
HOSP	
1. admissions	Admission-level information including admission type, timestamps, insurance, language, ethnicity, and linkage to ED encounters.
2. transfers	Timestamped records of intra-hospital location changes, enabling identification of ICU admission, discharge, and readmission events.
3. patients	Immutable demographic information such as sex, date of birth, and mortality indicators (in-hospital and post-discharge).
ICU	
1. icustays	ICU encounter-level records including ICU admission/discharge times, unit type, and unique stay identifiers (stay_id).
2. caregiver	Contains a single deidentified caregiver identifier (caregiver_id) representing the provider who documented data in the MetaVision ICU system. This table can be used to link ICU event records to documenting caregivers but is not used directly as a predictive feature in this study.
3. d_items	Dictionary table containing metadata for all itemids used in ICU event tables, including item names, units, and item categories.
4. datetimeevents	Timestamped ICU clinical events recorded as date/time entries, such as ventilator settings and procedure-related timestamps.
5. inputevents	Detailed records of ICU medication administrations and IV fluid events, including doses, rates, units, and event timestamps.

2. Study Design and Study Population

We conducted a retrospective cohort study using electronic health record (EHR) data from MIMIC-IV-ED v2.2 and MIMIC-IV v3.1. For each patient's earliest eligible ED visit during the study period was defined as the index encounter, and subsequent hospital and ICU events were linked using subject_id, hadm_id, and stay_id to reconstruct complete care trajectories.

Phase 1: Prediction of Early ICU Admission (12-hour window)

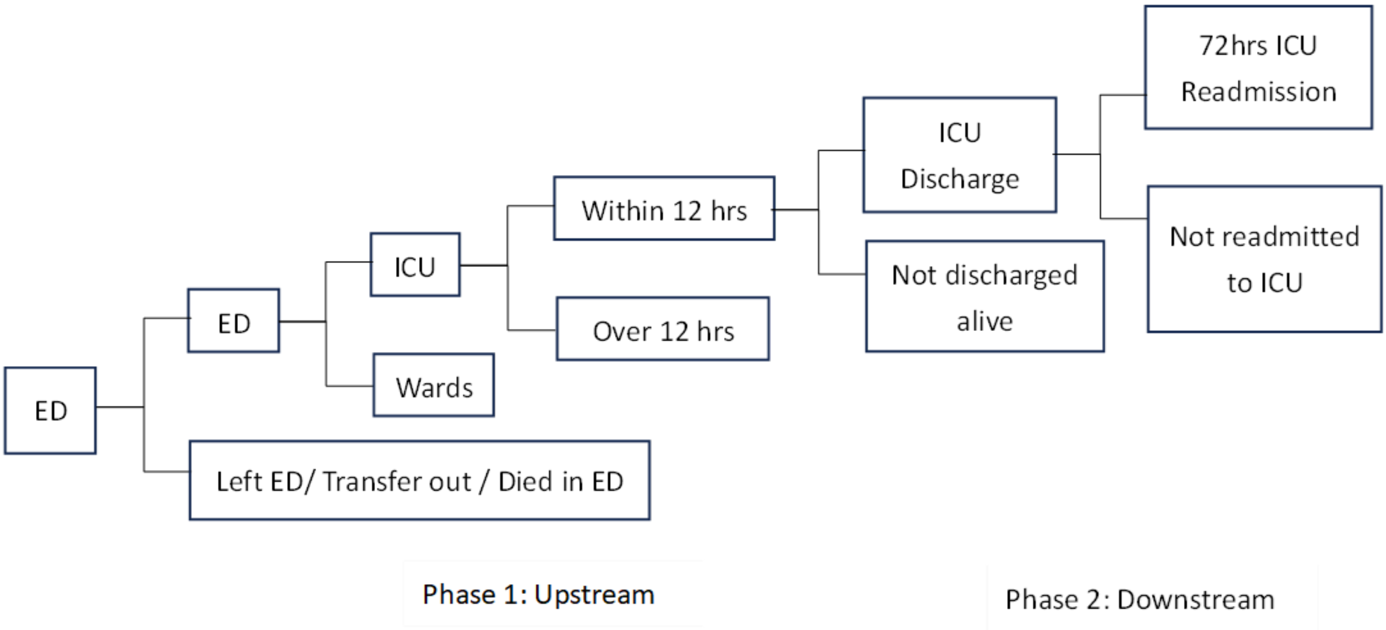
Phase 1 focused on identifying patients at risk of needing ICU care soon after arriving in the ED. We included all ED visits that led to hospital admission and excluded patients who left before receiving care, were transferred to other hospitals, or died in the ED. Among the admitted group, those moved to the ICU within 12 hours of ED arrival were labeled as positive cases. Patients admitted to regular wards or to the ICU after more than 12 hours were used as the comparison group.

Phase 2: Prediction of ICU Readmission Within 72 Hours

Phase 2 is conditional on the Phase 1 positive cohort and therefore represents a clinically distinct subpopulation. This phase focused on identifying patients at risk of deterioration after their first ICU stay. To ensure consistent care for patients, we included only those whose ICU admission came directly from the ED, excluding ICU stays that began in the operating room, inpatient wards, or outside hospitals. Among patients discharged alive from their first ICU admission, those readmitted within 72 hours were labeled as positive cases and patients without early readmission considered as a comparison group. This phase captures the challenge of predicting early instability after ICU discharge.

Together, the two phases would enhance the ED to ICU pathway: anticipating ICU needs early during ED evaluation, and assessing short-term readmission risk at the time of ICU discharge. We also ensure each model was trained on a well-defined, clinically meaningful patient group to have actionable forecasting in real-world settings.

Figure 1. Study Structure for Phase 1 and Phase 2



3. Variable Construction and Feature Engineering

Cohort Selection and Data Integration

To support both models: ICU admission within 12 hours of ED arrival in Phase 1 and identifying ICU readmission within 72 hours after ICU discharge in Phase 2, we constructed an integrated data processing workflow using multiple variables in the MIMIC-IV database. The major steps are summarized below.

We started by building a base cohort using data merged from the ed-stays, patients, admissions, transfers, and icustays tables. We aligned timestamps across these sources to maintain the correct order of events and to calculate time-to-ICU metrics needed for defining the cohort.

Phase 1: Upstream ICU Admission Cohort: For Phase 1, we included ED visits that led to hospital admission. The outcome variable, icu_within_12h, was marked positive if the patient was admitted to any ICU within 12 hours of ED triage. We excluded patients who left the ED before treatment, were transferred to another facility, or died in the ED.

Phase 2: Downstream ICU Readmission Cohort: Phase 2 focused on downstream risk among patients identified as needing early ICU care in Phase 1. It included only patients who were admitted to the ICU within 12 hours of arriving at the ED which means the same group identified as positive cases in Phase 1. From this group, we included only patients who were discharged alive after their first ICU stay. To avoid mislabeling administrative transfers as clinical readmissions, ICU-to-ICU moves within 72 hours (such as MICU to SICU or SICU to CCU) were excluded, as these typically reflect service changes or bed availability rather than true deterioration. The outcome variable, *readmission_72h*, was defined as a return to any ICU within 72 hours after discharge.

This two-phase framework traces the full care path from ED arrival through initial ICU treatment and short-term post-ICU outcomes, allowing both models to train on proper population.

Triage Data Processing

We standardized the MIMIC IV triage data to create consistent and reliable inputs for analysis. The pain score field contained mixed formats, including numbers, descriptive text, spelling variations, and clinically meaningful phrases. To address this variability, we designed a cleaning module that normalizes text, groups similar expressions, and extracts valid numeric values whenever possible. When the text contains a number such as "5," "5/10," or "3 to 4," the parser selects the value that best represents the intended pain score and ignores numbers that obviously refer to vital signs such as 98.6 or 110. Descriptive terms are also mapped onto a numeric scale with values representing no, mild, moderate, and severe pain.

If the patient is unable to provide a score, as indicated by entries such as "unable," "refused," "intubated or nonverbal," or "unknown," the value is set to negative one. This preserves its clinical meaning and prevents confusion with a true score of zero. The final pain value follows a simple rule: it first uses any valid numeric score, then uses the descriptive mapping when needed, and finally assigns negative one when no interpretable information is available.

Chief complaints were originally recorded as free text with inconsistent formatting. We converted each entry into a cleaned list, standardized the text, removed duplicates, and excluded administrative terms such as "TRANSFER." After calculating term frequencies, we selected the twenty five most common complaint terms and created a binary indicator for each one. This produces a compact and clinically meaningful representation of chief complaints. Vital signs, including temperature, heart rate, respiratory rate, oxygen saturation, and blood pressure, were kept as continuous variables without further transformation because they are already standardized.

Medication Feature Engineering

The medication features were derived from the medrecon table, which lists the home medications recorded at triage. This dataset presented a challenge because the raw text was unstructured and high dimensional. A single patient often had multiple rows of data representing different prescriptions, so we needed to aggregate this information into a single row per hospital stay.Our feature engineering process involved three main steps to turn this text into usable model inputs.

First, we performed text normalization to reduce noise. We converted all drug names to lowercase and stripped out dosage information, such as numbers and units like "50mg" or "tablet". This was necessary because the same drug is often recorded with different dosages, and treating them as separate features would create too many unique variables.

Second, we addressed the issue of sparsity by grouping drugs into broader categories. Using a keyword based dictionary, we mapped thousands of unique drug names into 22 clinical categories. This step served as a form of dimensionality reduction. It allowed the model to detect patterns related to chronic diseases, such as hypertension or diabetes, without needing to learn the name of every specific brand or generic drug.

Finally, we encoded these features for the machine learning model. We used a count encoding method where each category became a column representing the number of medications the patient was taking in that group. For patients with no records in the medrecon table, we imputed these features with zero, assuming they were not taking any home medications.

Final Dataset Construction

Following the preprocessing of triage and medication data, all engineered features were integrated into the base emergency department cohort using the unique stay identifier (stay_id). The resulting analytic dataset consolidated demographic characteristics, encounter attributes, triage information, early medication exposures, and diagnosis-derived clinical markers into a single structured framework suitable for predictive modeling.

Categorical variables generated during feature engineering were assigned a dedicated "Missing" category when values were unavailable, allowing potential information contained in documentation gaps to be preserved. Numerical variables, including vital signs and temporal medication measures, retained their original missingness, which was handled directly by XGBoost during model development.

To address differing analytic requirements, two versions of the dataset were constructed. The full dataset incorporated all engineered variables, including processed triage features, chief complaint terms, medication class counts, temporal medication metrics, and diagnosis indicators. The compact dataset contained a reduced set of predictors consisting of demographic characteristics, core triage attributes, early medication categories, and high-risk diagnosis markers, providing a more interpretable structure for secondary analyses.

4. Independent Variables

The final feature set included three types of information available early in the ED visit. These covered patient characteristics, clinical status at presentation, and outpatient medication history.

Patient Demographics and Visit Information

This category included age, gender, insurance type, marital status, and arrival mode. These variables describe baseline patient characteristics and contextual information about the visit.

Triage Vital Signs and Clinical Assessment

We included standard triage vital signs such as, temperature, heart rate, respiratory rate, oxygen saturation, and blood pressure. We also used acuity level, a cleaned 0–10 pain score, and one-hot indicators for the most common chief complaints.

Home Medication History

We extracted medication features from the medrecon table, which lists home medications recorded before ED arrival. These medications offer an overview of chronic conditions and baseline health. We grouped drugs into broad categories using keyword matching with base-rule and added count features for commonly reported medications. A summary of these features is provided in Table S3.

Table S3. Feature Description

Category	**Feature Description**
Demographics	Age, Gender, Insurance type, Marital status, Transport mode
Triage Vitals	Temp, HR, RR, SpO_2, SBP, DBP

Triage Assessment	Acuity level, Cleaned pain score, top presenting complaints (binary)

5. Dependent Variables

This study used two outcome variables corresponding to the two prediction tasks in the modeling framework.

ICU Admission Within 12 Hours (icu_within_12h)

This variable indicates whether a patient was admitted to any ICU within 12 hours of ED arrival. It serves as the primary outcome for Phase 1 and reflects the need for early critical care escalation.

ICU Readmission Within 72 Hours (readmission_72h)

This variable indicates whether a patient returned to any ICU within 72 hours after being discharged alive from their first ICU stay. It serves as the primary outcome for Phase 2 and captures early post-ICU deterioration requiring renewed intensive care.

6. Statistical Analyses

We used descriptive statistics to compare the ED to ICU and ED to Non-ICU groups. Continuous variables are reported as mean (SD) and were compared using Student's t-tests. Categorical variables are reported as counts (%) and were compared using chi-square tests. Tests were two sided, with $p < 0.05$ considered significant. All analyses were performed in Python 3.12.

7. Predictive Modeling

Two predictive modeling frameworks were developed corresponding to the two phases of this study:

(1) prediction of ICU admission within 12 hours of ED arrival, and

(2) prediction of ICU readmission within 72 hours following ICU discharge.

For both tasks, we trained XGBoost models using structured clinical features that were generated after preprocessing and feature engineering. The dataset was split into training (80%) and testing (20%) sets with separated sampling to keep the original class balance. Since both outcomes were highly imbalanced, we set scale_pos_weight to the negative-to-positive ratio in the training data so the model gives more weight to positive cases.

We evaluated performance mainly with AUC score. We also report precision, recall, and F1-score on the test set to reflect positive case detection. To interpret better, we reviewed feature importance from the XGBoost models and tuned hyperparameters on a validation split to improve stability and reduce overfitting.

8. Features Selection

Both Model 1 and Model 2 initially contained approximately one hundred predictors after data preprocessing. To evaluate whether a more compact model could achieve comparable performance, we applied XGBoost's built-in feature importance scores to rank all predictors based on their contribution to the model.

After training the baseline model using the full feature set, we extracted the importance scores and sorted the variables in descending order. We then selected the top 30 most important features and used these variables to construct a reduced version of the model. A new XGBoost classifier was trained using only these 30 features, and its performance was compared against the baseline model.

Across both prediction tasks, the simplified models achieved AUC values very similar to the full models, with no meaningful performance degradation. This suggests that a relatively small subset of key predictors is sufficient to capture the majority of the predictive signal. Reducing the feature set not only improves model interpretability but also supports more efficient deployment in real-world hospital systems.

9. ICU Capacity Forecasting Framework

We developed an ICU capacity forecasting framework that integrates three key components: predicted ICU admissions from the emergency department, ICU length of stay, and early ICU readmission risk. Together, these elements provide the information necessary to anticipate short-term ICU bed demand and support hospital operations.

Handling Timestamp Shifting

MIMIC-IV version 3.0 expands its underlying data coverage to the years 2008 through 2022. However, none of the timestamps in the database are provided in their real calendar form. To protect patient privacy, all date-time fields undergo systematic date shifting, which moves each patient's entire timeline to a future range between the years 2100 and 2200. This shifting does not affect model development because every patient's dates are shifted by a constant offset, preserving all relative ordering of events such as ED arrival, ICU admission, ICU discharge, and subsequent readmission.

For the purpose of demonstrating how our ICU capacity forecasting framework operates, we treat the shifted timestamps as though they were actual calendar dates. This assumption allows us to visualize the dynamic flow of patients through the ED and ICU, including how many patients arrive in the ED, transition into the ICU, prepare for ICU discharge, or face potential readmission at any given moment in time.

Under this framework, we can illustrate time-specific demand estimates. For example, at 12:00 PM on January 1st, 2150, we can approximate how many ED patients are likely to require ICU admission (Model 1), how many ICU beds remain occupied due to historical LOS, and how many recently discharged patients may return based on early readmission risk (Model 2). Treating the shifted dates as operational timestamps allows us to demonstrate the feasibility of forecasting short-term ICU demand even though the absolute calendar years in MIMIC-IV do not reflect real historical dates.

Predicted ICU admissions from the ED (Model 1)

The first component is generated using the 12 hour ICU admission model, which assigns each ED patient a probability of requiring ICU level care within the next 12 hours. Adding these probabilities across current ED patients shows a short term estimate of expected ICU admissions and helps flag upcoming demand.

ICU length of stay

The second part uses ICU length of stay for simulation. When the 12-hour model flags an ED patient as they are more likely to need ICU care, that prediction would be linked to the patient's actual ICU entry time and historical LOS. So the values are used to estimate how long an ICU bed would be occupied during the simulated window. This does not affect model training. LOS is not used as a predictive input. It is only applied after predictions are generated to estimate bed use which create a more realistic view of ICU occupancy and short-term demand.

Early ICU readmission risk (Model 2)

The third component examines early ICU readmission risk using the 72 hour model. It targets patients who were recently discharged from the ICU but may return within a short time. These predictions help assess whether a discharge is suitable and whether additional ICU demand may follow soon, in order to support decisions about closer monitoring and safer discharge planning.

Taken together, predicted ICU admissions, historical ICU length of stay, and early readmission risk provide a more complete short term view of ICU demand. This combined approach shows how the two models can work together to support bed planning and operational decisions, including triage, discharge timing, and resource allocation.

Results

1. Descriptive Results

Among 425,053 ED patients included in the cohort, 31,915 (7.5%) were admitted to the ICU within 12 hours (ED–ICU), while 393,138 (92.5%) were not (ED–Non-ICU). Patients in the ED–ICU group were older on average (mean age 62.4 vs. 49.4 years, $p < 0.001$) and more likely to be male (54.3% vs. 45.2%, $p < 0.001$). Racial distribution also differed significantly, with a higher proportion of White and “Other” race patients in the ED–ICU group, whereas Black, Asian, and Hispanic/Latino individuals comprised larger shares of the ED–Non-ICU population ($p < 0.001$).

Marital status varied across groups: ED–ICU patients were more frequently married or widowed, while the ED–Non-ICU group had a substantially higher proportion classified as “other” marital status ($p < 0.001$). Insurance coverage also differed significantly. Medicare and private insurance were more common among ED–ICU patients, and Medicaid coverage was likewise higher in the ICU group (17.0% vs. 9.4%) ($p < 0.001$).

Arrival mode showed notable contrasts: 64.1% of ED–ICU patients arrived by ambulance compared with 34.4% of ED–Non-ICU patients, whereas walk-in arrivals were far more common among the Non-ICU group (62.3% vs. 21.4%) ($p < 0.001$). The ED–ICU group also had a shorter mean ED length of stay (5.6 vs. 7.3 hours, $p < 0.001$).

Finally, 30-day mortality was markedly higher in the ED–ICU group (15.3%) compared with 1.0% among ED–Non-ICU patients ($p < 0.001$).

Table 1. Baseline Characteristics between ICU and Non-ICU Groups

Characteristics	ED - ICU		ED – Non-ICU		P-Value
	N	%	N	%	
N (%)	31,915	7.5	393,138	92.5	
Age (Mean ± SD)	62.4	17.7	49.4	20.0	<0.001*
Gender					
Male	17,319	54.3	177,849	45.2	<0.001*
Female	14,596	45.7	215,289	54.8	<0.001*
Race					
White	20,462	64.1	225,522	57.4	<0.001*
Black	4,378	13.7	88,609	22.5	<0.001*
Asian	1,086	3.4	17,441	4.4	<0.001*
Hispanic/Latino	1,291	4.0	33,219	8.4	<0.001*
Others	4,698	14.7	28,347	7.2	<0.001*
Marital Status					
Married	12,794	40.1	62,760	16.0	<0.001*
Single	10,080	31.6	72,266	18.4	<0.001*
Widowed	4,173	13.1	20,031	5.1	<0.001*
Others	4,868	15.3	238,081	60.6	<0.001*
Insurance					
Medicaid	5,423	17.0	36,985	9.4	<0.001*
Medicare	17,683	55.4	77,279	19.7	<0.001*
Private	7,307	22.9	49,029	12.5	<0.001*
Others	1,502	4.7	229,845	58.5	<0.001*
Arrival Transport					
Ambulance	20,447	64.1	135,288	34.4	<0.001*
Helicopter	612	1.9	256	0.1	<0.001*
Walk in	6,836	21.4	244,999	62.3	<0.001*
Others	4,020	12.6	12,595	3.2	<0.001*
ED LOS, hours (Mean ± SD)	5.6	3.7	7.3	6.6	<0.001*
30-Day Mortality	4,888	15.3	4,020	1.0	<0.001*

*ED–ICU refers to ED patients who were admitted to the ICU within 12 hours.

*ED–Non-ICU refers to ED patients who were not admitted to the ICU within 12 hours.

2. Predictive Modeling Results

12-hr ICU Admission

We trained several variants of the inflow model to predict whether a patient would be admitted to the ICU within 12 hours of arriving in the ED. The best-performing model used ED encounter data, triage measurements, demographic features, and home medication history. We used XGBoost with categorical encoding enabled and applied a scale_pos_weight adjustment to address the class imbalance.

On the test set, the updated model achieved an AUC of 0.9630 and an overall accuracy of 0.8706. For the ICU class, the model reached a recall of 0.93, meaning it successfully identified the vast majority of patients who required intensive care. The precision for ICU cases was 0.30, which reflects the expected tradeoff when detecting a rare but clinically important outcome. The F1-score for ICU predictions was 0.45. For non-ICU patients, performance remained strong, with a precision of 0.99, recall of 0.87, and an F1-score of 0.93.

Overall, these results indicate that incorporating triage information and medication history greatly strengthens the model's ability to flag high-risk patients early in their ED stay. The high recall for ICU cases is especially valuable because it ensures that patients likely to need intensive care are identified promptly, enabling earlier clinical intervention and more informed ICU resource planning.

Table 2. Performance of the Best XGBoost Model (main + triage + med)

Metric	Class 0 (Non-ICU)	Class 1 (ICU)	Overall
Precision	0.99	0.30	-
Recall	0.87	0.93	-
F1-score	0.93	0.45	-
Accuracy	-	-	0.87
AUC	-	-	0.96

72-hr ICU Readmission

For the second model, we focused on predicting whether a patient would be readmitted to the ICU within 72 hours of being discharged. We trained an XBoost model using data from the ICU stay itself, which included specific features like medication ingredients and input events collected in the last 24 hours. Since readmission is a relatively rare event compared to a standard discharge, we applied a scale_pos_weight adjustment to help the model pay more attention to the minority class.

On the test set, the model achieved an AUC of 0.7205 and an overall accuracy of 0.79. The performance for the readmission class was more challenging than the admission model, resulting in a recall of 0.52 and a precision of 0.12. This means the model correctly identified about 52 percent of the patients who eventually returned to the ICU. For the patients who did not need readmission, the model performed well with a precision of 0.97 and a recall of 0.80.

These results suggest that predicting readmission is inherently difficult because it depends on many complex factors during the recovery process. Although the precision for the positive class is low, the moderate recall value indicates that the model can still serve as a screening tool. It can help clinical teams identify a portion of high risk patients who might benefit from a longer observation period before they are transferred to a general ward.

Table 3. Performance of the XGBoost Readmission Model (ICU Events)

Metric	Class 0 (No Readmission)	Class 1 (Readmission)	Overall
Precision	0.97	0.12	-
Recall	0.80	0.52	-
F1-score	0.88	0.20	-
Accuracy	-	-	0.79
AUC	-	-	0.72

ICU Dashboard

This forecasting module generates a unified ICU bed outlook by combining predicted ICU admissions from the ED, simulated ICU occupancy based on real patient timelines, and predicted ICU outflows that incorporate early readmission risk. The framework produces a time-indexed dataset that reflects expected census levels, bed releases, upcoming demand from the ED, and potential readmissions at every 12-hour interval. By generating these forecasts, the system offers a more realistic view of short-term ICU capacity. It helps clinicians anticipate bottlenecks, supports data-driven triage and discharge planning, and enables hospital operations teams to allocate beds and resources more effectively. We provide the dashboard results for September 1, 2150 as a reference.

Table 4. ICU Capacity Forecasting Dashboard Output

timestamp	total_ED_arriving_patients	Current_ICU_Bed_occupied	pred_ICU_inflow_12h	actual_ICU_inflow_12h	pred_ICU_outflow_12h	actual_ICU_outflow_12h	pred_ICU_net_change_12h	actual_ICU_net_change_12h	pred_ICU_readmission
2150/9/5 00:00	7	3	2	0	0	2	0	0	1
2150/9/5 12:00	11	3	2	1	3	-1	-2	0	0
2150/9/6 00:00	6	3	0	0	0	0	0	0	0
2150/9/6 12:00	7	3	2	1	3	0	-2	1	0
2150/9/7 00:00	1	2	0	0	2	-2	-2	0	0
2150/9/7 12:00	9	2	1	1	1	0	0	0	0
2150/9/8 00:00	4	2	1	0	0	1	0	0	0
2150/9/8 12:00	8	3	3	2	2	1	0	0	0
2150/9/9 00:00	7	4	1	1	0	1	1	0	0

Discussion

This study showed that two complementary machine learning models can help address important operational challenges in critical care. The first model predicts ICU admission within 12 hours of a patient's arrival in the emergency department, and it demonstrated strong performance. This suggests that early ED information contains meaningful signals that can identify patients who are at risk of rapid deterioration. The second model predicts ICU readmission within 72 hours after discharge. Its performance was more modest, which is consistent with previous research showing that post-ICU instability is influenced by many complex and sometimes unmeasured clinical factors.

Our findings contribute to existing evidence that structured electronic health record data can support early identification of high-risk patients in both ED and ICU settings. Unlike most prior studies that examine these problems separately, our work combines both prediction tasks and integrates them into a unified framework for short-term ICU capacity forecasting. By incorporating predicted ICU inflow, expected ICU bed occupancy based on retrospective LOS replay, and estimated readmission risk, our framework illustrates how predictive modeling can support real-time decisions about triage, discharge timing, and ICU resource allocation.

This study has several limitations. It was conducted using retrospective data from a single medical center, which may reduce how well the findings apply to other hospitals. In addition, because we did not have access to GPU computing resources and had limited experience with neural networks, we relied on traditional feature engineering to preprocess the input data. A neural network approach may allow future work to incorporate unstructured clinical notes, imaging information, or physiological time series, which could improve predictive performance.

The ICU capacity forecasting framework presented in this study should be viewed as a proof of concept rather than a system that can be deployed immediately. It uses actual lengths of stay from the dataset, and the timestamps in MIMIC-IV have been shifted forward, so they cannot be used for real-time forecasting. To move this research toward practical implementation, validation across multiple clinical sites is necessary, and a model that predicts length of stay must also be included. Only with these additions can the framework be applied in real operational environments within hospitals.

Overall, the results show that combining prediction of ED-to-ICU admission and ICU-to-ward readmission can improve situational awareness and help hospitals better anticipate short-term changes in ICU demand.

Conclusion

In conclusion, this study demonstrates that routinely collected electronic health record data can be used to build effective machine learning models for two important clinical transitions: ICU admission shortly after ED arrival and ICU readmission soon after discharge. The ICU admission model performed strongly, and the readmission model showed moderate but still relatively well accuracy. Together, these results suggest that early clinical signals can help identify higher risk patients and support more proactive ICU planning.

The combined framework also shows more than single patient prediction. By linking admission risk, expected ICU bed use, and short-term readmission risk, it offers a practical way to anticipate near term changes in ICU demand. This can inform triage priorities, discharge timing, and short-term staffing and bed allocation.

From a practical perspective, these results supporting using prediction models can help teams plan ahead for short term ICU demand. Future models may also benefit from richer data, especially like trends over time and clinical notes.

Future work should test this approach in different hospitals and assess real time use. It should also examine how these predictions can be used into existing decision support tools to show whether they improve patient safety and ICU flow.

References

[1] Glass, G. et al. (2021). *Dynamic data in the ED predict requirement for ICU transfer following acute care admission*. Crit. Care, 25(1): 123. [PMCID: PMC8145085]

[2] Gidari, A. et al. (2020). *Predictive value of National Early Warning Score 2 (NEWS2) for ICU admission in patients with SARS-CoV-2 infection*. Infect. Dis. (Lond), 52(10): 698–704. [PubMed: 32835597]

[3] Nguyen, M. et al. (2021). *Developing machine learning models to personalize care levels among emergency room patients for hospital admission.* J Am Med Inform Assoc., 28(11): 2423–2432. [PMCID: PMC8510323]

[4] Hu, H. et al. (2025). *ReAdmit: Predicting early unplanned ICU readmission using radiology notes and structured data*. Nurs. Crit. Care, 30(6): e70200. [PubMed]

[5] Ruppert, M.M. et al. (2023). *Predictive modeling for readmission to intensive care: a systematic review*. Crit. Care Explor., 5(1): e0848. [PubMed: 36723728]