



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

Team 6: Kayla Huang, WeiAn Huang, Zhao Peng, JyunRu Huang

# Table of contents

01

Introduction

02

Method

03

Results

04

Discussion

05

Conclusion

06

Future Steps



# Introduction

# Background



## ED Crowding Delays Critical Care

Overcrowding makes early deterioration harder to detect, causing some patients to worsen within hours and require urgent ICU transfer.

## ICU Strain Raises Mortality

Delays in escalation or premature ICU discharge increase mortality risk and lengthen hospital stays.

## Need for Short-Term ICU Capacity Planning

Hospitals need real-time ICU demand forecasting to support safe triage, timely discharge, and resource allocation.

Sources: [1], [2], [4], [5]

# Literature Review

Author/Year	Study focus	Method	Result (AUC)	Clinical impact
Hu et al., 2025	Predict ICU readmission within 72h	XGBoost + structured + radiology-note embeddings	0.78	Outperformed SWIFT; multimodal features improve prediction
Nguyen et al., 2021	Predict ICU need within 24h after inpatient admission	ML using EHR (vitals, labs, orders, dx codes)	0.82-0.88	Better than ESL; vital signs most predictive
Glass et al., 2021	Predicting ED→ICU transfer within 24h	Logistic regression using vitals + labs	0.70	Baseline performance for early deterioration detection
Gidari et al., 2020	Predicting in-hospital ICU admission among COVID-19 patients	NEWS2 early warning score	0.90	Shows vital signs can signal deterioration

# Problem Statement & Gap



1	Early ICU needs in the ED must be identified quickly, yet current tools (e.g., EWS, vitals-only scores) remain limited.
2	Predicting 72-hour ICU readmission is still challenging, even with modern ML models, leaving a safety risk after ICU discharge.
3	Existing models treat ED→ICU and ICU→ward transitions separately, hindering short-term resource planning.

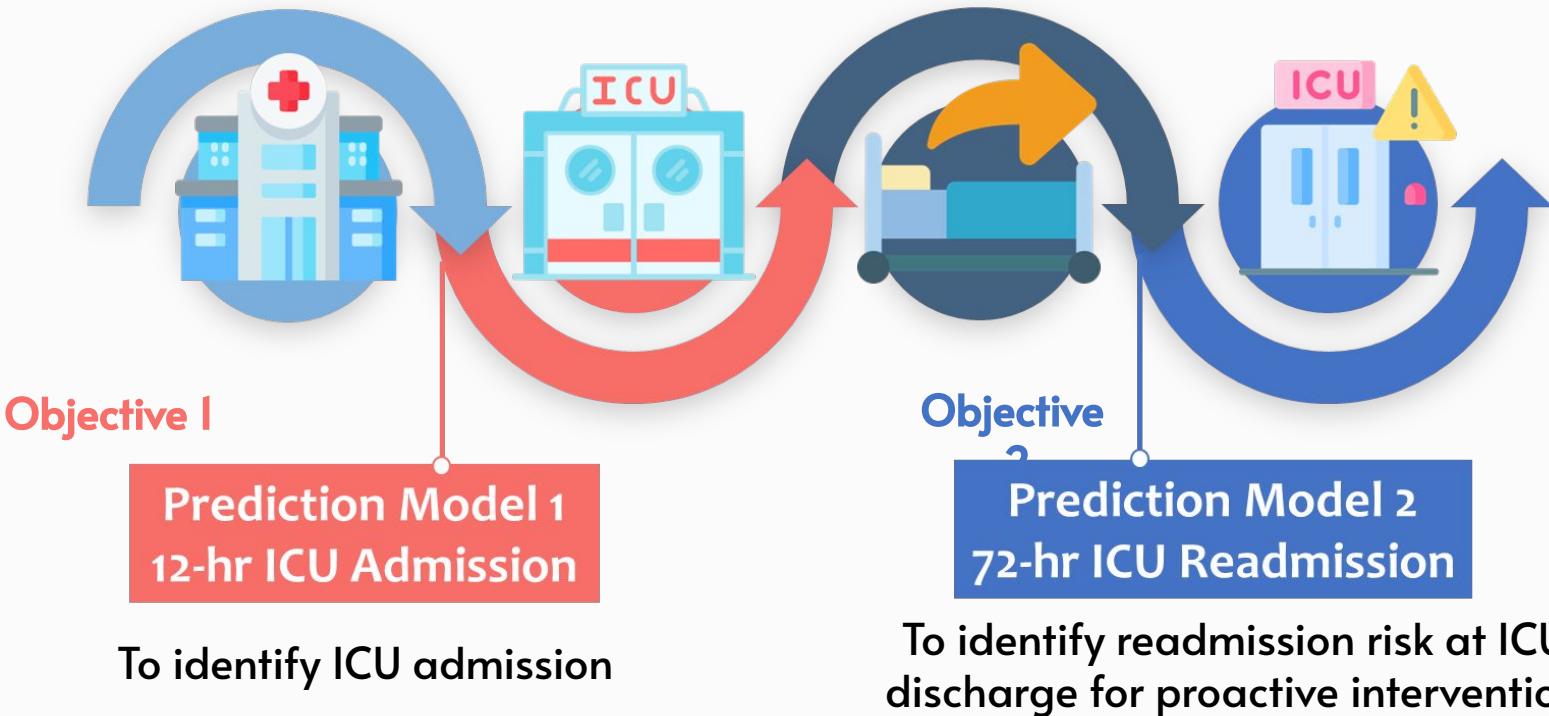
# Objective

ED Phase

ICU Phase

Discharge

ICU Readmission

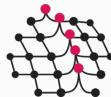




# Method

## Data Source

### PhysioNet



- MIMIC-IV-ED v2.2 — ED triage, initial diagnoses, early medications
- MIMIC-IV v3.1 — admissions, transfers, ICU stays, labs, clinical events
- Record linkage:  
subject\_id and hadm\_id used to integrate ED and hospital data

## Dataset we used

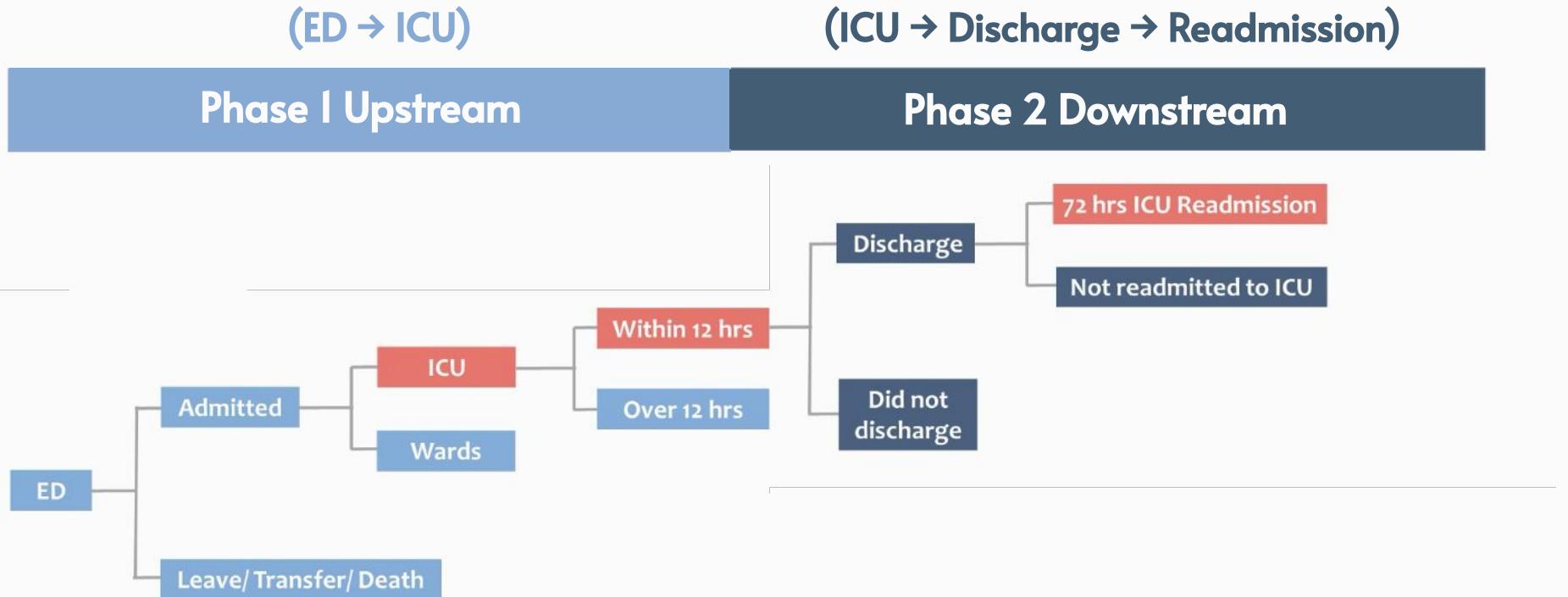
### HOSP

- admissions- basic admission info
- transfers- timestamped location changes
- patients- core demographic/ mortality info

### ICU

- icustays- icu level encounter records
- dateimeevents- icu time series clinical events
- inpuťevents- detailed medication/ IV records

# Study Design



- Cohort: ED visits → hospital admission
- Dependent Variable: `icu_within_12h`
- Negative: ward admission OR ICU after 12h

- Cohort: phase I positives + survived ICU stay
- Dependent Variable: `readmission_72h`
- Excluded: unit-to-unit transfers

# Feature Engineering/ Modeling/ Evaluation Metrics

## Feature Engineering

### ICU Stay Processing

- Length of Stay (LOS) is calculated using the ICU admission time (intime) and discharge time (outtime).

### Triage Processing

- Standardizing pain scores
- Chief-complaint normalization → top 25 categories

### Medication Feature Engineering

- Mapping >1,000 drug names → 22 pharmacologic classes
- Extracting class-level counts

## Modeling

XGBoost

## Evaluation Metrics

AUC, Accuracy,  
Recall, Precision



# Results

# Descriptive Results

## Demographics

### Key Differences between ICU & Non-ICU Groups:

- Demographics:

- Older (62 vs 49 years), male-predominant (54% vs 45%)
- More likely white (64% vs 57%) and married (40% vs 16%)
- Medicare-insured (55% vs 20%)

- Arrival Pattern:

- Ambulance transport (64.1% vs 34.4%)
- Walk-in less common (21.4% vs 62.3%)

- Clinical Outcomes:

- Shorter ED stay (5.6 vs 7.3 hours)
- Higher 30-day mortality rate (15.3% vs 1.0%)

Table1. 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*

# Predictive Modeling

## Results

### Model 1 12-hr ICU Admission

AUC  
**0.96**

Accuracy  
**0.87**

Recall  
**0.93**

Precision  
**0.30**

### Model 2 72-hr ICU Readmission

AUC  
**0.72**

Accuracy  
**0.79**

Recall  
**0.52**

Precision  
**0.12**

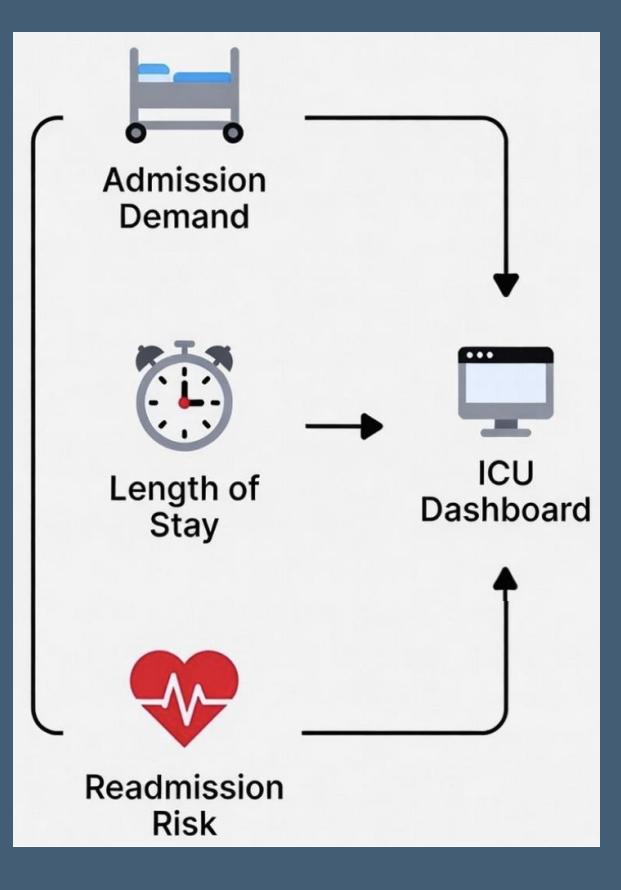
### Model Interpretation

- The model shows strong discrimination (AUC = 0.96)
- High recall (0.93) indicates it can successfully identify most patients who will require ICU care
- Low precision (0.30) reflects many false positives, likely due to the rarity of true ICU transfers

### Model Interpretation

- The model shows moderate discrimination (AUC = 0.72)
- Recall is low (0.52), meaning many true readmissions are still hard to detect
- Very low precision (0.12) reflects the rarity and complexity of post-ICU deterioration

# Combine Models: ICU Bed Demand Forecasting Framework



1

## Model 1 - 12-hr ICU Admission Demand

Assigns each ED patient a probability of requiring ICU care within 12 hours. Aggregating these **provides a short-term estimate of incoming ICU demand**

2

## ICU Length of Stay

Uses historical LOS data to **estimate how long ICU beds remain filled**

3

## Model 2 - 72-hr ICU Readmission Risk

Assess discharge suitability and flag patients who are likely to return, **improving short-term estimates of incoming ICU**

4

## ICU Dashboard

Offers a complete view of ICU patient flow and helps clinicians anticipate demand, allocate resources, and manage discharge timing

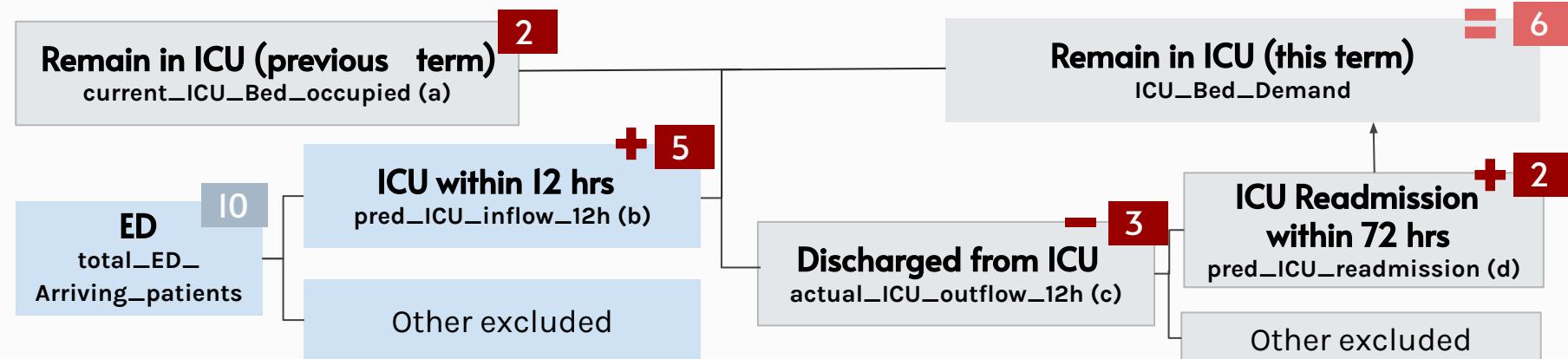
# Demo- ICU Bed Forecast Result

(ED → ICU)

(ICU → Discharge → Readmission)

Phase I Upstream

Phase 2 Downstream



timestamp	total_ED_Arriving_patients	current_ICU_Bed_occupied (a)	pred_ICU_inflow_12h (b)	actual_ICU_outflow_12h (c)	Pred_ICU_readmission (d)	ICU_Bed_Demand (a+b-c+d)
2150/9/1 12:00	10	2	5	3	2	6
2150/9/2 00:00	4	6	0	0	0	6
2150/9/2 12:00	11	6	3	1	0	8



# Discussion & Conclusion

# Discussion

## Key Contribution

- Integrated ED-to-ICU admission and ICU readmission prediction into one forecasting framework
- Achieved strong performance for early ICU admission prediction using ED data alone (AUC 0.96)
- Demonstrated a proof-of-concept ICU capacity forecasting system combining predicted inflow, LOS, and readmission risk

## Study Limitations

- Limited Generalizability- MIMIC-IV originates from one hospital
- Timestamp Shifting- MIMIC-IV dates are shifted into the 2100–2200 range which may affect time-dependent features
- Proof-of-Concept Framework- Deployment requires LOS prediction + multi-site validation

# Conclusion

## Key Findings

- Routine EHR data predicts ICU admission (AUC 0.96) and readmission (AUC 0.72)
- Triage features provide early signals of deterioration
- Framework connects ED inflow, ICU occupancy, and readmission risk
- Models offer practical decision support for triage and ICU discharge planning, helping anticipate bed shortages and reduce unsafe discharges.

## Future Work

- Multi-site validation using real-time clinical data
- Real-time deployment with streaming prediction pipelines
- LOS prediction model to replace LOS for accurate capacity forecasting
- Richer data inputs (notes, labs, physiological trends, imaging) to improve performance

## 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] He, S. 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): 2423–2431. [PMCID: PMC8561936]
- [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]

Thank you.

# ICU Bed Forecast Results

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

## Data Source



## MIMIC-IV-ED v2.2

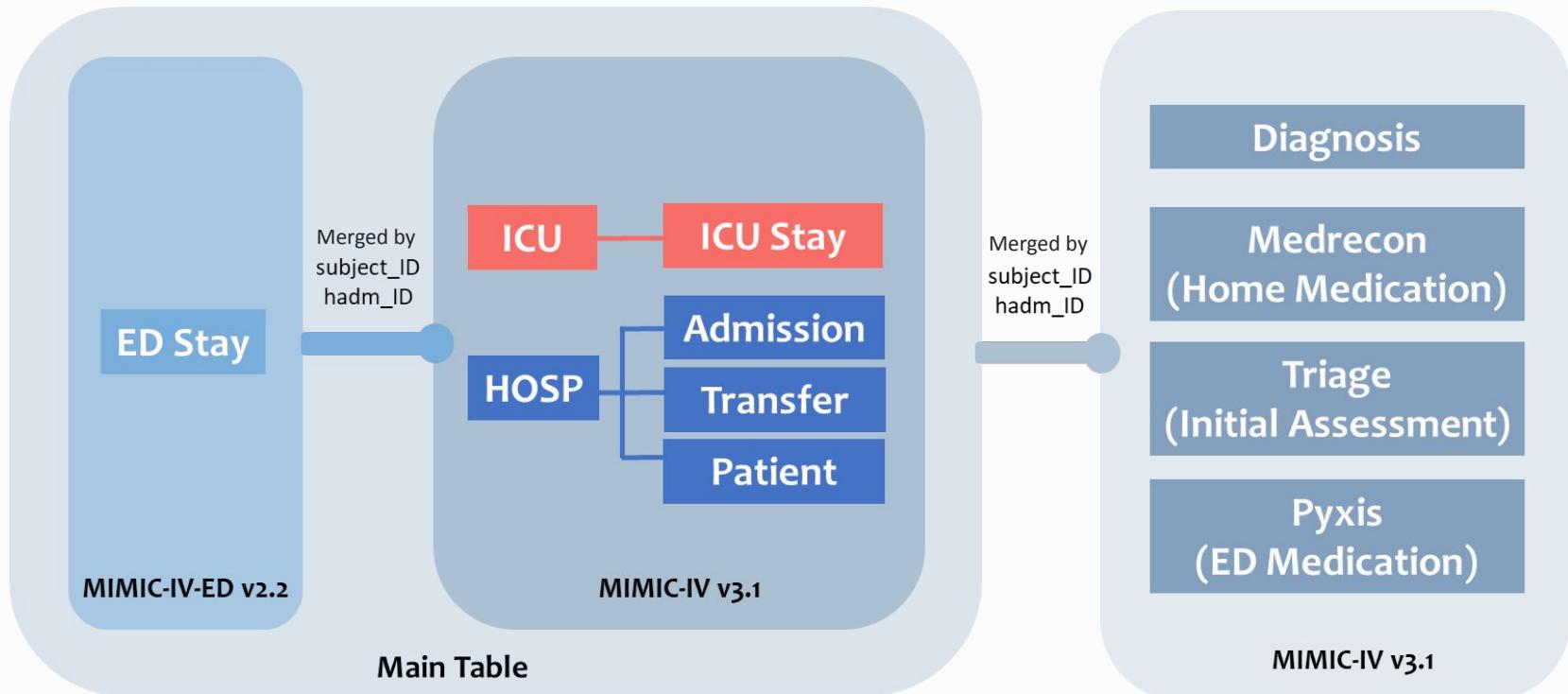
Section	ED (Emergency Department)
Original Data Name	Ed stays, diagnosis, triage (Initial Assessment), vital sign, medrecon (Home Medication), pyxis (ED Medication)
Key	subject_id (unique id for each patient), hadm_id (unique id for each hospital admission), Ed_stay (unique id for each ED visit)
Content Description	Emergency department visit data

## MIMIC-IV v3.1

Section	HOSP (Hospitalization)	ICU (Intense Care Unit)
Original Data Name	Admissions, transfers, patient (Demography)	ICU Stays
Key	Subject_id, hadm_id	subject_id, hadm_id
Content Description	all hospitalization records, transfer records, and patients' informations	Intensive Care Unit visit data

## 4. Data

# MIMIC-IV Data Merging



## 6. Current Status & Future Steps

### Future Steps

#### Model Phase 1 12-hr ICU Admission

- I. Train a logistic regression model using 0–3 hrs ED features
- II. Predict the probability of ICU admission within 12 hours
- III. Determine an optimal probability threshold as the alert cutoff

#### Model Phase 2 72-hr ICU Readmission

- I. Train an XGBoost model using ICU features from a patient's last 6–12 hrs
- II. Predict the probability of ICU readmission within 72 hours after discharge
- III. Apply SHAP to identify key factors that help alert physicians during discharge decisions

#### Make Scores Operational

- I. Convert predicted probabilities into hourly ICU bed demand
- II. Evaluate against a seasonal or moving-average baseline using simple time-series methods