Deep Learning for Early Prediction of Multi-drug Resistant Organisms in the Emergency Department

YooJin Chloe Sheen[1*], Kyra D. Bowden[2*], Boaz Goldberg[2*], Yuhan Gui[2*], Sidharth Raghavan[2*]
Joseph L. Greenstein[2], Casey Overby Taylor[1,2], Eili Y. Klein[3], Jeremiah S. Hinson[3]
[1]Johns Hopkins University School of Medicine, [2]Whiting School of Engineering, Johns Hopkins University,
[3]Department of Emergency Medicine, Johns Hopkins University School of Medicine | Baltimore, MD

Introduction

- 1 in 10 hospital patients acquire a **healthcare-associated infection** (HAI).[1]
- Multi-drug Resistant Organism (MDRO) infections increase risk of patient readmission, which typically leads to higher mortality rates.[2]
- The emergency department (ED) is a fast-paced decision-making environment, and patients are often prescribed **inappropriate antibiotics**.
- Current statistical models often overlook crucial transmission factors and have limited predictive accuracy.[3]
- Clinicians need **automated MDRO risk assessments** for patients in the ED to improve empirical prescriptions of antibiotics and reduce the risk of antibiotic resistance.

Transmissions occur via contacts

Our approach enhances the hospital context by incorporating **contact networks**.

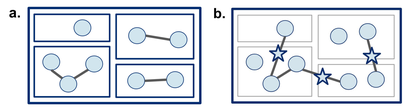

Fig. 1 **Graph-based contact model.** (a) Room-mediated contacts: edges connect patients who shared a room concurrently or sequentially within an optimized lookback period. (b) HCW-mediated graph: edges connect patients seen by the same healthcare worker within the lookback period.

Objectives

We propose a ML model that provides interpretable risk predictions to guide antibiotic decisions in the ED.

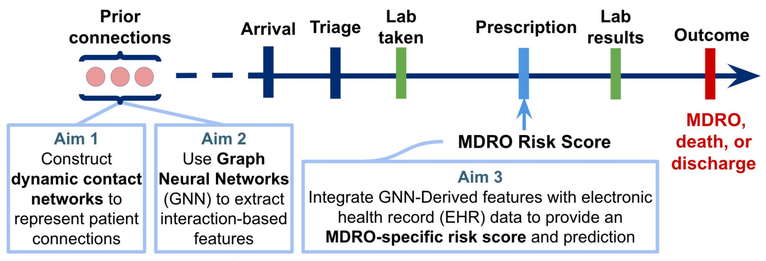

Fig 2. High-risk MDRO patient model timeline with targeted interventions.

Cohort Characterization

Features	MDRO+ (n = 31172)	non-MDRO (n = 119802)
Age (Mean)	53.8	54.8
Gender (Male)	15728	70581
Gender (Female)	15436	49188
Length of stay in hospital (days)	158.8	59.0

Table 1. **Selected characteristics of patient cohort** (n = 128852) derived from six hospitals within the Johns Hopkins Medical System from 2016 - 2024.*

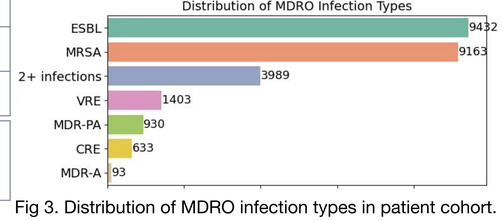

Fig 3. Distribution of MDRO infection types in patient cohort.

**Due to dynamic nature of MDRO infections, the MDRO+ and non-MDRO infection groups include patient overlap.*

Methods

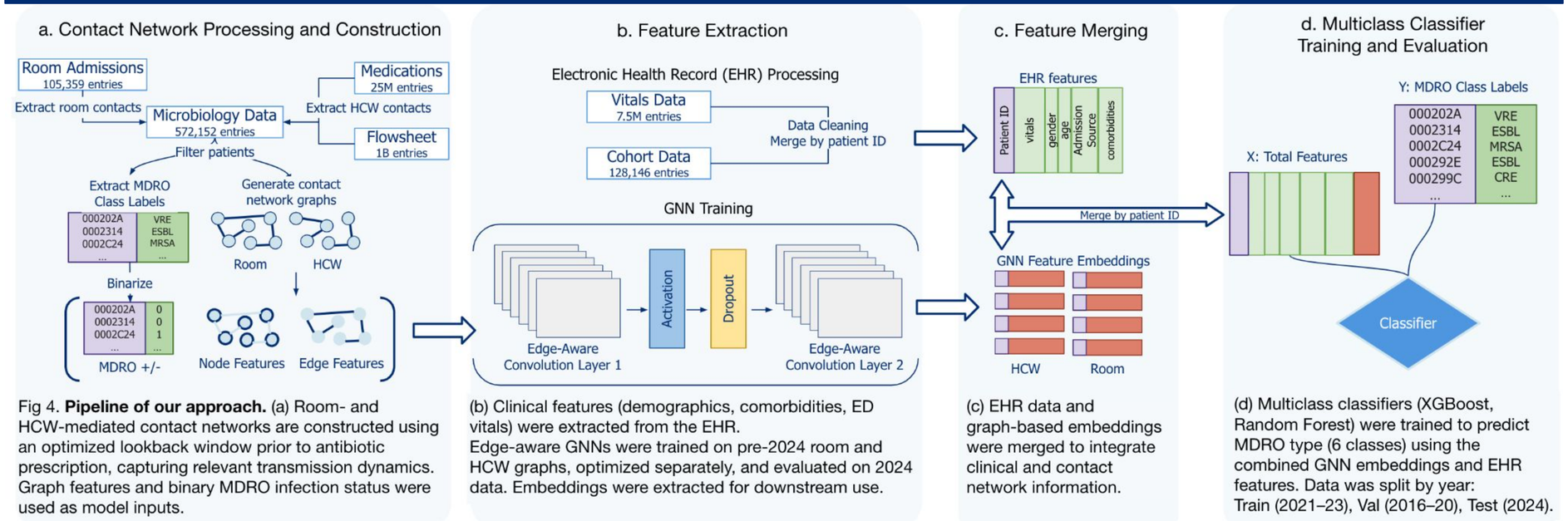

Fig 4. **Pipeline of our approach.** (a) Room- and HCW-mediated contact networks are constructed using an optimized lookback window prior to antibiotic prescription, capturing relevant transmission dynamics. Graph features and binary MDRO infection status were used as model inputs.

(b) Clinical features (demographics, comorbidities, ED vitals) were extracted from the EHR. Edge-aware GNNs were trained on pre-2024 room and HCW graphs, optimized separately, and evaluated on 2024 data. Embeddings were extracted for downstream use.

(c) EHR data and graph-based embeddings were merged to integrate clinical and contact network information.

(d) Multiclass classifiers (XGBoost, Random Forest) were trained to predict MDRO type (6 classes) using the combined GNN embeddings and EHR features. Data was split by year: Train (2021–23), Val (2016–20), Test (2024).

Results

Contact networks reveal connections between MDRO-infected patients.

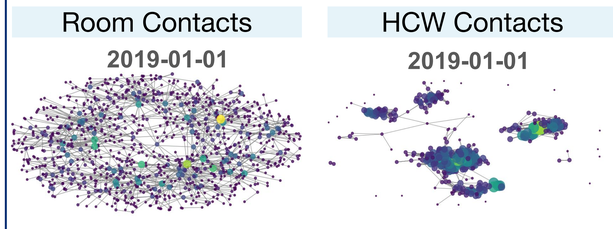

Fig 5. **Two-day window contact networks for 2019-01-01.** The higher degree and fewer components in the HCW -mediated contact networks compared to the room-mediated contact networks indicate greater connectivity between patient groups through healthcare workers.

Our classifier can distinguish between MDRO and non-MDRO infections.

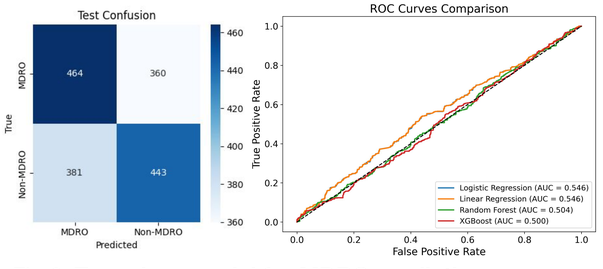

Fig 6. **Two-class model for MDRO prediction.** Patients were classified by logistic regression on 8 PCs of each contact network's GNN embedding type. Accuracy: 55%; Sensitivity: 56%; Specificity: 54%; F1: 56%; Cohen's Kappa: 10%.

Interpretability

- 64-dim GNN embeddings were reduced to 8 Principal Components per connection type (16 total).
- PCA improved ROC-AUC vs. using 2, 4, 16, 32, or 64 raw features.
- Logistic regression outperformed XGBoost and Random Forest on validation.
- PCs are linear combinations of GNN features and improve model performance but limit direct interpretability and further work is needed to map them back.

Future Directions and Use Case

- Addition of MDRO class prediction to final classifier.
- Federated learning to incorporate data from multiple sites into MDRO prediction.
- Integrate model predictions into real-time clinical workflows to support empiric antibiotic decision-making.

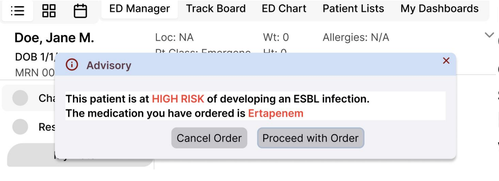

Fig 7. Concept of order entry screen with best advisory warning.

References

1. Cui J et al. *Identifying Importation and Asymptomatic Spreaders of MDROs in Hospitals.* medRxiv. 2024.
2. Burnham JP et al. *Readmissions with MDRO Infection in Patients with Prior MDRO Infection.* Infect Control Hosp Epidemiol. 2018;39(1):12–19.
3. Prakash BA et al. *Identifying Importation and Asymptomatic Spreaders of MDROs in Hospitals.* Res Square. 2024 Jul 18.