

# Predicting outcome in emergency room patients with Suspected Gastrointestinal Infection using OMOP-CDM

So Hee Lee<sup>1</sup>, Byungjin Choi, MD<sup>1,2</sup>, Min Ho An, MD<sup>1,2</sup>, Junhyuk Chang<sup>1</sup>, Harrin Kim<sup>1</sup>, Rae Woong Park, MD, Ph.D<sup>2</sup>

<sup>1</sup>Department of Biomedical Sciences, Ajou University Graduate School of Medicine, Suwon, Republic of Korea

<sup>2</sup>Department of Biomedical Informatics, Ajou University School of Medicine, Suwon, Republic of Korea

## Background

Gastroenteritis is inflammation of the gastrointestinal tract, primarily caused by viruses such as rotaviruses and noroviruses, which account for 50% to 70% of cases.<sup>1</sup> They affect people of all ages, are very contagious, and can cause serious complications if diagnosis or treatment is delayed. There is consensus that early prediction is important to improve patient outcomes.<sup>2</sup> In this study, we aimed to predict the estimated duration of treatment and the severity of suspected acute gastroenteritis to prevent transmission through prompt and appropriate treatment.

## Methods

We used Ajou University School of Medicine (AUSOM) Observation Medical Outcomes Partnership Common Data Model (OMOP-CDM) database, which had been collected from 1994 to 2024. We extracted the data of patients who visited the emergency room (ER) with symptoms of abdominal pain, vomiting, and diarrhea accompanying fever. Outcomes of interests were ER revisits, Intensive care unit (ICU) admissions, and mortality within 7 days. Demographic variables, conditions, visits, medications and measurements were used to develop machine models to predict outcome variables. We developed prediction models using the Gradient Boosting Machine (GBM) and a Least Absolute Shrinkage and Selection Operator algorithms (LASSO). The area under the receiver operating characteristic curves (AUROC) was calculated to evaluate the models' performance. The study population was randomly split into the train (75%) and test set (25%) in 3-fold cross validation. The Cox proportional model was developed for calculating the hazard ratio and the Kaplan-Meier method was applied for plotting the survival curve. We used ATLAS 2.8.0 and R version 4.1.0 for all analyses in this study.<sup>3</sup>

## Results

Among 196,779 patients selected based on the inclusion and exclusion criteria, 101,518 (50.77%) were male and their mean age was  $19.63 \pm 24.17$ . For ER revisits, ICU admissions, and mortality, the number of each outcome patients was 230, 1,091 and 788, respectively.

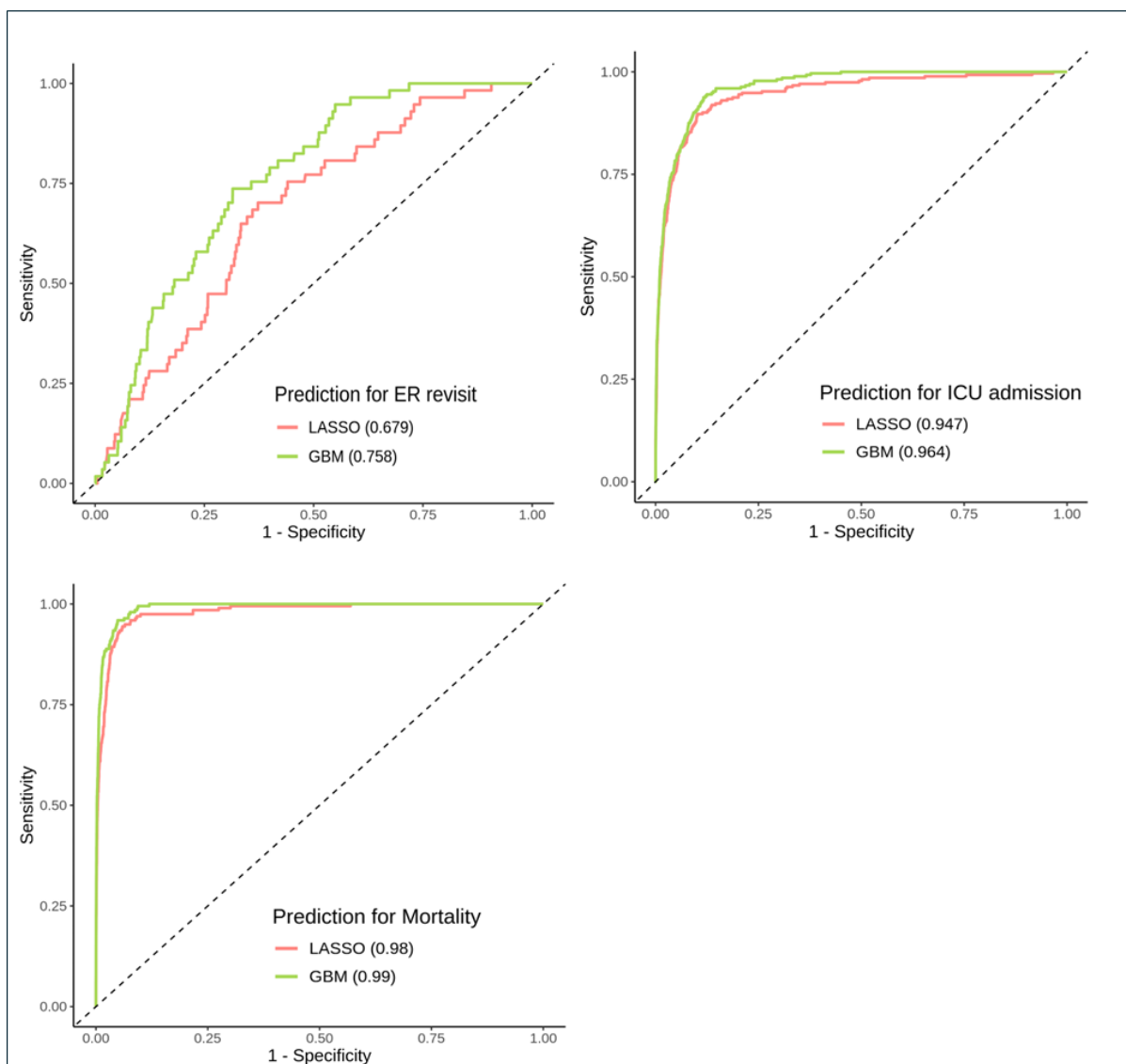
For all outcomes, the GBM model outperformed LASSO model for all outcomes. LASSO's performance shows AUROCs of 0.679, 0.947, and 0.980, respectively, while GBM's performance shows AUROCs of 0.758, 0.964, and 0.990, respectively (Table 1, Figure 1).

In the additional analysis, we classified the high and low-risk groups in the test set of the ICU admission model developed using the GBM algorithm and performed a survival probability analysis on the ICU admission outcome. The low-risk group had a lower hazard ratio (HR) compared to the high-risk group for ICU admission (HR 0.013, 95% CI [0.009 – 0.019];  $p < 0.0001$ , Figure 2).

**Table 1. Performance of LLR and GBM prediction model.**

Model	Outcome	Target Counts	Outcome Counts	AUROC
<b>LASSO</b>	ER revisit	186,749	230	0.679 (0.617-0.741)
	ICU admission	186,749	1,091	0.947 (0.932-0.962)
	Mortality	186,973	788	0.980 (0.972-0.987)
<b>GBM</b>	ER revisit	186,749	230	0.758 (0.710-0.807)
	ICU admission	186,749	1,091	0.964 (0.957-0.972)
	Mortality	186,973	788	0.990 (0.988-0.993)

Abbreviations: LASSO, Least Absolute Shrinkage and Selection Operator; GBM, Gradient Boosting Machine; ER, Emergency room; LOS, Length of stay; ICU, Intensive care unit.



**Figure 1. Performance of LASSO and GBM prediction model for severity of gastroenteritis**

Abbreviations: LASSO, Least Absolute Shrinkage and Selection Operator; GBM, Gradient Boosting Machine; ER, Emergency room; ICU, Intensive care unit;

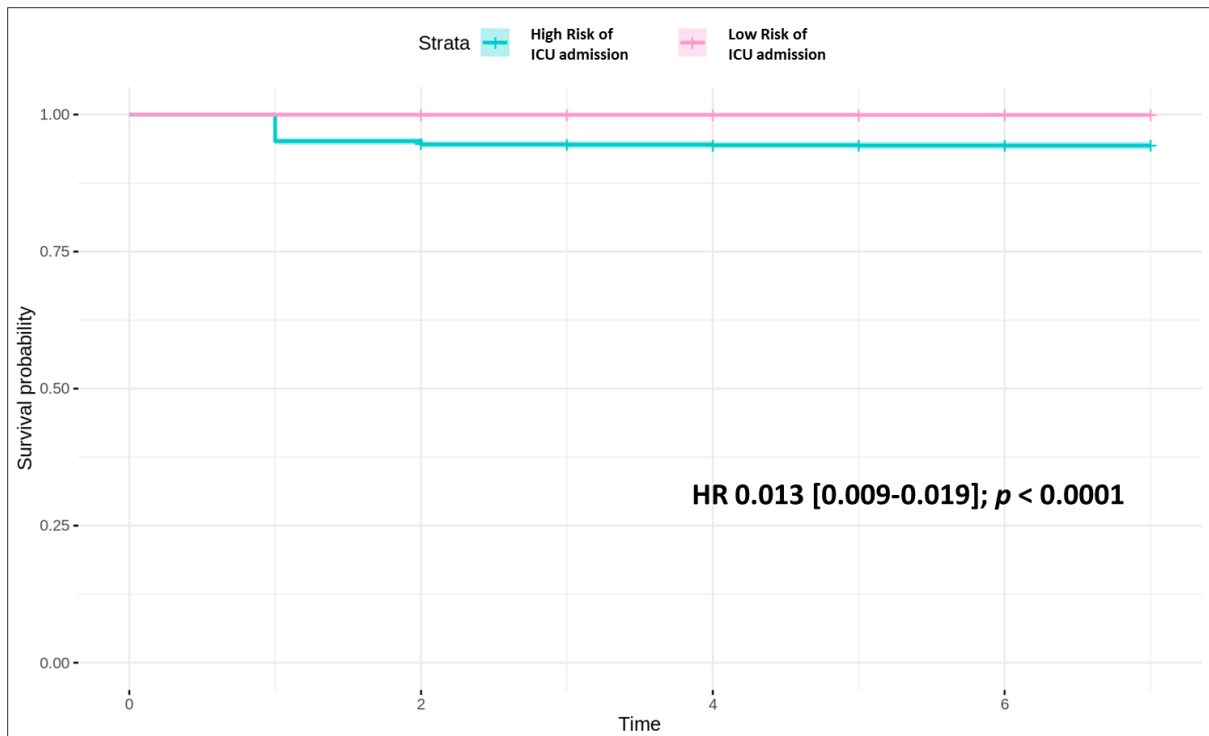


Figure 2. Kaplan-Meier survival analysis of Gradient boosting machine for ICU admission.

### Conclusion

In this study, we developed CDM-based prediction models to assess the severity of outcomes such as ER revisits, ICU admissions, and mortality in patients presenting with symptoms of acute gastroenteritis. These models exhibited high model performance in predicting the severity of gastroenteritis within 7 days and effectively differentiated between high-risk and low-risk groups for ICU admission survival probability. By facilitating rapid and appropriate treatment, these models have the potential to significantly improve patient outcomes for those with suspected infectious disease symptoms.

### Acknowledgement

This research was funded by a grant from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HR16C0001) and this research was supported by a Government-wide R&D Fund project for infectious disease research (GFID), Republic of Korea (grant number: HG22C0024, KH124685).

### References/Citations

1. Graves NS. Acute gastroenteritis. Prim Care. 2013 Sep;40(3):727-41.
2. Bányai K, Estes MK, Martella V, Parashar UD. Viral gastroenteritis. Lancet. 2018 Jul 14;392(10142):175-186.
3. Atlas Version 2.8.0 <https://github.com/OHDSI/Atlas/releases/tag/v2.8.0>