Beyond Traditional Scoring Systems: Developing an AI-Powered Tool for Trauma Prognosis

Sally Hong, MIDS [a], H. Claude Sagi, MD [c], Charles Butts, MD [d], Adam J. Starr, MD [a f], Caroline Park, MD [a e], Vikas Chowdhry, MS MBA [a], and Michael D. Goodman, MD [b]

[a] TraumaCare.AI, INC, Dallas, TX

[b] Department of Surgery, University of Cincinnati, Cincinnati, OH

[c] Department of Orthopaedic Surgery, University of Cincinnati, Cincinnati, OH

[d] Department of Trauma & Acute Care Surgery, University of South Alabama College of Medicine, Mobile, AL

[e] Department of Surgery, University of Texas Southwestern Medical Center, Dallas, TX

[f] Department of Orthopedic Surgery, University of Texas Southwestern Medical Center, Dallas, TX

Note*: Abstract based on unpublished work with University of South Alabama College of Medicine and recently published research with University of Cincinnati [1]*

Introduction

Trauma is the leading cause of death for individuals under 45 years of age in the United States. Timely and accurate evaluation of polytrauma patients is critical to prioritizing treatment and optimizing outcomes. Traditional scoring systems, such as the Trauma and Injury Severity Score, have supported clinical decision-making but are limited by low specificity, static nature, and reliance on data not routinely captured in Electronic Health Records (EHRs) *[8,9]*. Advances in artificial intelligence (AI) and EHR integration and data-interoperability standards present an opportunity to create predictive tools tailored to local patient data, including demographics, injury patterns, and dynamic physiological changes. Such tools can deliver real-time, actionable insights to enhance trauma patient management *[2,3,6]*.

Approach

We utilized retrospective data from two Level 1 trauma centers to develop machine learning models predicting 48-hour mortality in polytrauma patients. Variables were selected based on real-time availability in trauma workflows and feedback from clinical experts (trauma surgeons, ortho-trauma surgeons, and critical care nurses) regarding their clinical utility *[1]*. To align with workflow requirements, a proof-of-concept application was developed to enhance model interpretability and actionability (*Figure 2*). Additionally, to test generalizability, we trained and tested the machine learning model using variables previously utilized in published studies at Parkland Health *[1,3]*.

Machine Learning Methods and Performance

Our models aimed to predict 48-hour mortality using real-time, routinely collected data. The target statistical threshold, set in consultation with clinicians, was maximum precision at 80% recall, with predictions initiated within the first hour post-admission.

Patient data were processed into hourly snapshots, with an 80/20 temporal train/test split. Feature engineering addressed clinical data variability, applied PCA for feature correlation, and generated summary statistics (e.g., min/max/mean) for frequent measurements. XGBoost was chosen for its scalability, robustness with missing data, and interpretability. A bivariate model was developed to approximate clinicians' gestalt, estimating that clinicians achieve 10% precision at the target recall threshold. This benchmark contextualized the performance improvements offered by the machine learning model.

The models achieved precision rates of 35% and 48% at 80% recall for the two trauma centers, significantly surpassing the clinician-estimated baseline *[Figure 1]*. Notable top predictors included GCS, blood pressure, and lactate at one center, and GCS, BUN, and potassium at the other. Performance variability was influenced by differences in data quality, availability, and patient demographics.

Conclusion & Future Work

We successfully developed machine learning models for predicting mortality in polytrauma patients and demonstrated the generalizability of model variables across institutions. Collaboration with clinical experts enhanced the interpretability and actionable nature of the model outputs *[1]*.

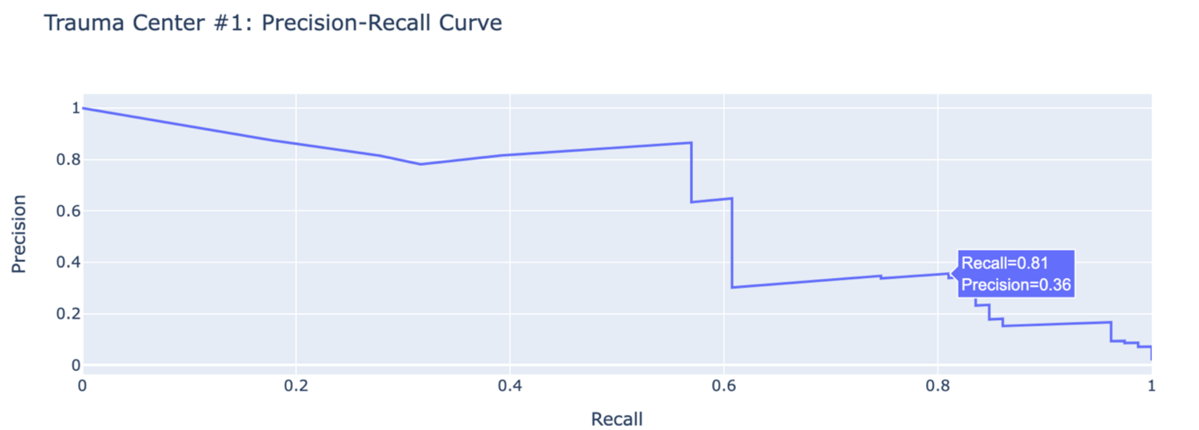
Future work focuses on integrating the proof-of-concept application with major EHR platforms using interoperability standards such as FHIR APIs and SMART on FHIR. Once deployed, we plan to conduct prospective evaluations to measure the model's impact on clinical outcomes and process improvement metrics.

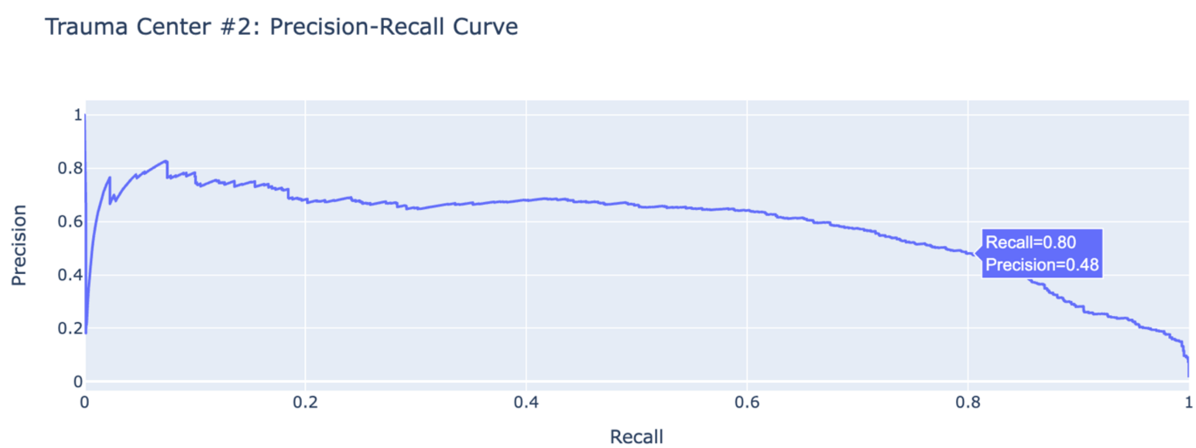
Tables and Figures

Table 1: Dataset Overview

	Trauma Center #1	Trauma Center #2
Number of Encounters	4,800	7,000
Mortality Rate	4.5%	7.0%
Model Precision at ~80% Recall	35.6%	48.1%

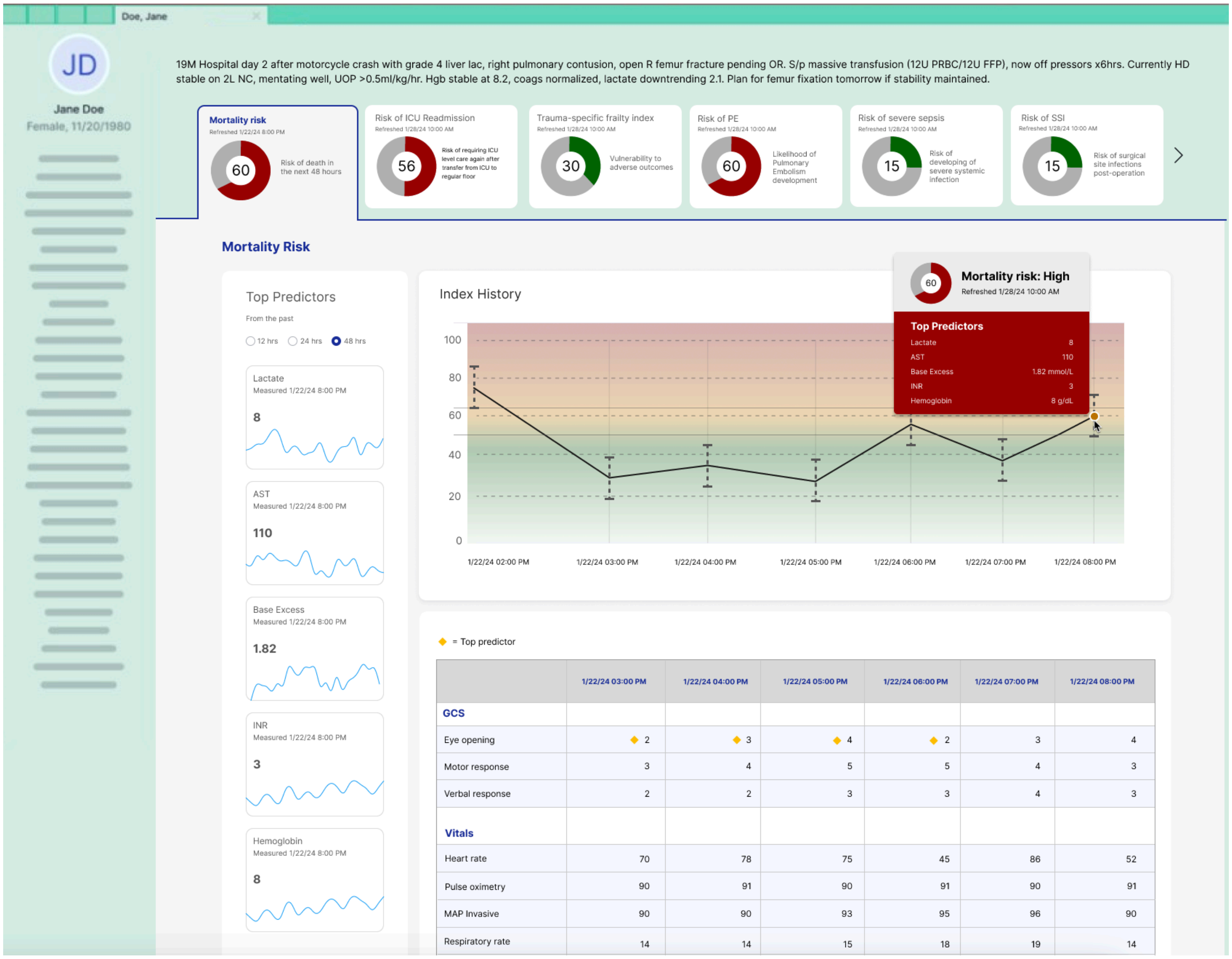
Figure 1: Precision Recall Curves





Note: The PR curve is smoother for Trauma Center #2 due to more data points (more encounters).

Figure 2: User interface to convey model output (© 2025 TraumaCare.AI, INC).



References

[1] Becker ER, Price AD, Barth J, Hong S, Chowdhry V, Starr AJ, Sagi HC, Park C, Goodman MD. External validation of predictors of mortality in polytrauma patients. *J Surg Res.* 2024;301:618-622. doi:10.1016/j.jss.2024.07.024

[2] Park C, Loza-Avalos SE, Harvey J, Hirschkorn C, Dultz LA, Dumas RP, Sanders D, Chowdhry V, Starr A, Cripps M. A real-time automated machine learning algorithm for predicting mortality in trauma patients: Survey says it's ready for prime-time. *Am Surg.* 2023;0(0):1-7. doi:10.1177/00031348231207299

[3] Starr AJ, Julka M, Nethi A, Watkins JD, Fairchild RW, Rinehart D, Park C, Dumas RP, Box HN, Cripps MW. Parkland Trauma Index of Mortality: Real-time predictive model for trauma patients. *J Orthop Trauma.* 2022;36(6):280-286. doi:10.1097/BOT.0000000000002290

[4] Qi J, Bao L, Yang P, Chen D. Comparison of base excess, lactate and pH predicting 72-h mortality of multiple trauma. *BMC Emerg Med.* 2021;21(80):1-10. doi:10.1186/s12873-021-00465-9

[5] Mica L, Niggli C, Bak P, Yaeli A, McClain M, Lawrie CM, Pape HC. Development of a visual analytics tool for polytrauma patients: Proof of concept for a new assessment tool using a multiple layer Sankey diagram in a single-center database. *World J Surg.* 2020;44(3):764-772. doi:10.1007/s00268-019-05267-6

[6] Hamad DM, Subacius H, Thomas A, Guttman MP, Tillmann BW, Jerath A, Haas BM, Nathens AB. A multidimensional approach to identifying high-performing trauma centers across the United States. *J Trauma Acute Care Surg.* 2024;97(1):125-133. doi:10.1097/TA.0000000000004319

[7] Bläsius FM, Laubach M, Andruszkow H, Lichte P, Pape HC, Lefering R, Horst K, Hildebrand F, Trauma Register DGU. Strategies for the treatment of femoral fractures in severely injured patients: trends in over two decades from the TraumaRegister DGU. *Eur J Trauma Emerg Surg.* 2021;47(2):487-495. doi:10.1007/s00068-020-01599-4

[8] Tran Z, Verma A, Wurdeman T, Burruss S, Mukherjee K, Benharash P. ICD-10 based machine learning models outperform the Trauma and Injury Severity Score (TRISS) in survival prediction. *PLoS One.* 2022;17(10):e0276624. doi:10.1371/journal.pone.0276624

[9] Halvachizadeh S, Baradaran L, Cinelli P, et al. How to detect a poly-trauma patient at risk of complications: a validation and database analysis of four published scales. PLoS One. 2020;15:e0228082.