UNCLASSIFIED

The Development and Evaluation of a Machine Learning Model to Predict Trauma Related Infections

Sidney Cox MS[1], Nicole Caldwell MA, RN[1], David Luellen MBA, BS[1], Alicia Williams MD[1], Jose Salinas PhD[1], David Tribble MD, DrPH[2], Laveta Stewart PhD[2,3], Nusrat Epsi PhD[2,3], Sena Mike MS[1]

[1]US Army Institute of Surgical Research, JBSA, San Antonio, TX

[2]Infectious Disease Clinical Research Program, Department of Preventive Medicine and Biostatistics, Uniformed Services University of the Health Sciences, Bethesda, MD

[3] Henry M. Jackson Foundation for the Advancement of Military Medicine, Inc, Bethesda, MD

Introduction

- From 2009 to 2012, one third of combat casualties developed an infection during their initial hospitalization [1].
- Early identification of infection is severely hindered by a reliance on outdated technology that may not deliver timely detection in a combat environment [2].
  - Diagnostic delays are compounded in remote, combat environments as technology and access to experienced providers is limited.
- Machine learning (ML) models have shown promise in predicting important clinical outcomes such as admittance to ICU and cancer prognosis [3-4].
- For this study, we explored the medical information mart for intensive care (MIMIC IV) database which provides prehospital clinical data that is necessary for early prediction of infection compared to fragmented military datasets.

Objectives

- Develop and compare ML models for early prediction of infection in trauma patients.
- Identify the most important predictors associated with infection.
- Evaluate the performance of machine learning models for infection prediction.

Methods

- Patient data were extracted from the MIMIC IV database.
- Patients were included if they received service by the trauma team, stayed in both the emergency department and the intensive care unit, were under the age of 65, had a total length of stay over 72 hours, and did not have a cardiac or cancer related diagnosis (Figure 1).
- Data were randomly divided into training and test sets using an 80/20 split.
- Synthetic minority over-sampling technique (SMOTE) was used to balance the data classes.
- Features were down selected using the least absolute shrinkage and selection operator (LASSO) penalty.
- Logistic regression, random forest, and gradient boosting ML models were evaluated.
- Performance was evaluated using accuracy, area under the curve (AUC), precision, sensitivity, and specificity.

Results

- A total of 341 patients were utilized for modeling (Table 1).
  - Our collaborators at USU ran similar ML models using a military specific dataset that displayed a significant difference in injury patterns [5].
- Two prediction models were developed: one using only demographic and injury variables and another incorporating early vital signs to determine whether physiological data improved prediction.
  - Vital signs included in the second model were heart rate, systolic blood pressure, diastolic blood pressure, temperature, respiratory rate, and oxygen saturation.

Characteristics	ISR Population
Sample size, N	341
Male	70%
Median age, years	49
Median hospital stay, days	7.19
Penetrating injuries	21
Mechanical ventilation	39%
Median First GCS*	15
Median First Shock Index	0.71

Table 1: MIMIC Patient Population Characteristics.
*Median time to first GCS recorded was 5.76 hours.

Figure 2: Patient Injury Breakdown.

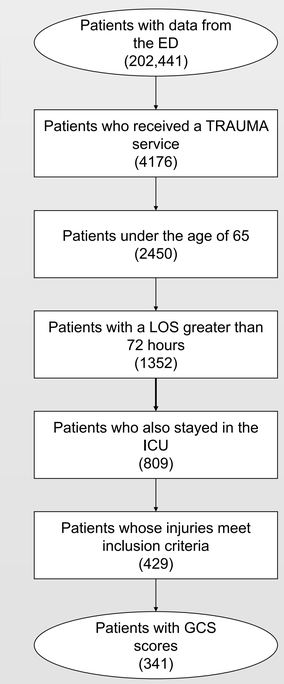

Figure 1: Patient Inclusion Criteria.

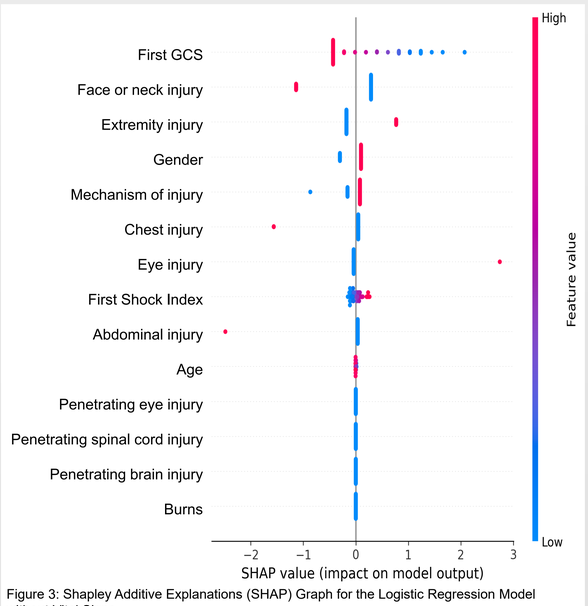

Figure 3: Shapley Additive Explanations (SHAP) Graph for the Logistic Regression Model without Vital Signs.

Results Cont'd

ML Models			Accuracy	AUC	Precision	Recall	Specificity	F1
Model with Vital Signs	*Logistic Regression*	Value (%)	77%	72%	19%	50%	79%	27%
	Gradient Boosting	Value (%)	**88%**	72%	**25%**	17%	**95%**	20%
	Random Forest	Value (%)	86%	63%	17%	17%	92%	17%
Model without Vital Signs	*Logistic Regression*	Value (%)	71%	**81%**	18%	**67%**	71%	**29%**
	Gradient Boosting	Value (%)	80%	54%	17%	33%	84%	22%
	Random Forest	Value (%)	80%	54%	17%	33%	84%	22%

Table 2: Machine Learning Model Results Comparison.

Conclusions

- The logistic regression ML model without vital sign data produced the best prediction ability (Table 2) with the highest AUC, recall, and F1.
- A patient's first GCS measurement was the most influential predictor of infection risk (Figure 3).
- The logistic regression model without vital sign data takes into consideration the class imbalance of the training data in order to effectively predict infection.
- A limitation is that our patient population lack the injury patterns and operational conditions seen in a combat environment.
- Future model development will include additional data from the BAMC trauma registry as well as data from our collaborators.
- Additionally, we are developing a time series ML model that can predict the risk of high consequence infection based on an accumulation of data over a patient's length of stay.

References:

1. Weintrob AC, Murray CK, Xu J, et al. Early infections complicating the care of combat casualties from Iraq and Afghanistan. *Surgical infections* 2018; **19**(3): 286-97.
2. Murray CK, Obremskey WT, Hsu JR, et al. Prevention of infections associated with combat-related extremity injuries. The Journal of trauma 2011; 71(2 Suppl 2): S235-57.
3. Escobar GJ, Turk BJ, Ragins A, et al. Piloting electronic medical record-based early detection of inpatient deterioration in community hospitals. Journal of hospital medicine: an official publication of the Society of Hospital Medicine 2016; 11 Suppl 1: S18-S24.
4. Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI. Machine learning applications in cancer prognosis and prediction. Comput Struct Biotechnol J 2015; 13: 8-17.
5. Stewart L, Epsi NJ, Chen D, et al. Infection Risk Prediction for Triage Decision Support of Combat Casualties. Poster presented at MHSRS 2026.

This study was conducted under an exempt protocol reviewed and approved by the US Army Medical Research and Development Command Institutional Review Board and in accordance with the approved protocol. The IDCRP has been funded by the National Institute of Allergy and Infectious Disease (NIAID), National Institute of Health (NIH), under Inter-Agency Agreement YI-AI-5082 and the Defense Health Program, U.S. DoW, under award HU0001190002. This project was supported through Defense Health Program award HU00012320031.

To learn more, visit: https://usaisr.health.mil