

A Real-Time Automated Machine Learning Algorithm for Predicting Mortality in Trauma Patients: Survey Says it's Ready for Prime-Time

The American Surgeon™
2023, Vol. 0(0) 1–7
© The Author(s) 2023



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/00031348231207299
journals.sagepub.com/home/asu



Caroline Park, MD, MPH¹, Sandra E. Loza-Avalos, MPH¹,
Jalen Harvey, MD¹, Carol Hirschhorn, ACNP², Linda A.
Dultz, MD, MPH¹, Ryan P. Dumas, MD¹, Drew Sanders,
MD³ , Vikas Chowdhry, MBA², Adam Starr, MD³, and
Michael Cripps, MD¹

Abstract

Background: Though artificial intelligence (“AI”) has been increasingly applied to patient care, many of these predictive models are retrospective and not readily available for real-time decision-making. This survey-based study aims to evaluate implementation of a new, validated mortality risk calculator (Parkland Trauma Index of Mortality, “PTIM”) embedded in our electronic healthrecord (“EHR”) that calculates hourly predictions of mortality with high sensitivity and specificity.

Methods: This is a prospective, survey-based study performed at a level I trauma center. An anonymous survey was sent to surgical providers and regarding PTIM implementation. The PTIM score evaluates 23 variables including Glasgow Coma Score (GCS), vital signs, and laboratory data.

Results: Of the 40 completed surveys, 35 reported using PTIM in decision-making. Prior to reviewing PTIM, providers identified perceived top 3 predictors of mortality, including GCS (22/38, 58%), age (18/35, 47%), and maximum heart rate (17/35, 45%). Most providers reported the PTIM assisted their treatment decisions (27/35, 77%) and timing of operative intervention (23/35, 66%). Many providers agreed that PTIM integrated into rounds and patient assessment (22/36, 61%) and that it improved efficiency in assessing patients’ potential mortality (21/36, 58%).

Conclusions: Artificial intelligence algorithms are mostly retrospective and lag in real-time prediction of mortality. To our knowledge, this is the first real-time, automated algorithm predicting mortality in trauma patients. In this small survey-based study, we found PTIM assists in decision-making, timing of intervention, and improves accuracy in assessing mortality. Next steps include evaluating the short- and long-term impact on patient outcomes.

Keywords

artificial intelligence, machine learning, trauma mortality

Introduction

Artificial intelligence and machine learning have become integrated in direct patient care for years and for a variety of applications, but with mixed durability and barriers to implementation. Traditionally, machine learning algorithms utilize information input from a clinician or extracted from the electronic health record (EHR) to develop a result to a specific query. This technology has been applied to a wide variety of clinical scenarios, including risk of surgical complications, burn burden assessment, and mortality amongst trauma patients.^{1,2} While many of these machine learning algorithms have been developed

¹Division of Burns, Trauma and Acute Care Surgery, Department of Surgery, University of Texas Southwestern Medical Center, Dallas, TX, USA

²Parkland Memorial Hospital, Dallas, TX, USA

³Department of Orthopedic Surgery, University of Texas Southwestern Medical Center, Dallas, TX, USA

Corresponding Author:

Caroline Park, MD, MPH, Department of Surgery, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390, USA.

Email: Caroline.Park@UTSouthwestern.edu

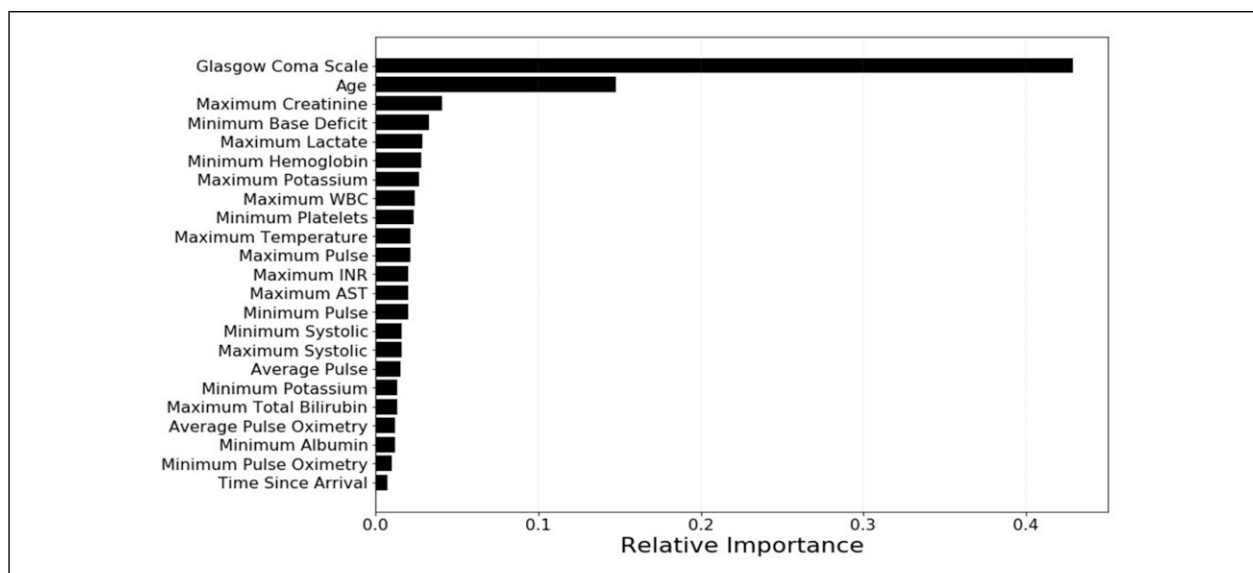


Figure 1. 23 variables included in the Parkland Trauma Index of Mortality (PTIM) model. Sample scoring with most significant variable listed at the top.

and validated, few have created a meaningful and lasting impact on provider practices and patient care. This is largely due to the retrospective nature of many existing machine learning algorithms and barriers to implementation (lack of integration into the EHR). Effective use of an algorithm requires real-time acquisition of data and prediction of outcomes to effectively deliver prophylactic treatment and adjust treatment courses.³ In this prospective, survey-based study, we evaluated provider practices on implementation of a real-time, automated machine learning algorithm embedded in the EHR to help predict mortality. We hypothesized that it would help improve communication between teams and impact utilization of resources and timing of interventions.

Methods

The Parkland Trauma Index of Mortality (PTIM) was developed and validated to predict the mortality risk of trauma patients over the next 48 hours from calculating a score based on objective variables (Figures 1 and 2).⁴ This tool is novel in that it is integrated directly into the EHR, extracts the data automatically, and calculates a PTIM score, requiring no input from the clinician (Figure 3). The clinician can then utilize this mortality prediction for planning, including operative intervention timing, deciding between damage control or definitive management, and goals of care. Surgical intensive care unit (SICU) providers, in addition to surgeon-specialists, were briefed on the implementation and utility of the PTIM score. For each instance, if the PTIM score was utilized in medical decision-making, an anonymous survey could be completed by the clinician via REDCap

(Research Electronic Data Capture, Vanderbilt University, Nashville, TN). The survey first queried what top 3 measurable factors (eg, Glasgow Coma Score (GCS), age, creatinine, and hemoglobin) the clinician felt contributed most to the patient's PTIM score. Next, the survey queried whether the score assisted in guiding the clinician's recommended treatment plan. Finally, the survey addressed the clinician's perceived ease of use, current utility, and future use. Summary statistics of survey outcomes were collected and organized using Stata IC 15 (StataCorp, College Station, TX). This study was approved by our institution's IRB and performed in compliance with the Declaration of Helsinki. Given the retrospective nature of patient data collection, informed consent was waived as the survey was anonymous; informed consent was waived by providers volunteering to participate.

Results

A total of 40 surveys were submitted, 35 of which reported using the PTIM score in medical decision-making. The factors that clinicians perceived as the top 3 predictors of mortality included GCS (22/38, 58%), age (18/35, 47%), and maximum pulse rate (17/35, 45%) which correlated with the algorithm's calculated actual most significant predictors of mortality (Figure 4(a) and (b)). Providers were also asked to provide what they felt were top predictors not scored on the algorithm, and they included "elevated intracranial pressure," "massive transfusion," "on-going bleeding," and "rising lactate." (Figure 5).

Respondents reported that the PTIM score assisted in determining the course of the treatment plan (27/35, 77%)

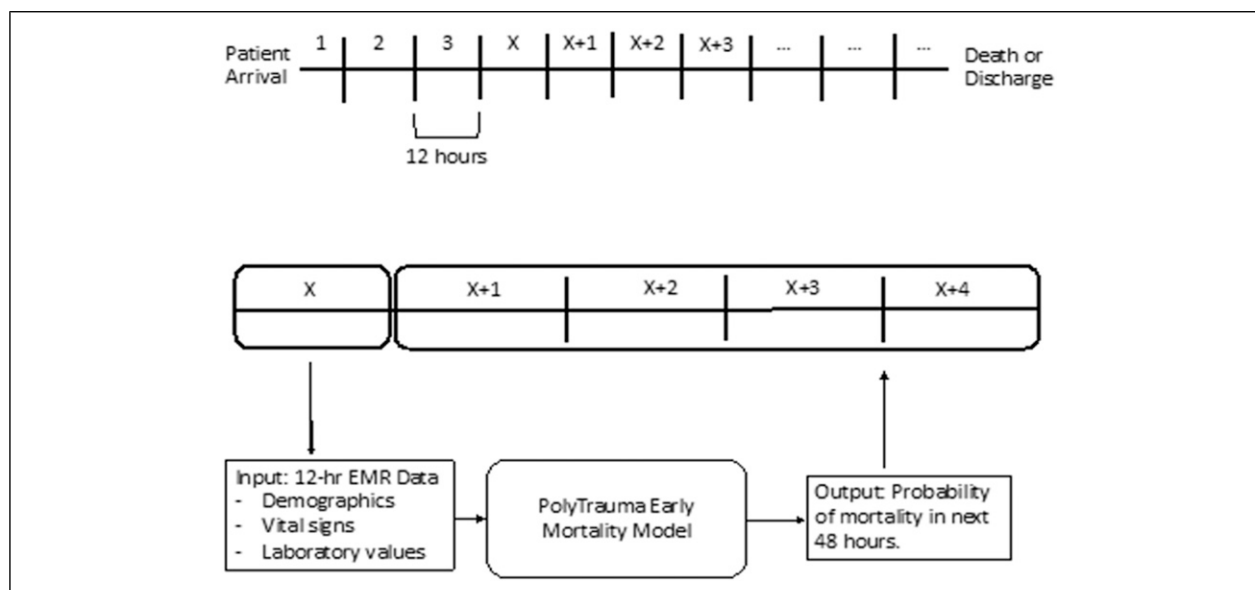


Figure 2. PTIM schematic. (Top) Patient encounters were divided into 12-hour time intervals from presentation to death or discharge. (Bottom) For each twelve-hour time interval, electronic health record data was used as PTIM input to predict mortality every hour in the next 48 hours.

and surgical intervention timing (23/35, 66%). When assessing the ease of use, clinicians reported (agree or strongly agree) that the PTIM score was easily integrated into rounds and patient assessment (22/36, 61%) and that it improved efficiency in assessing patients' potential mortality (21/36, 58%). Several clinicians reported that they would either continue to use the PTIM score when assessing patients (21/36, 58%) or were neutral on utilizing the score (15/36, 41%).

Discussion

Artificial intelligence, machine learning, and neural networks have been applied to patient care for over almost 3 decades, ranging from melanoma and burn wound recognition, predicting sepsis and transfusion needs, and mortality in trauma patients. These methods are becoming significantly more sophisticated and "intelligent," often surpassing the clinicians' ability to diagnose and treat. However, several of these predictive models are retrospective and are not readily available for real-time decision-making.

Several studies have applied teaching learning algorithms to retrospectively collected data on trauma and emergency general surgery patients, utilizing large databases from the National Trauma Data Bank and Trauma Quality Improvement Program, amongst smaller institutional databases. Methodologies vary from random-forest, logistic regression, and backwards elimination. At present, we have robust models to help predict mortality,^{5,6} venous thromboembolism, and sepsis.^{1,7} The data, however, is most useful when it

is delivered in real-time and reduces the "noise" and time required for manual calculations. One example of prospective use of machine learning algorithms in patient care can be seen with algorithms utilized for predicting a patient's likelihood of developing sepsis.⁸

In the trauma realm, there may be a decision point between definitive management and damage control for a patient's injuries. The indications for when damage control tactics might be utilized are still debated.^{9,10} When to operate for hemorrhage control is less controversial; however, once damage control surgery is done, when is it safe to proceed with other clinically significant injuries (eg, open femur fracture and pelvic fracture) that if not fixed or washed out in a timely fashion can result in infection, non-union, or pain? For example, some patients in this situation, and also at end-of-life, are either too high-risk or may not want to pursue invasive procedures given the risk over benefit. Several calculators exist to describe perioperative risk in surgical patients, such as the American College of Surgeons' National Surgical Quality Improvement Program or the non-linear POTTER surgical-risk calculator.^{2,11–13} However, these are lengthy, may require manual input, or are retrospective, only describing after the care of the patient that they were likely to die or have complications. To our knowledge, there is no live, fully automated, and integrated machine learning algorithm that prospectively estimates mortality risk in trauma patients.

Many existing trauma mortality scoring systems or risk assessments are retrospective, or may require lengthy calculations that normally do not occur in real-

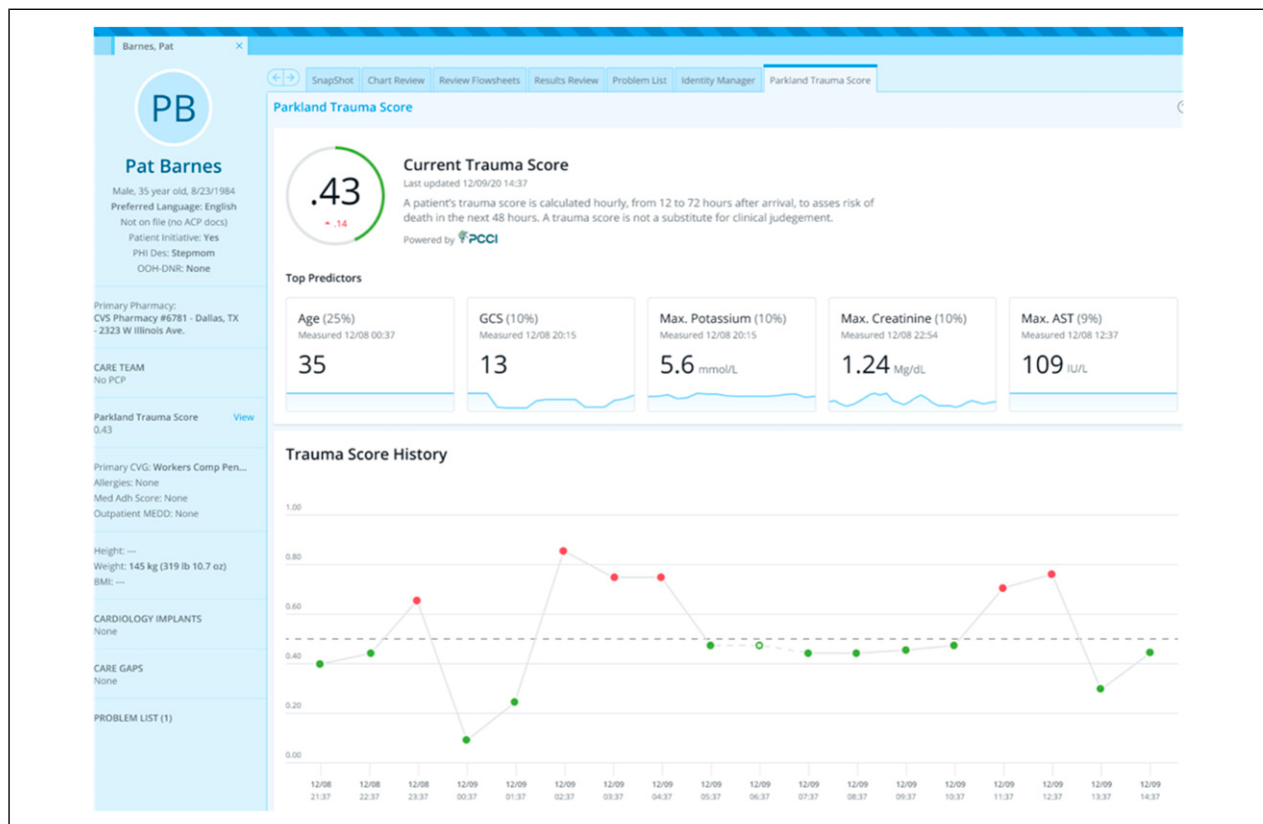


Figure 3. Sample patient dashboard, snapshot of PTIM metric, and top 5 predictors with hourly calculations (age, GCS, maximum potassium, maximum creatinine, and maximum AST).

time and are calculated retrospectively. Scoring systems including the TRISS,¹⁴ Trauma Score,¹⁵ the Injury Severity Score, Abbreviated Injury Scales, and ASCOT¹⁶ are critical in assessing severity of disease but may require lengthy calculations, are external to the EHR, and may not be as helpful with real-time evaluation. In trauma and acute care, there is the historically described “golden hour” which encompasses the time prior to arrival and initial presentation to a trauma center where the focus is to resuscitate, perform life-saving maneuvers, and stabilize the patient.¹⁷ The PTIM score informs clinicians about the patient’s mortality risk over the next 48 hours, by synthesizing multiple physiologic and laboratory data every hour. Additionally, because the algorithm utilizes objective data gathered from the EMR, it reduces the amount of human or user error that occurs with manual input (tracking utilization of massive transfusion and diagnosis of acute respiratory distress syndrome). This allows the PTIM algorithm to be utilized with minimal effort by clinicians, allowing for more time spent planning and delivering patient care.

The PTIM algorithm can be distinguished from a more common “risk calculator,” which can be based on traditional statistical methods or machine learning approaches.

The distinction lies in the method used to derive the predictions or calculations.

A traditional risk calculator is typically developed using statistical methods, often based on regression models and the underlying model is usually simpler and interpretable. For example, the Framingham Risk Score for cardiovascular disease is based on a regression model using factors like age, cholesterol levels, and blood pressure. In the field of trauma, the ISS, RTS, and TRISS are examples of traditional risk and/or severity calculators.

A machine learning-based risk calculator utilizes machine learning algorithms to predict risk. These algorithms can range from simple linear models to more complex structures like neural networks, decision trees, or ensemble methods. Further, these models “learn” from data. Given a sufficiently large dataset, they capture intricate patterns or interactions between variables that might be harder to model with traditional statistical methods. The model can be updated or retrained as new data becomes available.

The algorithm for PTIM was developed utilizing a balanced bagging ensemble of decision tree classifiers on a large dataset from a level 1 trauma center. Thus, it is a machine learning model. Further, the model implementation methodology allows for model updates and

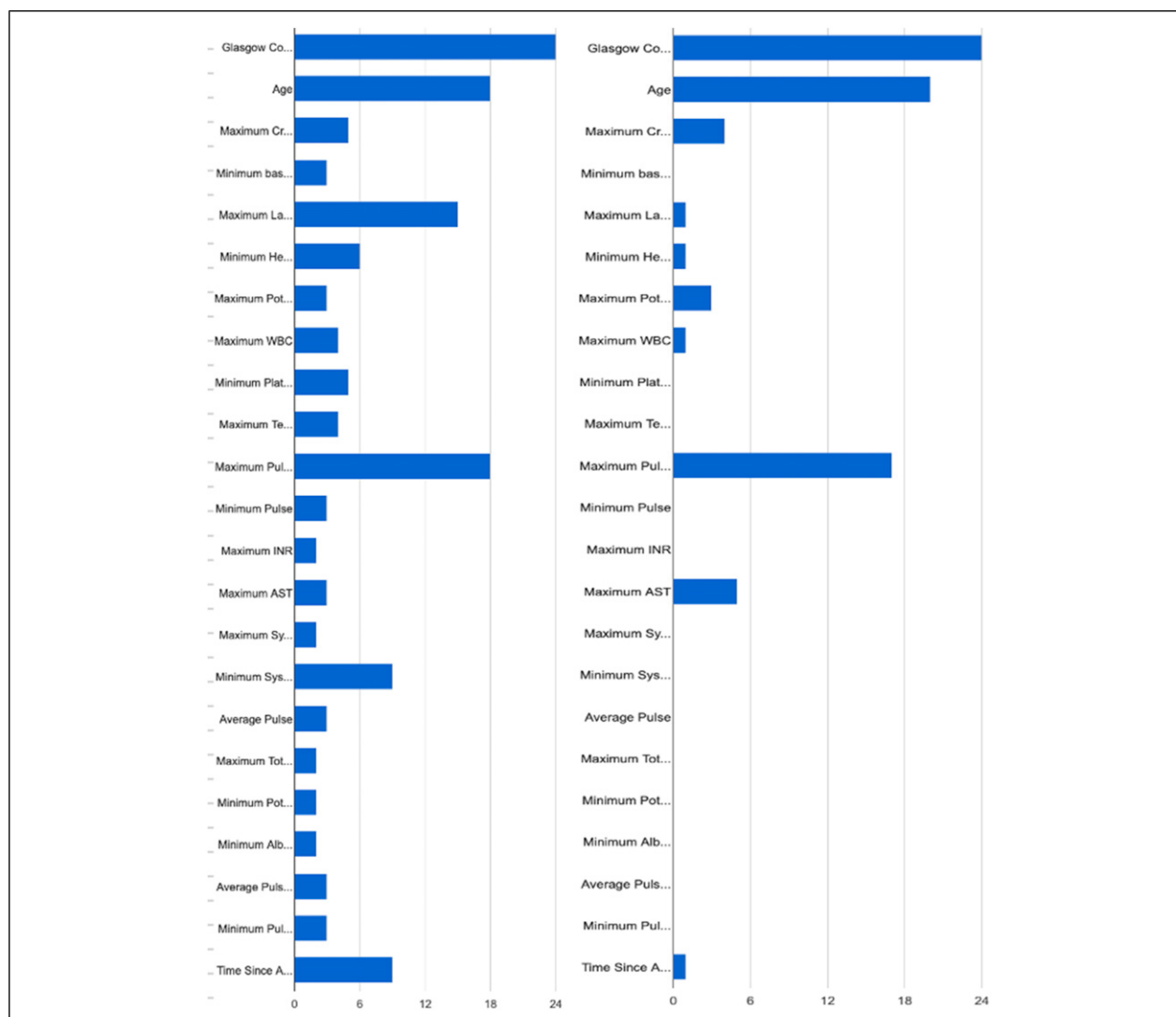


Figure 4. a, b proposed top 3 predictors of mortality (Glasgow Coma Score, age, and maximum pulse) vs actual top 3 predictors of mortality (Glasgow Coma Score, age, and maximum pulse).

retraining if the model performance falls out of acceptable bounds. Further details about the machine learning algorithm driving the risk calculator are available to the public.¹⁸

Our study is not without limitations. Our sample size was small, and as such, may not represent all the clinicians that rotate through or work in our ICU. Additionally, our study is likely affected by selection bias in the responses of the clinicians that completed the survey. Some of the surveys were completed by the same clinician, as clinicians could complete a survey each time they utilized the

PTIM tool during their care of a patient. Thus, some responses may be from the same clinician, leading to possible selection bias in our results. Lastly, the PTIM score was likely to play a more impactful role in the care of complex patients, which also contributes to a selection bias.

Despite these limitations, the PTIM tool was useful in the majority of instances it was used. The tool is already integrated into the EHR and because it is updated automatically, requires no input from clinicians on their part save for noting the hourly score. Machine learning

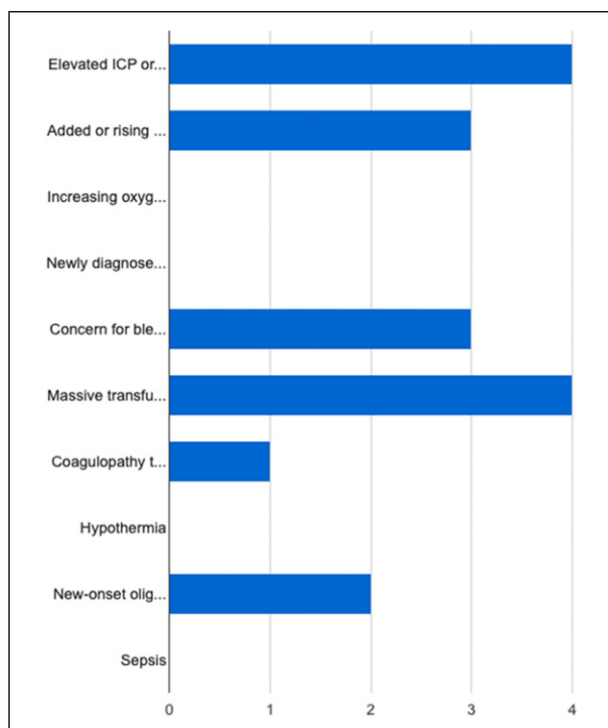


Figure 5. Top variables not accounted for by the model and ranked high by clinicians (elevated ICP, massive transfusion, concern for bleeding, rising lactate, coagulopathy, and new-onset oliguria).

algorithms used for predicting patient mortality or outcomes prospectively have been developed in other areas of medicine, such as acute kidney injury prediction in pediatric and adult populations, or neurosurgical interventions.¹⁹⁻²¹

Conclusion

In this study, we prospectively surveyed clinicians on the use of a real-time, hourly, automated machine learning algorithm embedded into the electronic healthrecord for more effective triage, management, and mortality risk assessment of trauma patients. We found concordance between perceived predictors and actual predictors of mortality scored by PTIM. Providers found it easy to use and assisted in treatment and prediction of mortality. We hope to address clinician barriers to utilizing the PTIM tool, and standardizing how it is incorporated into patient care—where to look for the tool in notes and interpretation of the score across all specialties. To date, the tool has been used mostly for communication between the SICU and orthopedic surgery services and establish “readiness” for the operating room. Lastly, we hope to answer the critical question—how does this help our patients? Can we improve our resuscitation and utilization of resources and decrease ICU length of stay, complications, and

mortality? The answer may lie in more data and through challenging the algorithm with more diverse patients through our current multi-center work.

Author’s Note

The preliminary findings of this paper were presented at the American College of Surgeons Clinical Congress, October, 2022 in San Diego, CA.

Authors’ Contributions

Study Design: CP, DS, VC, and AS.

Data extraction and analysis: SLA, JH, CH, and LAD.

Writing and critical revisions: CP, SLA, JH, RPS, DS, VC, AS, and MC.

Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: State: - Advisor to TraumaCare.AI, machine-learning entity. This relationship occurred after this research was performed.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Disclosures

CP and AS are advisors to TraumaCare.AI.

ORCID iD

Drew Sanders  <https://orcid.org/0000-0001-7186-5915>

References

1. Nederpelt CJ, Mokhtari AK, Alser O, et al. Development of a field artificial intelligence triage tool: Confidence in the prediction of shock, transfusion, and definitive surgical therapy in patients with truncal gunshot wounds. *J Trauma Acute Care Surg.* 2021;90(6):1054-1060. doi:10.1097/TA.0000000000003155
2. Bertsimas D, Dunn J, Velmahos GC, Kaafarani HMA. Surgical risk is not linear: Derivation and validation of a novel, user-friendly, and machine-learning-based Predictive Optimal Trees in Emergency Surgery Risk (POTTER) calculator. *Ann Surg.* 2018;268(4):574-583. doi:10.1097/SLA.0000000000002956
3. Benz D, Balogh ZJ. Damage control surgery: Current state and future directions. *Curr Opin Crit Care.* 2017;23(6):491-497. https://journals.lww.com/co-criticalcare/Fulltext/2017/12000/Damage_control_surgery_current_state_and_future.8.aspx
4. Starr AJ, Julka M, Nethi A, et al. 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](https://doi.org/10.1097/BOT.0000000000002290)
5. He L, Luo L, Hou X, et al. Predicting venous thromboembolism in hospitalized trauma patients: a combination of the Caprini score and data-driven machine learning model. *BMC Emerg Med*. 2021;21(1):60. PMID: 33971809; PMCID: PMC8111727. doi:[10.1186/s12873-021-00447-x](https://doi.org/10.1186/s12873-021-00447-x)
 6. Lang EW, Pitts LH, Damron SL, Rutledge R. Outcome after severe head injury: An analysis of prediction based upon comparison of neural network versus logistic regression analysis. *Neurol Res*. 1997;19(3):274-280. PMID: 9192380. doi:[10.1080/01616412.1997.11740813](https://doi.org/10.1080/01616412.1997.11740813)
 7. Maurer LR, Bertsimas D, Bouardi HT, et al. Trauma outcome predictor: An artificial intelligence interactive smartphone tool to predict outcomes in trauma patients. *J Trauma Acute Care Surg*. 2021;91(1):93-99. PMID: 33755641. doi:[10.1097/TA.0000000000003158](https://doi.org/10.1097/TA.0000000000003158)
 8. Roberts DJ, Bobrovitz N, Zygun DA, Kirkpatrick AW, Faris PD, Stelfox HT. Indications for use of damage control surgery and damage control interventions in civilian trauma patients: A scoping review. *J Trauma Acute Care Surg*. 2015;78(6):1196. https://journals.lww.com/jtrauma/Fulltext/2015/06000/Indications_for_use_of_damage_control_surgery_and.19.aspx
 9. Ben-Israel D, Jacobs WB, Casha S, et al. The impact of machine learning on patient care: A systematic review. *Artif Intell Med*. 2020;103:101785. Epub 2019 Dec 31. PMID: 32143792. doi:[10.1016/j.artmed.2019.101785](https://doi.org/10.1016/j.artmed.2019.101785)
 10. Sinha M, Jupe J, Mack H, Coleman TP, Lawrence SM, Fraley SI. Emerging technologies for molecular diagnosis of sepsis. *Clin Microbiol Rev*. 2018;31(2):e00089-e000817. PMID: 29490932; PMCID: PMC5967692. doi:[10.1128/CMR.00089-17](https://doi.org/10.1128/CMR.00089-17)
 11. Ko C. ACS NSQIP conference and semiannual report overview. Presentation at the 2009 ACS NSQIP National Conference. 2009.
 12. Steinberg S, Popa M, Michalek JA, Bethel MJ, Ellison EC. Comparison of risk-adjustment methodologies in surgical quality improvement. *Surgery*. 2008;144(4):662-667; discussion 662-7.
 13. Wick EC, Shore AD, Hirose K, et al. Readmission rates and cost following colorectal surgery. *Dis Colon Rectum*. 2011; 54(12):1475-1479.
 14. Boyd CR, Tolson MA, Copes WS. Evaluating trauma care: The TRISS method. Trauma score and the injury severity score. *J Trauma*. 1987;27(4):370-378.
 15. Champion HR, Sacco WJ, Carnazzo AJ, Copes W, Fouty WJ. Trauma score. *Critical Care Medicine: September*. 1981;9:672-676.
 16. Champion HR, Copes WS, Sacco WJ, et al. A new characterization of injury severity. *J Trauma*. 1990;30(5): 539-545. PMID: 2342136. doi:[10.1097/00005373-199005000-00003](https://doi.org/10.1097/00005373-199005000-00003)
 17. Clark DE. R A Cowley, the “golden hour,” the “momentary pause,” and the “third space”. *Am Surg*. 2017;83(12): 1401-1406. PMID: 29336762.
 18. Nethi A. Parkland Trauma Index of Mortality (PTIM): Real-time predictive model for trauma patients. Quantitative methods. Cornell University. <https://arxiv.org/abs/2010.03642>. Accessed September 12, 2023.
 19. Dong J, Feng T, Thapa-Chhetry B, et al. Machine learning model for early prediction of Acute Kidney Injury (AKI) in pediatric critical care. *Crit Care*. 2021;25(1):288. PMID: 34376222. PMCID: PMC8353807. doi:[10.1186/s13054-021-03724-0](https://doi.org/10.1186/s13054-021-03724-0)
 20. Zhang Z, Ho KM, Hong Y. Machine learning for the prediction of volume responsiveness in patients with oliguric acute kidney injury in critical care. *Crit Care*. 2019; 23(1):112. PMID: 30961662. PMCID: PMC6454725. doi:[10.1186/s13054-019-2411-z](https://doi.org/10.1186/s13054-019-2411-z)
 21. Celtikci E. A systematic review on machine learning in neurosurgery: The future of decision-making in patient care. *Turk Neurosurg*. 2018;28(2):167-173. PMID: 28481395. doi:[10.5137/1019-5149.JTN.20059-17.1](https://doi.org/10.5137/1019-5149.JTN.20059-17.1)