Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Validation of the artificial intelligence–based trauma outcomes predictor (TOP) in patients 65 years and older

Majed El Hechi, MDa,b, Anthony Gebran, MDa,b, Hamza Tazi Bouardi, MSc, Lydia R. Maurer, MDa,b, Mohamad El Moheb, MDa,b, Daisy Zhuo, PhDd, Jack Dunn, PhDb, Dimitris Bertsimas, PhDb,d, George C. Velmahos, MD, PhDa, Haytham M.A. Kaafarani, MD, MPHa,b,*

a Division of Trauma, Emergency Surgery, and Surgical Critical Care, Massachusetts General Hospital, Boston, MA
b Center for Outcomes and Patient Safety in Surgery, Massachusetts General Hospital, Boston, MA
c Massachusetts Institute of Technology, Cambridge, MA
d Interpretable AI, Boston, MA

Article info
Article history:
Accepted 16 November 2021
Available online 23 December 2021

ABSTRACT

Background: The Trauma Outcomes Predictor tool was recently derived using a machine learning methodology called optimal classification trees and validated for prediction of outcomes in trauma patients. The Trauma Outcomes Predictor is available as an interactive smartphone application. In this study, we sought to assess the performance of the Trauma Outcomes Predictor in the elderly trauma patient.

Methods: All patients aged 65 years and older in the American College of Surgeons Trauma Quality Improvement Program 2017 database were included. The performance of the Trauma Outcomes Predictor in predicting in-hospital mortality and combined and specific morbidity based on incidence of 9 specific in-hospital complications was assessed using the c-statistic methodology, with planned subanalyses for patients 65 to 74, 75 to 84, and 85+ years.

Results: A total of 260,505 patients were included. Median age was 77 (71–84) years, 57% were women, and 98.8% had a blunt mechanism of injury. The Trauma Outcomes Predictor accurately predicted mortality in all patients, with excellent performance for penetrating trauma (c-statistic: 0.92) and good performance for blunt trauma (c-statistic: 0.83). Its best performance was in patients 65 to 74 years (c-statistic: blunt 0.86, penetrating 0.93). Among blunt trauma patients, the Trauma Outcomes Predictor had the best discrimination for predicting acute respiratory distress syndrome (c-statistic 0.75) and cardiac arrest requiring cardiopulmonary resuscitation (c-statistic 0.75). Among penetrating trauma patients, the Trauma Outcomes Predictor had the best discrimination for deep and organ space surgical site infections (c-statistics 0.95 and 0.84, respectively).

Conclusion: The Trauma Outcomes Predictor is a novel, interpretable, and highly accurate predictor of in-hospital mortality in the elderly trauma patient up to age 85 years. The Trauma Outcomes Predictor could prove useful for bedside counseling of elderly patients and their families and for benchmarking the quality of geriatric trauma care.

Introduction

The Census Bureau estimates that 49 million individuals in the United States were 65 years and older in 2016. By 2034, this number is projected to rise to 77 million, outnumbering the 76.5 million individuals under the age of 18.1 As the US population continues to age, the number of traumatic injuries in the elderly patient population will inevitably continue to increase. Decreased physiologic reserve, frailty, malnutrition, and frequent comorbidities all contribute to poor outcomes in this patient cohort.2,3 In turn, the high rate of morbidity and mortality in the elderly patient
often translates to longer hospital stays and contributes to increased health care utilization.\(^4\)

An important aspect of the delivery of care to the injured geriatric patient and family includes an early discussion regarding the goals of care and treatment plan.\(^5\) To guide discussions and facilitate decision making, accurate knowledge about the patient’s prognosis is crucial. Frequently, discussions are guided by the personal experience of the treating physician. At other times, the surgeon may choose to use validated prognostic calculators after trauma, such as the Trauma Injury Severity Score (TRISS)\(^6\) or the Geriatric Trauma Outcome Score (GTOS).\(^7,8\)

Although these models have predictive power, they share a common methodology whereby every variable in the model contributes the same amount of risk toward the outcome, regardless of other variables in the model. This makes the predictive model linear and thus insensitive to the clinical context to which it is applied. However, the clinical reality is that the interaction among patient demographics, comorbidities, vital signs, and the severity of injury at presentation in determining patient outcome is almost never linear or additive.\(^9\)

To illustrate how the interaction between demographics can be nonlinear, consider a 70-year-old man who fell from standing. The severity of injury at presentation may determine what factors influence his risk of mortality. For instance, if the patient presents alert and cooperative, certain comorbidities such as cirrhosis or heart failure might play significantly into the probability of surviving his hospital stay. Whereas when the injury is more severe and the patient presents with a depressed Glasgow Coma Scale (GCS), the extent of injury becomes a major determinant in survival, and the relative importance of comorbidities wanes.

Novel machine learning (ML) technology has the ability to generate models that capture nonlinear patterns in data. Our group has created the Trauma Outcomes Predictor (TOP),\(^10\) a smartphone-based risk prediction tool for trauma patients that uses ML optimal classification trees (OCTs), and validated the tool for predicting outcomes among all trauma patients. Although the accuracy of TOP in predicting hospital mortality and 9 in-hospital complications has been validated across all trauma patients, it has not been examined specifically in the geriatric trauma population, in which prognostic discussions are of particular importance but risk prediction is historically more challenging. In this study, we aimed to assess the performance of TOP in trauma patients aged 65 years and older.

**Methods**

**Database and patient population**

The American College of Surgeons Trauma Quality and Improvement Program (ACS-TQIP) database year 2017 was selected for this validation study. With contributions from more than 800 level I or II ACS-verified or state-designated trauma center centers across the United States, the database consists of more than 100 data items, including patient demographics, comorbidities, type and mechanism of injury, Injury Severity Score (ISS), Abbreviated Injury Scale (AIS), prehospital and emergency department (ED) vital signs, diagnoses, complications, and mortality. All patients older than 65 years included in the 2017 ACS-TQIP database were included. Three age subgroups were also created: 65–74, 75–84, and ≥85 years old. This study was approved by the Mass General Brigham Institutional Review Board.

![The optimal classification tree used by Trauma Outcome Predictor (TOP) to predict inpatient mortality, with a magnification on one of the terminal nodes.](image)

**The OCT methodology and the TOP smartphone application**

The OCT methodology used for the development of the TOP calculator has been previously described in detail.\(^10\)–\(^12\) Briefly, OCTs are a subset of ML methodologies that examine all input variables and select the most important ones for each tree to optimize accuracy. To do this, OCTs reboot themselves with each set of variables to try all conceivable combinations and determine the most impactful ones. In the process, OCTs establish specific cutoffs for each node based on knowledge about downstream and upstream impacts, as opposed to top-down approaches used by other tree-based models such as classification and regression trees. A unique OCT is generated for every outcome of interest. The structure of the mortality prediction OCT with a magnification of one of its terminal
nodes is presented in Figure 1. The example terminal node shows a split on the age of 42. For patients older than 42 years, the next question is whether the patient has liver cirrhosis. If so, the predicted mortality is 9.76%, compared with a patient who does not have cirrhosis, for whom there is an additional question about GCS on presentation. For patients with a GCS greater than 13, the predicted mortality is 1.26%, and for those less than 13, the predicted mortality is 2.82%. The splits on each of these nodes, in addition to the nodes themselves, are not determined by the developers but instead generated by the OCT algorithm in a way that optimizes overall performance of the model. An additional advantage of OCTs is their interpretability, unlike other “black-box” ML models used in risk prediction that risk introducing existing bias and disparities into the models. With OCTs, the clinicians are able to easily follow the algorithms’ reasoning and understand how the final predictions were made.

The user-friendly TOP smartphone application that is built on the OCTs is presented in Figure 2. This example demonstrates how the questions generated by the application differ depending on an individual patient’s characteristics. For instance, the patient on the left has a GCS of 15 on presentation, whereas the patient on the right has a GCS of 10, and the subsequent questions differ to ultimately provide individual mortality estimates for each patient.
Primary and secondary outcomes

The primary outcome was the accuracy of prediction of in-hospital mortality for blunt and penetrating trauma patients. Secondary outcomes included the prediction of combined morbidity (i.e., the occurrence of any of 9 in-hospital complications) as well as individual complications across both mechanisms of injury. Individual complications included acute respiratory distress syndrome, cardiac arrest requiring cardiopulmonary resuscitation, organ space surgical site infections, deep surgical site infections, unplanned intubation, severe sepsis, acute kidney injury, deep venous thrombosis, and pulmonary embolism.

Measuring the performance of TOP

The area under the curve (AUC) metric, or c-statistic, was used to measure the ability of algorithms to predict in-hospital mortality.
Results

A total of 260,505 patients were included, of which 39%, 40%, and 21% were in the 65–74, 75–84, and ≥85 years age groups, respectively. The median patient age was 77 years (71–84), 56.7% were females, and hypertension (64.5%), diabetes mellitus (25.8%), and chronic obstructive pulmonary disease (13.0%) were the most common comorbid conditions (Table I). Most injuries resulted from falls (80.6%) and motor vehicle collisions (12.4%). The median ISS was 9 (5–10), and a severe head injury (AIS >3) was present in 18.2% of patients.

Most patients were admitted to the floor (54.7%) or ICU (23.4%) from the ED (Table II).

The in-hospital mortality rate was 4.2%, and 3.2% of patients developed 1 or more complications. The most common in-hospital complications were unplanned intubation (1.3%), acute kidney injury (0.7%), deep venous thrombosis (0.5%), and cardiac arrest requiring CPR (0.7%). The most common discharge dispositions were rehabilitation or long-term care (53.2%) and home (43.2%).

Table III

| Patient characteristics | Patients ≥65 | Age 65–74 | Age 75–84 | Age ≥85 | P value |
|-------------------------|--------------|-----------|-----------|--------|---------|
| Hospital mortality      | 10,840 (4.2%)| 3,426 (3.4%)| 4,729 (4.5%)| 2,685 (4.9%)| <.001   |
| Combined morbidity      | 8,275 (3.2%) | 3,611 (3.6%)| 3,266 (3.1%)| 1,398 (2.6%)| <.001   |
| Acute kidney injury     | 1,759 (0.7%) | 678 (0.7%)  | 745 (0.7%) | 336 (0.6%) | 0.97    |
| ARDS                    | 653 (0.3%)   | 345 (0.3%)  | 231 (0.2%) | 77 (0.2%)  | <.001   |
| Cardiac arrest requiring CPR | 1,704 (0.7%) | 709 (0.7%)  | 689 (0.7%) | 306 (0.6%) | 0.06    |
| Deep SSI                | 93 (<0.1%)   | 47 (<0.1%)  | 33 (<0.1%) | 13 (<0.1%) | 0.53    |
| Deep venous thrombosis  | 1,349 (0.5%) | 643 (0.6%)  | 506 (0.5%) | 200 (0.4%) | <.001   |
| Organ space SSI         | 53 (<0.1%)   | 33 (<0.1%)  | 17 (<0.1%) | 3 (<0.1%)  | <.001   |
| Pulmonary embolism      | 654 (0.3%)   | 314 (0.3%)  | 235 (0.2%) | 105 (0.2%) | <.001   |
| Unplanned intubation    | 3,469 (1.3%) | 1,547 (1.5%)| 1,387 (1.3%)| 535 (1.0%) | <.001   |
| Severe sepsis           | 950 (0.4%)   | 434 (0.4%)  | 361 (0.3%) | 155 (0.3%) | <.001   |
| Hospital discharge disposition | 101,324 (43.2%) | 52,823 (54.4%)| 38,726 (39.1%)| 15,274 (29.9%)| <.001   |
| Home (with or without services) | 131,549 (53.2%) | 41,172 (42.4%)| 56,809 (57.3%)| 33,568 (65.7%)| <.001   |
| Rehab or long-term care | 8,874 (3.6%) | 3,043 (3.1%)| 3,575 (3.6%)| 2,256 (4.4%)| <.001   |
| Other                   | 54,362 (21%) | 104,819 (39.1%)| 101,324 (43.2%)| 131,549 (53.2%)| <.001   |

ARDS, acute respiratory distress syndrome; CPR, cardiopulmonary resuscitation; SSI, surgical site infection.

Figure 3. Trauma Outcome Predictor (TOP) model area under the receiver operator characteristic (ROC) curve; for (A) blunt injury mortality and (B) penetrating injury mortality.
hospital complications were unplanned intubations (1.3%), acute kidney injuries (0.7%), and cardiac arrest (0.7%) (Table III).

Prediction of in-hospital mortality

Blunt trauma patients

The c-statistic for predicting in-hospital mortality in blunt trauma patients ≥65 was 0.83. On subgroup analysis, the performance of TOP was the highest among patients aged 65 to 74 (c-statistic: 0.86) and lowest among patients ≥85 (c-statistic: 0.77) (Figure 3, A).

Penetrating trauma patients

The c-statistic for predicting in-hospital mortality in penetrating trauma patients ≥65 was 0.92. On subgroup analysis, the performance of TOP was the highest among patients aged 65 to 74 (c-statistic: 0.93) and lowest among patients 75 to 84 (c-statistic: 0.88) (Figure 3, B).

Prediction of in-hospital morbidity

Blunt trauma patients

The c-statistic for predicting in-hospital morbidity in blunt trauma patients ≥65 was 0.69. On subgroup analysis, the performance of TOP was the highest among patients aged 65 to 74 (c-statistic: 0.73) and lowest among patients ≥85 (c-statistic: 0.63) (Figure 4, A).

Penetrating trauma patients

The c-statistic for predicting in-hospital morbidity in penetrating trauma patients ≥65 was 0.71. On subgroup analysis, the performance of TOP was the highest among patients aged 65 to 74 (c-statistic: 0.72) and lowest among patients ≥85 (c-statistic: 0.68) (Figure 4, B).

Prediction of individual complications

Figure 5 is a radar plot that shows the performance of TOP in predicting individual complications. Among patients with blunt trauma, TOP best predicted acute respiratory distress syndrome (c-statistic: 0.75) and cardiac arrest (c-statistic: 0.75). Among patients with penetrating trauma, TOP best predicted organ space (c-statistic: 0.95) and deep surgical site infections (c-statistic: 0.84).

Discussion

In this study, we examine the ability of TOP to predict in-hospital outcomes in the vulnerable geriatric trauma patient. We found that TOP predicts in-hospital mortality very well in blunt and penetrating trauma patients older than 65 years of age. In addition, TOP has moderate performance in predicting in-hospital morbidity among patients 65 to 74 years old, but its performance drops in older patients, especially those older than 85 years. Despite having good to excellent performance among certain geriatric patients, TOP performed better in the general cohort of trauma patients, when compared to geriatric trauma patients as evidenced by the respective AUCs.
Prognostication of outcome and handling family discussions in the injured geriatric patient is never an easy task. In addition to the physical components of injury, the physician has to take into account the cumulative deficit burden model, which consists of preinjury physical, social, functional, and cognitive factors when counseling family. To add to the challenge, it has been shown that some families overestimate their loved one’s odds of recovery and are reluctant to accept poor prognoses, whereas others tend to underestimate them and are more inclined to withdraw care. Consequently, surrogate decision makers tend to pursue aggressive treatment options early in the hospitalization course, but often prefer palliative care once the poor prognosis is apparent. This translates into prolonged intensive care unit stays and numerous invasive interventions for the severely injured geriatric patient, which unfortunately do not translate to the survival of hospital stay. Additional evidence suggests that patients typically underestimate their remaining life expectancy, whereas physicians have been found to overestimate patients’ life expectancy. These data reinforce the accuracy of TOP in predicting in-hospital mortality and major complications in the geriatric trauma patient. Compared to relying on personal experience or anecdotal data alone, TOP can substantially help surgeons achieve objective and rapid assessments of risk and set realistic patient and family expectations early in the hospital course. This is mostly applicable to geriatric patients younger than 85 years, where TOP proved to have the best performance.

Two other notable risk calculators that have found application in the prognostication of geriatric mortality after trauma are TRISS and the GTOS. The TRISS method employs a logistic regression model that includes the Injury Severity Score and Revised Trauma Score variables to generate a mortality estimate for blunt and penetrating injuries. The discrimination of TRISS for prediction of mortality as determined by the area under the curve ranges from 0.69 to 0.90 in geriatric patients. The greatest advantage of TRISS is its durability, where it remains the most recognized and accepted mortality estimator in trauma patients since its creation more than 30 years ago. Although the Dutch Trauma Registry has recently created a modified version of TRISS to extend its use to the elderly population, it has been shown that TRISS has high misclassification rates in patients older than 54 years, which can be attributed to the dichotomization of age at 55 years. TOP, on the other hand, takes into account the nuances of advanced age, where the trees allow splits at various age cutoffs at different levels of the tree. Furthermore, TRISS takes first recorded physiologic parameters, which is problematic because it relies on the effort of emergency medical personnel to capture multiple variables, which might not always be available and might not be missing-at-random in the severely injured patients. TOP, meanwhile, available as a smartphone application, can be applied for prognostication as soon as the patient is admitted from the emergency department. Moreover, it is important to note that although the 2 above calculators provide mortality estimates to geriatric trauma patients, they are not capable of predicting the occurrence of a complication during hospitalization, an ability that is
possessed by TOP. Additionally, TOP has the potential to be used as a tool for benchmarking the quality of care among geriatric trauma patients, an area that is being currently explored by the authors.

Our study has a number of limitations. First, patients with less severe injuries are not included in the ACS-TQIP database. Second, an AI or ML methodology only performs as well as the data it trains on. For the elderly trauma patient, especially those older than 85, the data points currently collected by the ACS-TQIP are not sufficient. Additional information on patients’ cognitive levels, for example, might considerably improve predictions in the geriatric population. Third, the aggregate of all patients in the United States, on which the model is trained, might not be an accurate representation of any specific patient demographic at a local hospital level. Fourth, TOP focuses solely on in-hospital mortality and morbidity risk, which does not provide a complete prognostic picture, since severe traumatic injuries are known to have an increased risk of mortality beyond the index admission.25

In conclusion, TOP is an interpretable and accurate calculator used to predict in-hospital mortality and complications in the geriatric trauma patient up to the age 85 years. TOP could prove useful for bedside counseling of elderly patients and their families regarding the risk of mortality and could guide important goals of care discussions in this population.

Funding/Support

This study was supported by a grant from CRICO/Risk Management Foundation of the Harvard Medical Institutions.

Conflicts of interest/Disclosure

Drs Zhou, Dunn, and Bertsimas are founders of Interpretable AI.

References

1. United States Census Bureau. Older people projected to outnumber children for first time in U.S. history. https://www.census.gov/newsroom/press-releases/2018/cb18-41-population-projections.html. Accessed October 16, 2021.
2. Garwe T, Albrecht RM, Stoner JA, Mitchell S, Morghare P. Hypoalbuminemia at admission is associated with increased incidence of in-hospital complications in geriatric trauma patients. Am J Surg. 2016;212:109–115.
3. Bhattacharya B, Maung A, Schuster K, Davis KA. The older they are the harder they fall: injury patterns and outcomes by age after ground level falls. Injury. 2016;47:1955–1959.
4. Taylor MD, Tracy JK, Meyer W, Pasquale M, Napolitano LM. Trauma in the elderly: intensive care unit resource use and outcome. J Trauma Acute Care Surg. 2002;53:407–414.
5. Madni TD, Ekeh AP, Brakenridge SC, et al. A comparison of prognosis calculators for geriatric trauma: a Prognostic Assessment of Life and Limitations After Trauma in the Elderly Consortium study. J Trauma Acute Care Surg. 2017;83:90–96.
6. Boyd CR, Mann Tolson, Copes WS. Evaluating trauma care: the TRISS method. J Trauma Acute Care Surg. 1987;27:370–378.
7. Cook AC, Joseph B, Inaba K, et al. Multicenter external validation of the geriatric trauma outcome score: a study by the Prognostic Assessment of Life and Limitations After Trauma in the Elderly (PALLIATE) Consortium. J Trauma Acute Care Surg. 2016;80:204–209.
8. Zhao FZ, Wolf SE, Nakonezny PA, et al. Estimating geriatric mortality after injury using age, injury severity, and performance of a transfusion: the geriatric trauma outcome score. J Palliat Med. 2015;18:677–681.
9. Chen JH, Asch SM. Machine learning and prediction in medicine: beyond the peak of inflated expectations. N Engl J Med. 2017;376:2507–2509.
10. 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;91:93–99.
11. 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;268:574–583.
12. Bertsimas D, Dunn J. Optimal classification trees. Mach Learn. 2017;106:1039–1082.
13. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 1982;143:29–36.
14. Cesari M, Araujo de Carvalho I, Amuthavalli Thiagarajan J, et al. Evidence for the domains supporting the construct of intrinsic capacity. J Gerontol A. 2018;73:1653–1660.
15. Mitrakos AB, Mogilner AJ, Rockwood K. Accumulation of deficits as a proxy measure of aging. Sci World J. 2001;1:323–336.
16. You JJ, Downar J, Fowler RA, et al. Barriers to goals of care discussions with seriously ill hospitalized patients and their families: a multicenter survey of clinicians. JAMA Intern Med. 2015;175:549–556.
17. O’Connell K, Maier R. Palliative care in the trauma ICU: current opinion in critical care. https://journals.lww.com/co-criticalcare/Fulltext/2016/12000/Palliative_care_in_the_trauma_ICU.13.aspx. Accessed October 19, 2021.
18. Wright A, Schurr M. Geriatric trauma: review and recommendations. JAMA Intern Med. 2015;175:1986–1988.
19. Stiel S, Bertram L, Neuhaus S, et al. Barriers to goals of care discussions with seriously ill hospitalized patients and their families: a multicenter survey of clinicians. JAMA Intern Med. 2015;175:549–556.
20. O’Connell K, Maier R. Palliative care in the trauma ICU: current opinion in critical care. https://journals.lww.com/co-criticalcare/Fulltext/2016/12000/Palliative_care_in_the_trauma_ICU.13.aspx. Accessed October 19, 2021.
21. Wright A, Schurr M. Geriatric trauma: review and recommendations. WMJ. 2001;100:57–59.
22. Champion HR, Sacco WJ, Copes WS, Gann DS, Gennarelli TA, Flanagan ME. The trauma score. J Trauma. 1989;29:384.
23. de Munter L, ter Jongh MAC. Improvement of the performance of survival prediction in the ageing blunt trauma population: a cohort study. PLoS One. 2018;13. e0209099.
24. Demetriades D, Chan LS, Velmahos G, et al. TRISS methodology in trauma: the need for alternatives. Br J Surg. 1998;85:379–384.
25. Center for Health Statistics. With special feature on racial and ethnic health disparities. https://www.cdc.gov/nchs/data/hus/hus15.pdf. Accessed October 16, 2021.