Motor vehicle collisions (MVCs) account for 1.35 million deaths and cost $518 billion US dollars each year worldwide, disproportionately affecting young patients and low-income nations. The ability to successfully anticipate clinical outcomes will help physicians form effective management strategies and counsel families with greater accuracy. The authors aimed to train several classifiers, including a neural network model, to accurately predict MVC outcomes.

**METHODS** A prospectively maintained database at a single institution’s level I trauma center was queried to identify all patients involved in MVCs over a 20-year period, generating a final study sample of 16,287 patients from 1998 to 2017. Patients were categorized by in-hospital mortality (during admission) and length of stay (LOS), if admitted. All models included age (years), Glasgow Coma Scale (GCS) score, and Injury Severity Score (ISS). The in-hospital mortality and hospital LOS models further included time to admission.

**RESULTS** After comparing a variety of machine learning classifiers, a neural network most effectively predicted the target features. In isolated testing phases, the neural network models returned reliable, highly accurate predictions: the in-hospital mortality model performed with 92% sensitivity, 90% specificity, and a 0.98 area under the receiver operating characteristic curve (AUROC), and the LOS model performed with 2.23 days mean absolute error after optimization.

**CONCLUSIONS** The neural network models in this study predicted mortality and hospital LOS with high accuracy from the relatively few clinical variables available in real time. Multicenter prospective validation is ultimately required to assess the generalizability of these findings. These next steps are currently in preparation.

https://thejns.org/doi/abs/10.3171/2022.1.FOCUS21739

**KEYWORDS** length of stay; machine learning; motor vehicle collision; passenger mortality

**ABBREVIATIONS** AIS = Abbreviated Injury Scale; AUROC = area under the receiver operating characteristic curve; CARE = Clinical Administrative Research Education; GCS = Glasgow Coma Scale; ISS = Injury Severity Score; LOS = length of stay; MAE = mean absolute error; MVC = motor vehicle collision; NTDB = National Trauma Data Bank; TMPM = trauma mortality prediction model; TTA = time to admission.

regions of the AIS. This method produces a score ranging from 0 to 75, with 75 representing the most profound trauma. The ISS soon became a widely accepted standard metric to reproducibly describe trauma and anticipate mortality. Since that time, a number of researchers have attempted to augment the predictive capabilities of the ISS by integrating demographic or clinical data to contextualize traumatic injuries in specific populations.

Predictive models to date, however, have been limited by the conventional statistical methods used in their generation. Increasingly, machine learning algorithms such as neural networks have been applied to biomedical science. These tools allow researchers to probe large data sets of diverse variables whose relationships may be nonlinear or otherwise counterintuitive.

We recently developed a neural network–based model to predict clinically meaningful outcomes following MVC, including in-hospital mortality and overall hospital length of stay (LOS). These models can be applied as a novel tool for retrospective performance analysis by trauma centers based on clinical outcomes projection and may, with further validation, serve as a prognostic tool in clinical trauma settings.

Methods

Study Design and Databases

We queried two prospectively collected databases maintained at Ryder Trauma Center at Jackson Memorial Hospital, a level I free-standing trauma center, including 1) local data from the American College of Surgeons National Trauma Data Bank (NTDB) and 2) our institution’s Clinical Administrative Research Education (CARE) database. The NTDB is a national-level database consisting of voluntarily submitted data from participating centers, while the CARE database consists of prospective data collected at our institution before our participation in the NTDB. Our center began collecting data for the NTDB in 2013, with complete data available through the end of 2017.

As our study used de-identified patient data, it was considered exempt by our institution’s IRB.

Patient Cohorts

We compiled all records from the CARE database from 1998 to 2013 with those from the NTDB database from 2013 to 2017, thereby generating a 2-decade sample of MVC data. Together, these databases captured patients presenting to our trauma center during their respective collection periods, including a variety of demographic information and clinical variables describing the presentation, hospital course, and outcomes of individual patients. As these databases collect variables throughout a patient’s hospital course, the time of data collection does not necessarily reflect the time variables recorded into the electronic health record.

Outcome Assessment

Two specific outcomes were targeted for prediction: in-hospital mortality and overall hospital LOS. We sought to build an independent predictive model for each outcome.

We defined in-hospital mortality as any mortality that occurred after a patient was admitted to the hospital but before discharge. This outcome was analyzed by grouping hospital discharge disposition labels (e.g., “expired” and “died”) into a single proxy variable. We defined hospital LOS as the difference in days between hospital admission and discharge. Separate predictive models were constructed for each of these outcomes.

Both models included all patients who did not die in the trauma bay and were subsequently admitted to the hospital. As our patient population was heavily right skewed, only patients within the 95th percentile (LOS < 40 days) were included for final LOS model development.

Model Training and Validation

Data sets were randomly divided into outcome-balanced training, validation, and test sets in a 64/20/16 training-validation-testing split (Supplemental Table 1).

The in-hospital mortality and hospital LOS models included the following features: age, Glasgow Coma Scale (GCS) score, ISS, and time to admission (TTA) in minutes. The in-hospital mortality model was a classification model with a categorical prediction variable of mortality. Its performance was evaluated by conventional sensitivity/specificity and the area under the receiver operating characteristic curve (AUROC). The LOS model was a regression model with a continuous target variable of the time to hospital discharge. As its target variable is continuous (time, in days), its performance was evaluated by mean absolute error (MAE). Further details of model development and validation are given in the Supplemental Material.

Results

Overall Sample Characteristics

In total, 22,662 records with 59 variables were obtained from the CARE database (1998–2013), and 3494 records with 85 variables were obtained from the NTDB database (2013–2017). In CARE, 17,047 MVCs occurred, while 2989 MVCs occurred in NTDB, creating an initial MVC cohort of 20,036 patients; 52 variables were reported in both databases. After feature selection, 16,287 records were used for development of the in-hospital mortality and LOS models. Patient characteristics of the pre-exclusion cohort (all MVCs) are reported in Table 1.

Model Characteristics and Performance

The in-hospital mortality cohort had a mortality incidence of 6.6%, with 1070 patient deaths from the 16,287 records. The training, validation, and testing sets comprised 10,424, 3257, and 2606 patients, respectively. There was no significant difference in input variables across these sets. The in-hospital mortality model had an AUROC of 0.976 (95% CI 0.974–0.978), sensitivity of 92%, and specificity of 90%.

Our LOS model drew from the same subgroup as our in-hospital mortality model, as described above. In these patients, the mean hospital LOS was 9.9 days (SD 22.1 days, range 0–380 days). A majority of patients were admitted directly to the main hospital floor (52.3%). Other major dispositions included the intensive care unit (27.1%)
and the operating room (19.1%). Only a small proportion of patients (0.5%) were held in observation in the trauma center. Initially, our model was able to predict hospital LOS in the admitted subgroup with high accuracy (MAE 4.23 days). However, after further restricting this subgroup to patients with LOS within the 95th percentile of the entire sample, the model’s accuracy improved significantly (MAE 2.23 days).

Sample input variable characteristics for each model cohort are given in Table 2. Performance statistics for each model are given in Table 3. The workflow process, including sample construction, feature selection, and model development, is depicted graphically in Fig. 1. Model performance is represented in Fig. 2.

**Discussion**

Our models represent the first application of a neural network–based system to predict clinically meaningful outcomes following MVC trauma, including patient mortality and hospital LOS. The accuracy of these models in combination with reliance on a few, readily available clinical parameters make this a powerful tool to inform decision-making during hospital admission following MVC, as well as to retrospectively evaluate trauma center performance against a clinically based benchmark.

Although previous attempts have been made to predict trauma outcomes such as mortality, most have not targeted MVC as a specific mechanism, favoring a generalized approach to all mechanisms of injury. Most of these models have been anatomically based, following the example of the long-accepted ISS; however, since its development and general adoption, a number of anatomically derived injury severity measures have been shown to predict mortality with higher accuracy.

One such model, the trauma mortality prediction model (TMPM), has demonstrated superior accuracy in predicting mortality (AUROC 0.866) compared with the ISS (AUROC 0.832). The TMPM was initially developed to overcome long-recognized shortcomings of the ISS, including, among other factors, its reliance on AIS codes, its restriction to three injuries, and its emphasis on severity of injury per se rather than importance of the anatomical site of injury. Drawing from a large patient population in the NTDB, the TMPM was derived by regression modeling to predict mortality from ICD-9-CM injury codes directly. In their initial study, Glance et al. reported superior accuracy in mortality prediction using the TMPM compared with an ISS-based model (AUROC 0.880 vs 0.850). Later, Cook et al. redemonstrated superior accuracy of the TMPM compared with the ISS and its variants, whether scores were computed from AIS codes (AUROC 0.888 vs 0.851) or from ICD-9 codes (AUROC 0.872 vs 0.830). Recently, however, Wang et al. developed the injury mortality prediction model, demonstrating superior accuracy in mortality prediction compared with the TMPM (AUROC 0.903 vs 0.890). The authors also attempted to incorporate age, sex, and injury mechanism but found little significant impact on the models’ function, whereas age was a significant feature in developing our own model. In a recent analysis of trauma prediction models specifically applied to in-hospital mortality following MVC, Van Belleghem et al. found superior results with scores derived from ICD-9-CM codes (AUROC 0.94–0.96) compared with scores derived from AIS codes (AUROC 0.72–0.76).

While the prediction of mortality following trauma has been extensively researched, less attention has been paid to the hospital course of surviving patients. Specifically, there have been few attempts to predict hospital LOS following trauma. Moore et al. recently determined that discharge destination, age, transfer status, and injury severity (AIS-based) were the strongest determinants of hospital LOS following trauma. These findings were based on a multilevel linear regression in a large, multicenter cohort. Conversely, Hwabejire et al. suggest that system-level fac-

**TABLE 1. Pre-exclusion clinical features of the overall cohort**

<table>
<thead>
<tr>
<th>Value</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of pts</td>
<td>20,036</td>
</tr>
<tr>
<td>Mean age, yrs</td>
<td>37.0 ± 18.4</td>
</tr>
<tr>
<td>Mean GCS score</td>
<td>13.3 ± 2.9</td>
</tr>
<tr>
<td>Mean ISS</td>
<td>14.0 ± 11.2</td>
</tr>
<tr>
<td>Mean TTA, mins</td>
<td>555.2 ± 425.8</td>
</tr>
<tr>
<td>Death, n (%)</td>
<td>1333 (6.7)</td>
</tr>
</tbody>
</table>

Pts = patients.
Mean values are presented as the mean ± SD.

**TABLE 2. Model subgroup clinical features**

<table>
<thead>
<tr>
<th>Value</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of pts</td>
<td>16,287</td>
</tr>
<tr>
<td>Mean age, yrs</td>
<td>37.2 ± 19.4</td>
</tr>
<tr>
<td>Mean GCS score</td>
<td>13.3 ± 3.4</td>
</tr>
<tr>
<td>Mean ISS</td>
<td>14.0 ± 13.1</td>
</tr>
<tr>
<td>Death, n (%)</td>
<td>1070 (6.6)</td>
</tr>
<tr>
<td>Mean TTA, mins</td>
<td>555.3 ± 426.6</td>
</tr>
</tbody>
</table>

Mean values are presented as the mean ± SD.

**TABLE 3. Performance statistics for each model**

<table>
<thead>
<tr>
<th>Value</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>In-hospital mortality model, %</td>
<td></td>
</tr>
<tr>
<td>AUROC (95% CI)</td>
<td>97.6 (0.974–0.978)</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>92</td>
</tr>
<tr>
<td>Specificity</td>
<td>90</td>
</tr>
<tr>
<td>PPV</td>
<td>76</td>
</tr>
<tr>
<td>NPV</td>
<td>99</td>
</tr>
<tr>
<td>LOS model, days</td>
<td></td>
</tr>
<tr>
<td>MAE</td>
<td>4.23</td>
</tr>
<tr>
<td>≤95th percentile</td>
<td>2.23</td>
</tr>
</tbody>
</table>

NPV = negative predictive value; PPV = positive predictive value.
tors (e.g., discharge placement and in-hospital operational delays) are the primary determinants of prolonged hospital LOS in these patients. However, no group has thus far offered a reliable, clinically applicable tool to predict hospital LOS at the moment of admission following trauma. Our aim was to construct a larger sample by including all patients with trauma following an MVC over a 2-decade period, thereby generating a stronger model. Although our model is generalizable to MVC, our scope of interest remains in the model’s applicability to the field of neurotrauma. MVCs are responsible for up to 70% of neurotraumatic injuries, with devastating physical, psychosocial, and financial implications. Therefore, we advocate for the adoption of advanced data modeling and metrics in the neurotrauma population.

Ultimately, these tools can significantly inform management of patients following MVCs in the general trauma setting. During admission, patients with greater mor-

FIG. 1. Process diagram showing sample parameters, feature and classifier selection, and model architecture. From 16,287 MVC case records, 52 variables were considered for inclusion in our models (text boxes, upper left). Preliminary surveys of variable correlation identified features most strongly correlated with target outcomes (heat map, upper center). A variety of classifier types were compared for accuracy in predicting target outcomes (boxplot, lower center). Having identified features of interest and the most effective classifier type, our models were constructed using alternating dense and dropout layers as described in the article (network diagram, right). ER = emergency room; ROC = receiver operating characteristic.

FIG. 2. Model performance. Left: AUROC is shown for the in-hospital mortality model. Right: Iterative MAE over sequential training cycles is shown for the LOS model.
tality risk may require more frequent monitoring or more aggressive management, while the projected hospital LOS can inform the choice of analgesia, discharge disposition and planning, and general management scheduling. Figure 3 highlights the distinct points at which the model can meaningfully impact management decisions during patient hospitalization following an MVC.

Limitations and Future Directions

Although our model was derived from a large sample over an extended period of time, all patients were seen and treated at a single, freestanding trauma center. This geographic and clinical limitation may restrict the generalizability of our findings and thus our model's applicability. Although we did test our model on records withheld from the training data set, this is ultimately a retrospectively constructed and validated model; therefore, its true accuracy in prospective forecasting has yet to be demonstrated.

ISS is typically calculated at discharge, which would limit our model's potential secondary application for prognostication. However, clinical data normally obtained in patients with trauma (e.g., whole-body scans) can generate a working ISS.

The strongly right-skewed distribution of our initial patient population presented a mathematical dilemma in our LOS model construction; 95% of the initial patient cohort had an LOS within 39 days, while outliers above the 95th percentile reached a maximum LOS of 380 days. Initially, we attempted to create a model that could predict the entire patient population's LOS. We found, however, that any model attempting to include predictions for outliers lost significant final accuracy for the majority of patients. We elected to exclude outliers above the 95th percentile, thereby achieving maximum accuracy for a greater number of individuals. Our model is, therefore, less capable of accurately quantifying LOS in extreme outliers. However, we believe that this specific loss of accuracy is well balanced by the rarity of those cases.

Finally, and perhaps most importantly, variability in data recording obviously introduces some degree of confounding to our models. This is a fundamental limitation of any database study.

Conclusions

Our models effectively predict clinically meaningful outcomes for patients following an MVC. This may aid clinicians by benchmarking site performance and potentially forecasting outcomes, thereby optimizing immediate and expeditious patient care. In particular, these tools have great potential for application by neurosurgeons, who are frequently consulted in the management of patients involved in an MVC. Despite high performance presently, augmentation with larger samples from outside centers would enhance the generalizability of our models.

Acknowledgments

We are grateful to the research staff of Ryder Trauma Center at Jackson Memorial Hospital for their assistance in acquiring and interpreting data, and to Roberto Suazo for his assistance in graphic design and web development.

References


**Disclosures**

The authors report no conflict of interest concerning the materials or methods used in this study or the findings specified in this paper.

**Author Contributions**

Conception and design: Jagid, Kolcun. Acquisition of data: Kolcun, Covello. Analysis and interpretation of data: Covello, Gernsback, Cajigas. Drafting the article: Kolcun, Covello. Critically revising the article: all authors. Reviewed submitted version of manuscript: Jagid, Kolcun, Gernsback, Cajigas. Statistical analysis: Covello, Cajigas. Administrative/technical/material support: Jagid, Gernsback, Cajigas. Study supervision: Jagid, Kolcun.

**Supplemental Information**

Videos


Online-Only Content

Supplemental material is available online.


**Correspondence**

Jonathan R. Jagid: University of Miami Miller School of Medicine, Miami, FL. jjagid@med.miami.edu.