

RESEARCH ARTICLE

Mortality prediction in patients with isolated moderate and severe traumatic brain injury using machine learning models

Cheng-Shyuan Rau¹✉, Pao-Jen Kuo²✉, Peng-Chen Chien², Chun-Ying Huang³, Hsiao-Yun Hsieh², Ching-Hua Hsieh²*

1 Department of Neurosurgery, Kaohsiung Chang Gung Memorial Hospital and Chang Gung University College of Medicine, Taiwan, **2** Department of Plastic Surgery, Kaohsiung Chang Gung Memorial Hospital and Chang Gung University College of Medicine, Taiwan, **3** Department of Trauma Surgery, Kaohsiung Chang Gung Memorial Hospital and Chang Gung University College of Medicine, Taiwan

✉ These authors contributed equally to this work.

* m93chinghua@gmail.com



Abstract

OPEN ACCESS

Citation: Rau C-S, Kuo P-J, Chien P-C, Huang C-Y, Hsieh H-Y, Hsieh C-H (2018) Mortality prediction in patients with isolated moderate and severe traumatic brain injury using machine learning models. PLoS ONE 13(11): e0207192. <https://doi.org/10.1371/journal.pone.0207192>

Editor: Yu Ru Kou, National Yang-Ming University, TAIWAN

Received: April 11, 2018

Accepted: October 28, 2018

Published: November 9, 2018

Copyright: © 2018 Rau et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript.

Funding: This research was supported by a grant from Chang Gung Memorial Hospital CMRPG8G0491. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Background

The purpose of this study was to build a model of machine learning (ML) for the prediction of mortality in patients with isolated moderate and severe traumatic brain injury (TBI).

Methods

Hospitalized adult patients registered in the Trauma Registry System between January 2009 and December 2015 were enrolled in this study. Only patients with an Abbreviated Injury Scale (AIS) score ≥ 3 points related to head injuries were included in this study. A total of 1734 (1564 survival and 170 non-survival) and 325 (293 survival and 32 non-survival) patients were included in the training and test sets, respectively.

Results

Using demographics and injury characteristics, as well as patient laboratory data, predictive tools (e.g., logistic regression [LR], support vector machine [SVM], decision tree [DT], naive Bayes [NB], and artificial neural networks [ANN]) were used to determine the mortality of individual patients. The predictive performance was evaluated by accuracy, sensitivity, and specificity, as well as by area under the curve (AUC) measures of receiver operator characteristic curves. In the training set, all five ML models had a specificity of more than 90% and all ML models (except the NB) achieved an accuracy of more than 90%. Among them, the ANN had the highest sensitivity (80.59%) in mortality prediction. Regarding performance, the ANN had the highest AUC (0.968), followed by the LR (0.942), SVM (0.935), NB (0.908), and DT (0.872). In the test set, the ANN had the highest sensitivity (84.38%) in mortality prediction, followed by the SVM (65.63%), LR (59.38%), NB (59.38%), and DT (43.75%).

Conclusions

The ANN model provided the best prediction of mortality for patients with isolated moderate and severe TBI.

Background

Identifying patients with traumatic brain injury (TBI) with high risk of mortality is important to maximize the resource for trauma care, and so that family members receive appropriate counsel and treatment decisions [1, 2]. One widely applied predictor of mortality outcome is the Trauma and Injury Severity Score (TRISS), which shows good discrimination in identifying the patients with TBI at high risk of mortality [3]. However, the model was not specifically designed for use in patients with TBI and is not always associated with high performance [4]. For patients with TBI, two prediction models (Corticosteroid Randomization after Significant Head Injury [CRASH] and the International Mission for Prognosis and Analysis of Clinical Trials in Traumatic Brain Injury [IMPACT]) based on large clinical trial datasets have shown good discrimination and have enabled accurate outcome predictions [5–7]. However, these two models lack the precision required for use on the individual patient level [8, 9].

Currently, machine learning (ML) has been successfully applied to aid in clinical diagnosis and prognosis prediction [10, 11]. In some field-specific datasets, many ML techniques have shown significantly better predictive power than more conventional alternatives [12]. Applied ML techniques in the clinical setting include, but are not limited to, LR, support vector machine (SVM), decision trees (DT), Bayes classification, and artificial neural networks (ANN). This study aimed to construct an optimal model of ML for mortality prediction among patients with moderate and severe TBI by using data from a population-based trauma registry in a level I trauma center.

Methods

Subject and data preparation

This study was approved by the Institutional Review Board (IRB) of Chang Gung Memorial Hospital with approval number 201700014B0. Informed consent was waived according to the regulation of the IRB. Detailed patient information between January 2009 and December 2015 was retrieved from the Trauma Registry System of the hospital. The adult patient cohort included those who were ≥ 20 years of age and hospitalized for the treatment of moderate and severe TBI, defined as an AIS score ≥ 3 points in the head (moderate TBI, AIS 3–4; severe TBI, AIS 5) [13, 14]. Polytrauma patients who had additional AIS scores ≥ 3 points corresponding to any other region of the body were excluded from this study [14]. Enrolled patients were divided into a training set (a 6-year span between 2009 and 2014) for generation of a plausible model under supervised classification, and a test set (a 1-year span in 2015) to test the performance of the model. Patients with missing data were not included in the dataset for analysis. The retrieved patient information included the following variables: age, sex, helmet-wearing status, pre-existed co-morbidities, such as coronary artery disease (CAD), congestive heart failure (CHF), cerebral vascular accident (CVA), diabetes mellitus (DM), end-stage renal disease (ESRD), and hypertension (HTN). Glasgow coma scale (GCS) score and vital signs, including temperature, systolic blood pressure (SBP), heart rate (HR), and respiratory rate (RR) were collected upon patient arrival at the emergency department. Blood-drawn laboratory data at the

emergency room, including white blood cell count (WBC), red blood cell count (RBC), hemoglobin (Hb), hematocrit (Hct), platelets, blood urine nitrogen (BUN), creatinine (Cr), alanine aminotransferase (ALT), aspartate aminotransferase (AST), sodium, potassium, and glucose was also collected. TBI-related diagnoses, such as epidural hematoma (EDH), subdural hematoma (SDH), subarachnoid hemorrhage (SAH), and intracerebral hematoma (ICH) were assessed, as were ISS and the in-hospital mortality of patients during admission. The in-hospital mortality included those caused by the injury directly to the brain or by the associated complication such as pneumonia or sepsis. Finally, a total of 1734 (1564 survival and 170 non-survival) and 325 (293 survival and 32 non-survival) patients comprised the training and test sets, respectively.

ML classifications

Logistic regression (LR). In this study, the LR classifier used `glm` function in the `stats` package in R3.3.3 (R Foundation for Statistical Computing, Vienna, Austria). A stepwise LR analysis was used to control the effects of confounding variables to identify independent risk factors for mortality.

Support vector machine (SVM). The SVM classifier used the `tune.svm` & `svm` function in the `e1071` package in R with the radial basis function to handle non-linear interactions [15]. The optimal operating point was estimated using a grid search with a 10-fold cross-validation varying the penalty parameter C , which determined the tradeoff between fitting error minimization and model complexity, and hyper-parameter γ , which defined the nonlinear feature transformation onto a higher dimensional space and controlled the tradeoff between error due to bias and variance in the model [15].

Decision tree (DT). In this study, the classification and regression trees (CART) of DT were used based on the Gini impurity index to achieve the best overall split [16], with the `rpart` function in the `rpart` package in R to enable better prediction through progressive binary splits in a combined approach for both nonparametric and nonlinear variables.

Naïve bayes (NB). The NB classifier—deemed as the simplest of Bayes classifiers—makes the assumption that input variables are conditionally independent of each other given the classification [17, 18]. Despite its simplicity, its performance is comparable to conventional or more sophisticated methods [19, 20], and it has yielded good results in mortality-classification settings [21]. In this study, the NB classifier used the `naiveBayes` function in the `e1071` package in R.

Artificial neural networks (ANN). The ANN is constructed from a set of neurons that exchange signals with each other via an interconnected network. Each connection has a numeric weight that can be adjusted during training of the network, making the system adaptive to input patterns and capable of revealing previously unknown relationships between given input and output variables [22–24]. In this study, the ANN classifier used a feed-forward neural network with the `nnet` function in the `nnet` package in R. The tuning parameters included the number of nodes in the hidden layer optimized between 1 and 20. For the training process, maximal iterations and decay were selected as 1000 and 0.001, respectively. To avoid over-fitting, iterations occurred until the error did not significantly decrease.

Performance of the ML classifiers

Model predictive performance regarding accuracy, sensitivity, specificity and the area under the curve (AUC) of the receiver operator characteristic curves (ROCs) corresponding to the two different models was measured. A nonparametric approach to the analysis of the AUC under correlated ROCs using the `roc` & `roc.test` function in the `pROC` package in R was pursued [25]. The predicted probabilities against binary events was validated using the `val.prob`

function in the rms package in R. Somers' Dxy measured probability of concordance minus the probability of discordance between predicted outcomes and observed outcomes and was used to assess the predictive discrimination [26]. The coefficient of determination, R^2 , is a statistic that will give some information about the goodness of fit of a model [27]. An R^2 of 1 indicates that the regression line perfectly fits the data. The Brier score indicated an overall measure of model performance by defining the mean squared error between the predicted probabilities and the actual outcomes [28]. Brier scores vary between 0 and 1, a lower score indicating higher accuracy.

Statistical analyses

All statistical analyses were performed using SPSS 20.0 (IBM Inc., Chicago, IL, USA) and R 3.3.3. For continuous variables, we used student t-tests to analyze normally distributed data, while Kolmogorov-Smirnov or Mann-Whitney U tests were used to compare non-normally distributed data. For categorical variables, we used Chi-square tests to determine the significance of the association between variables. Results are presented as mean \pm standard deviation, and a p-value < 0.05 was considered statistically significant.

Results

Patient demographics and injury characteristics

As shown in Fig 1, there was no significant difference in sex, CAD, CHF, CVA, and DM between patients that survived TBI and those that did not. In contrast, ESRD and HTN incidences were higher in the fatality group than in the group of patients that survived. In addition, fewer patients in the fatality group had worn a helmet when compared to patients in the survival group. A statistically significant difference in age, ISS, GCS, glucose, temperature, HR, RBC, Hb, Hct, platelets, and K was also found between groups (S1 Fig). Because correlation coefficients between Hb and Hct, BUN and Cr, and AST and ALT were similar, only one of the three representative variables (i.e., Hct, BUN, and AST) was selected for further ML classification to prevent the inclusion of duplicate parameters. Therefore, 25 total variables were used for ML classifier imputation.

Performance of ML classifiers in the training set

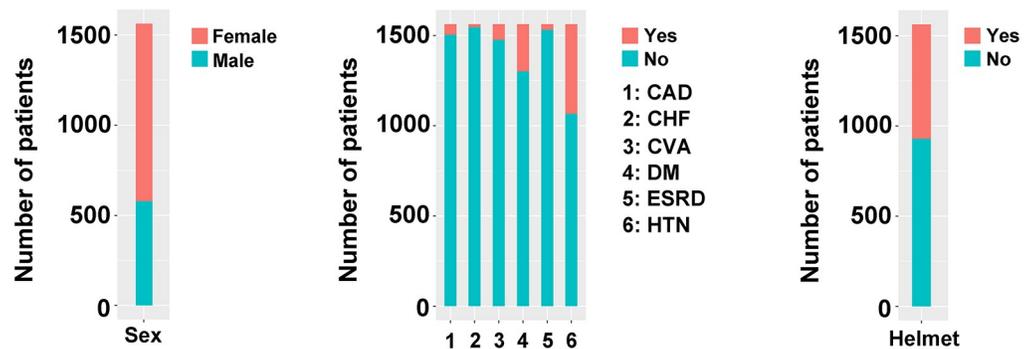
LR identified 12 predictors (age, helmet status, platelets, DM, SDH, Hct, temperature, ICH, GCS, glucose, ISS, and ESRD) as independent risk factors for mortality (S2 Fig).

The SVM classifier was performed for the prediction of mortality, taking input with all 25 variables with two parameters (C , γ) being determined by a grid search of 2^x , where x is an integer between -20 and 4 for C and between -20 and -4 for γ . The values which gave the highest 10-fold cross-validation accuracy were $C = 0.00003$ and $\gamma = 0.000977$.

In the DT model (Fig 2), the GCS was identified as the initial split variable, with an optimal cut-off value of > 4 . Among patients having a $GCS < 4$, high glucose levels of ≥ 216 mg/dL indicated a fatal outcome. In patients with glucose levels < 216 mg/dL, the next best predictor of mortality was age, with an optimal cut-off < 50 years. In patients ≥ 50 -years-old, existence of SDH presented as a predictor of fatal outcome. In addition, among patients with $GCS > 4$, an $ISS \geq 25$ and glucose levels ≥ 222 mg/dL were selected as significant variables for the prediction of a fatal outcome.

The constructed ANN model includes 25 inputs, one bias neuron in the input layer, eight hidden neurons, one bias neuron in the hidden layer, and one output neuron (Fig 3). A single output node indicated the probability of death.

Survival



Death

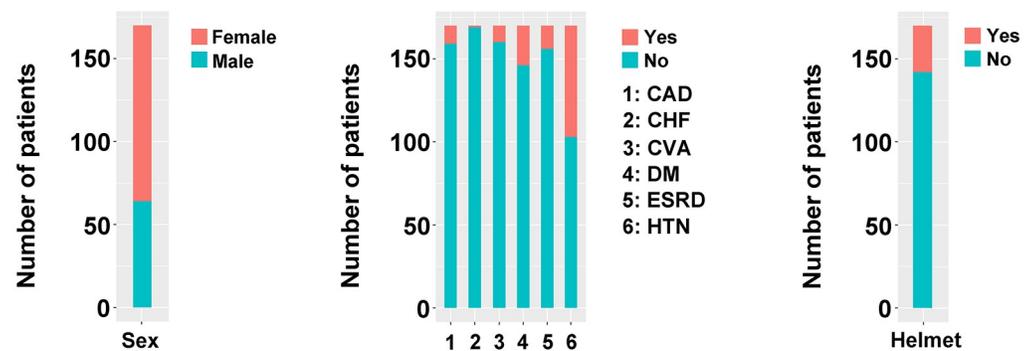


Fig 1. Patient demographics and injury characteristics.

<https://doi.org/10.1371/journal.pone.0207192.g001>

As shown in Table 1, all ML models, with the exception of the NB, achieved an accuracy of more than 90%. As the high survival among investigated patients rate likely accounts for the observed accuracy and specificity in mortality prediction, we further focused on the sensitivity of different ML models. In the established model, ANN had the highest sensitivity (80.59%), followed by the NB (73.53%), SVM (64.12%), DT (62.35%), and LR (53.89%). All five ML models had a specificity of more than 90%. In comparing AUCs of the ROCs among the five ML models for the training set (Fig 4), the ANN had a significantly higher AUC (0.968) than the other four ML models (S1 Table). Both LR (0.942) and SVM (0.935) had significantly higher AUCs than the NB (0.908) or DT (0.872); however, there was no significant difference in AUC between LR and the SVM. According to AUC comparisons, and in consideration of prediction sensitivity, the ANN was determined to be the best algorithm to predict mortality. The calibration curves of these five predictions demonstrated that LR, DT, and ANN plotted a nonparametric line close along the ideal diagonal line (Fig 5), while LR had the highest Dxy (0.884) and ANN had the highest R^2 (0.632) and lowest Brier score (0.036).

Performance of ML classifiers in the test set

For the test set, with the exception of the NB, all ML models achieved similar rates of accuracy (~92% to 93.5%) in terms of their ability to predict mortality (Table 1). In addition, these models achieved specificity rates that were greater than 92%. The ANN still displayed the highest

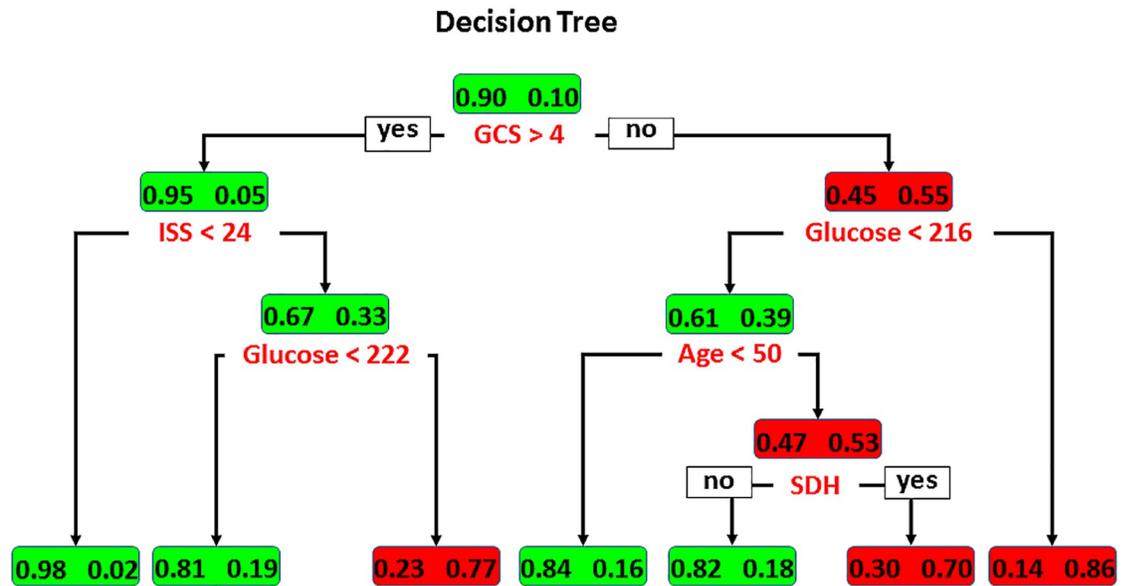


Fig 2. Illustration of the DT model for predicting mortality in patients with isolated moderate and severe TBI. Boxes denote the percentage of patients analyzed with discriminating variables; survivors and non-survivors are indicated by green and red colors, respectively.

<https://doi.org/10.1371/journal.pone.0207192.g002>

sensitivity (84.38%) in its ability to predict mortality, followed by the SVM (65.63%), LR and NB (59.38%), and DT (43.75%).

Discussion

In this study, both LR and the SVM had significantly higher AUCs than the NB and DT. It has been reported that the NB classifier operates under the assumption that independence is valid.

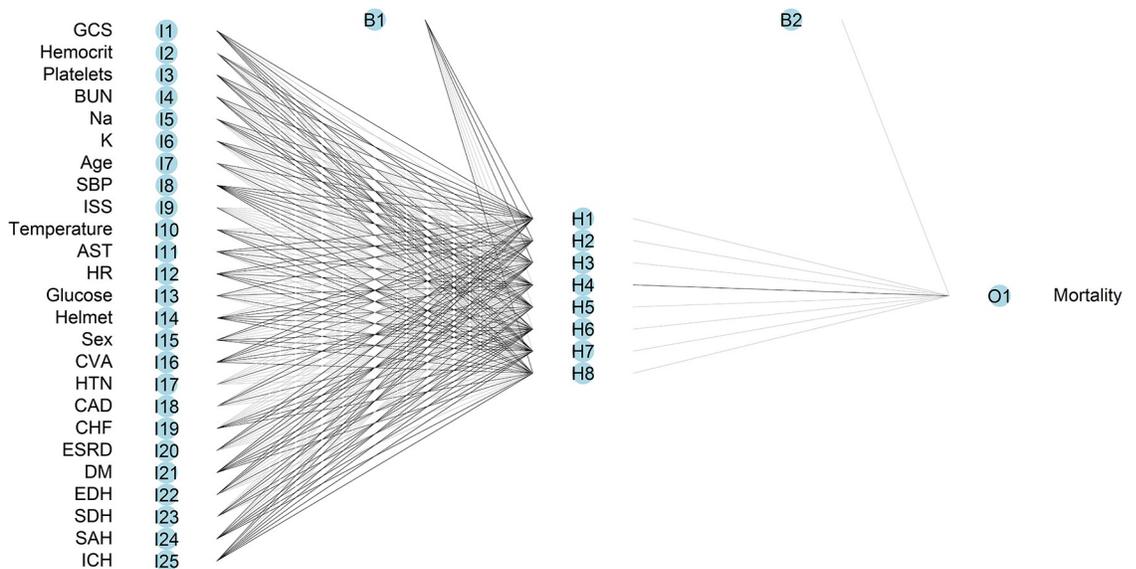


Fig 3. Architecture of the three-layered feed-forward ANN.

<https://doi.org/10.1371/journal.pone.0207192.g003>

Table 1. Mortality prediction performance (i.e., accuracy, sensitivity, and specificity) for the LR, SVM, DT, NB, and ANN models on training and test sets.

| Methods | Train | | | Test | | |
|---------|----------|-------------|-------------|----------|-------------|-------------|
| | Accuracy | Sensitivity | Specificity | Accuracy | Sensitivity | Specificity |
| LR | 93.66% | 53.89% | 98.08% | 93.54% | 59.38% | 93.54% |
| SVM | 92.96% | 64.12% | 96.10% | 92.50% | 65.63% | 95.22% |
| DT | 94.69% | 62.35% | 98.21% | 92.92% | 43.75% | 98.29% |
| NB | 89.56% | 73.53% | 91.30% | 86.15% | 59.38% | 89.08% |
| ANN | 93.94% | 80.59% | 95.40% | 92.00% | 84.38% | 92.83% |

LR, logistic regression; SVM, support vector machine; DT, decision trees; NB, Naive Bayes; and ANN, artificial neural networks.

<https://doi.org/10.1371/journal.pone.0207192.t001>

When two variables are related, however, the NB may place too much weight on them and too little weight on other variables, resulting in classification bias [17, 18]. In addition, the DT may lead to overestimation of the importance of included risk factors or may exclude other potential confounding factors that could influence actual risk [29]. On the other hand, the SVM boundary is only minimally influenced by outliers [30] and the employment of kernels help the model learn non-linear decision boundaries, allowing the classifier to solve more complex data than linear analyses methods, such as the LR model [31]. However, this advantage could not be identified in the current study, as there was no significant difference in AUC values

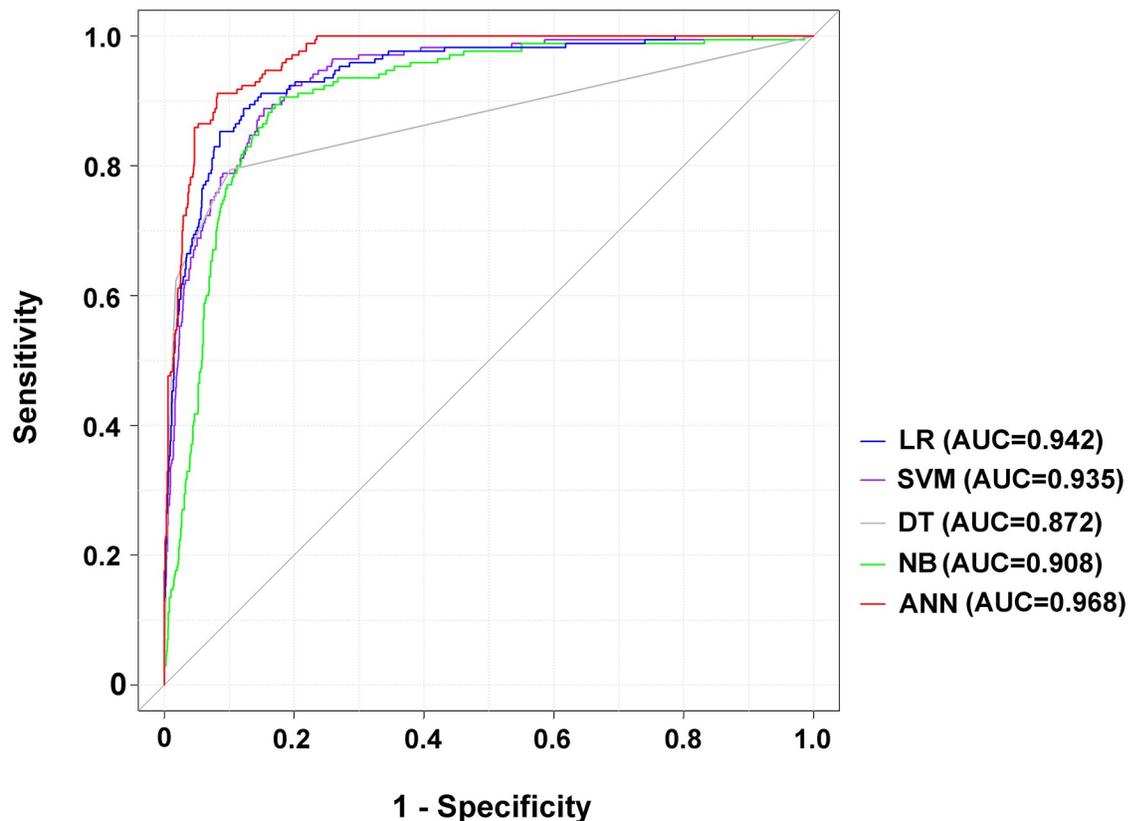


Fig 4. ROC curves for the LR, SVM, DT, NB, and ANN models in predicting the mortality of patients with isolated moderate and severe TBI.

<https://doi.org/10.1371/journal.pone.0207192.g004>

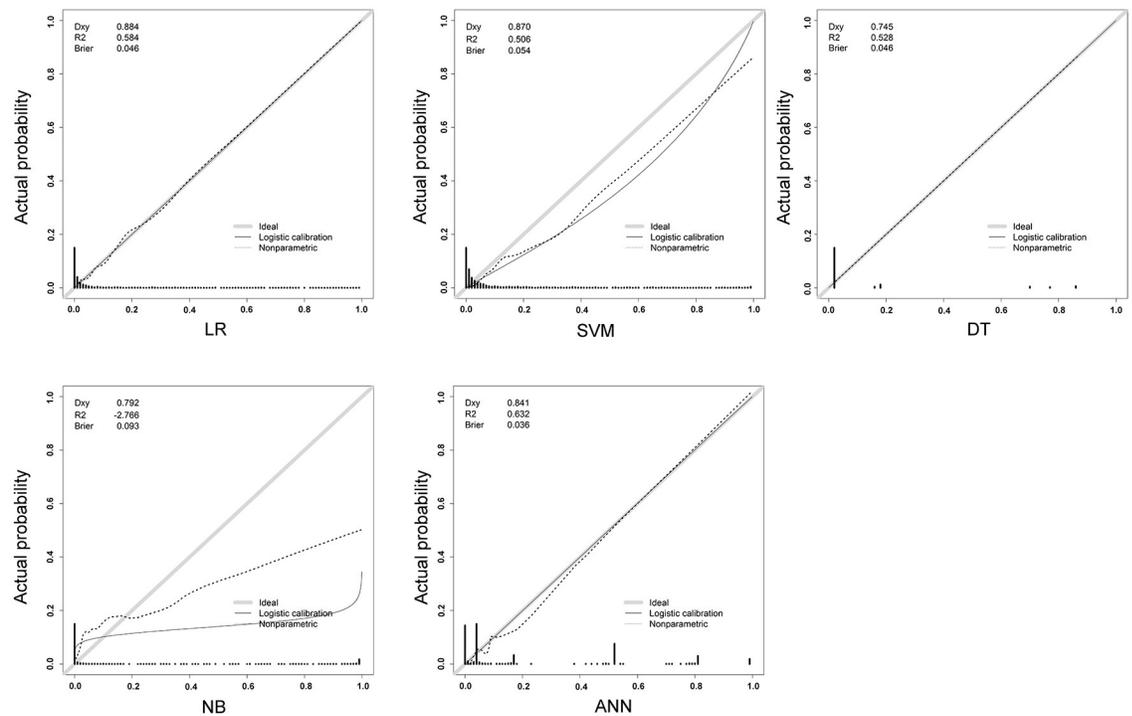


Fig 5. Calibration curves by the LR, SVM, DT, NB, and ANN models in predicting the mortality of patients with isolated moderate and severe TBI.

<https://doi.org/10.1371/journal.pone.0207192.g005>

between the LR and SVM. The exceptional performance of the LR model implies that most patients that died following TBI could be explained by a relatively small set of independent predictors that fit the logistic model assumptions well, such as old age [32], hyperglycemia [33], GCS scores [32, 34], and the presence of SDH [35]. Therefore, to improve the predictive performance of models than the LR, additional data or different approaches may be considered for the prediction task.

One study that predicted 10 609 trauma patient outcomes, using 16 anatomic and physiologic predictor variables, revealed that the ANN exceeded the TRISS model in terms of its ability to predict mortality, with an AUC of 0.912 compared to 0.895 for the TRISS [36]. Furthermore, the ANN model has been shown to be more accurate and to have better overall performance than LR model in predicting in-hospital mortality for patients receiving mechanical ventilation [37] as well as for patients in critical care [38]. For patients with TBI, the CRASH prognostic model shows good discrimination for 14-day mortality prediction, with an AUC of 0.89, while the IMPACT prognostic model shows good discrimination for 6-month mortality predictions, with an of AUC 0.80 [3]. In this study, we demonstrated that the ANN had a significantly higher AUC than the other four ML models when predicting the mortality of patients with isolated moderate and severe TBI. Moreover, the ANN retained the highest sensitivity (84.38%) among the investigated algorithms. Notably, the LR model provides odds ratio estimates for risk factors only under conditions with numbers of variables less than 20 [39]. A complex dataset with many predictors make LR model difficult to specify all possible interactions [40]. In contrast, the computational power of the ANN is derived from the distributed nature of its connections. As such, the ANN can successfully manages complex datasets, even when the sample size is small or the ratio between variables and records is unbalanced

[41], making it a natural modeling tool to examine survival in diverse populations [42]. In this study, we had included more variables such as preexisting comorbidities, helmet status, and laboratory data to improve prediction performance. Furthermore, the lack of specific mortality-related information in the trauma registry, including imaging characteristics of computed tomography scans [43], pupillary reactivity [7], and the existence of conscious deterioration, may have rendered the prediction model a space for improvement.

One criticism of the ANN is that it is difficult to assess the relative contribution of each variable to the final prediction put forth by the model [24]. Additionally, the ANN does not provide detailed information, such as the hazard ratio, which generally indicates the direction and magnitude of influence each variable has on the outcome [44]. Some other limitations included: First, patients declared dead on arrival at the hospital or at the scene of the accident were not recorded in the registered database [45, 46] and may have resulted in potential sample bias. The unknown status of inter-facility transfer or resuscitation for prehospital cardiac arrest may lead to a bias in the outcome measurement. Further, because the registered trauma data only had the in-hospital mortality but there were no information regarding the mortalities at 30 days, 3 months, or half a year data, there may exist some selection bias in the outcome measurement. Second, the imputation of physiological and laboratory data collected from the time of emergency department arrival does not reflect changes in hemodynamics and metabolic variables of patients who were under possible resuscitation. Third, this study was unable to assess the effects of any one particular treatment intervention, especially brain surgery. As such, we relied on the assumption that the assessment and management of patients—especially regarding operation quality—was uniform across the included population. Finally, the study sample was limited to a single urban trauma center in southern Taiwan, which may not be representative of other populations. However, because ANN generally could handle complex data which have interactions better than LR. Since ANN could do a better performance in mortality prediction than LR in this study according to the data from a single center, then we expect that ANN may still performed better than LR in dealing with the data from multicentric or multinational source. However, such opinion requires a further validation in the future study.

Conclusion

We demonstrated that the ANN model provided better performance in predicting the mortality of patients with isolated moderate and severe TBI. The results of studies published so far are encouraging and may provide the first steps towards the development of a prediction model that can be integrated into trauma care systems to identify patients at high risk for mortality.

Supporting information

S1 Fig. Patient injury characteristics, as well as physiological and laboratory data collected on patient arrival to the emergency department.

(TIF)

S2 Fig. The independent risk predictors factors in logistic regression model.

(TIFF)

S1 Table. Statistical *p*-value among AUC comparisons between LR, SVM, DT, NB, and ANN models in the training set.

(DOCX)

Acknowledgments

This research was supported by a grant from Chang Gung Memorial Hospital CMRPG8G0491. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author Contributions

Conceptualization: Ching-Hua Hsieh.

Formal analysis: Peng-Chen Chien.

Project administration: Hsiao-Yun Hsieh.

Validation: Chun-Ying Huang.

Writing – original draft: Cheng-Shyuan Rau.

Writing – review & editing: Pao-Jen Kuo.

References

1. Densmore JC, Lim HJ, Oldham KT, Guice KS. Outcomes and delivery of care in pediatric injury. *Journal of pediatric surgery*. 2006; 41(1):92–8; discussion -8. Epub 2006/01/18. <https://doi.org/10.1016/j.jpedsurg.2005.10.013> PMID: 16410115.
2. Rogers SC, Campbell BT, Saleheen H, Borrup K, Lapidus G. Using trauma registry data to guide injury prevention program activities. *The Journal of trauma*. 2010; 69(4 Suppl):S209–13. Epub 2010/10/22. <https://doi.org/10.1097/TA.0b013e3181f1e9fe> PMID: 20938310.
3. Wong GK, Teoh J, Yeung J, Chan E, Siu E, Woo P, et al. Outcomes of traumatic brain injury in Hong Kong: validation with the TRISS, CRASH, and IMPACT models. *Journal of clinical neuroscience: official journal of the Neurosurgical Society of Australasia*. 2013; 20(12):1693–6. Epub 2013/09/03. <https://doi.org/10.1016/j.jocn.2012.12.032> PMID: 23993210.
4. Demetriades D, Chan L, Velmanos GV, Sava J, Preston C, Gruzinski G, et al. TRISS methodology: an inappropriate tool for comparing outcomes between trauma centers. *J Am Coll Surg*. 2001; 193(3):250–4. Epub 2001/09/11. PMID: 11548794.
5. Han J, King NK, Neilson SJ, Gandhi MP, Ng I. External validation of the CRASH and IMPACT prognostic models in severe traumatic brain injury. *J Neurotrauma*. 2014; 31(13):1146–52. Epub 2014/02/27. <https://doi.org/10.1089/neu.2013.3003> PMID: 24568201.
6. Sun H, Lingsma HF, Steyerberg EW, Maas AI. External Validation of the International Mission for Prognosis and Analysis of Clinical Trials in Traumatic Brain Injury: Prognostic Models for Traumatic Brain Injury on the Study of the Neuroprotective Activity of Progesterone in Severe Traumatic Brain Injuries Trial. *J Neurotrauma*. 2016; 33(16):1535–43. Epub 2015/12/15. <https://doi.org/10.1089/neu.2015.4164> PMID: 26652051.
7. Perel P, Arango M, Clayton T, Edwards P, Komolafe E, Poccock S, et al. Predicting outcome after traumatic brain injury: practical prognostic models based on large cohort of international patients. *BMJ*. 2008; 336(7641):425–9. Epub 2008/02/14. <https://doi.org/10.1136/bmj.39461.643438.25> PMID: 18270239
8. Carter EL, Hutchinson PJ, Koliass AG, Menon DK. Predicting the outcome for individual patients with traumatic brain injury: a case-based review. *Br J Neurosurg*. 2016; 30(2):227–32. Epub 2016/02/09. <https://doi.org/10.3109/02688697.2016.1139048> PMID: 26853860.
9. Stevens RD, Sutter R. Prognosis in severe brain injury. *Crit Care Med*. 2013; 41(4):1104–23. Epub 2013/03/27. <https://doi.org/10.1097/CCM.0b013e318287ee79> PMID: 23528755.
10. Goldstein BA, Navar AM, Carter RE. Moving beyond regression techniques in cardiovascular risk prediction: applying machine learning to address analytic challenges. *Eur Heart J*. 2016. Epub 2016/07/21. <https://doi.org/10.1093/eurheartj/ehw302> PMID: 27436868.
11. Szlosek DA, Ferrett J. Using Machine Learning and Natural Language Processing Algorithms to Automate the Evaluation of Clinical Decision Support in Electronic Medical Record Systems. *EGEMS (Washington, DC)*. 2016; 4(3):1222. Epub 2016/09/30. <https://doi.org/10.13063/2327-9214.1222> PMID: 27683664
12. Yahya N, Ebert MA, Bulsara M, House MJ, Kennedy A, Joseph DJ, et al. Statistical-learning strategies generate only modestly performing predictive models for urinary symptoms following external beam

- radiotherapy of the prostate: A comparison of conventional and machine-learning methods. *Med Phys*. 2016; 43(5):2040. Epub 2016/05/06. <https://doi.org/10.1118/1.4944738> PMID: 27147316.
13. Savitsky B, Givon A, Rozenfeld M, Radomislensky I, Peleg K. Traumatic brain injury: It is all about definition. *Brain Inj*. 2016; 30(10):1194–200. Epub 2016/07/29. <https://doi.org/10.1080/02699052.2016.1187290> PMID: 27466967.
 14. Rau CS, Kuo PJ, Wu SC, Chen YC, Hsieh HY, Hsieh CH. Association between the Osteoporosis Self-Assessment Tool for Asians Score and Mortality in Patients with Isolated Moderate and Severe Traumatic Brain Injury: A Propensity Score-Matched Analysis. *Int J Environ Res Public Health*. 2016; 13(12). Epub 2016/12/06. <https://doi.org/10.3390/ijerph13121203> PMID: 27918475
 15. Gultepe E, Green JP, Nguyen H, Adams J, Albertson T, Tagkopoulos I. From vital signs to clinical outcomes for patients with sepsis: a machine learning basis for a clinical decision support system. *J Am Med Inform Assoc*. 2014; 21(2):315–25. Epub 2013/08/21. <https://doi.org/10.1136/amiajnl-2013-001815> PMID: 23959843
 16. Guilbault RWR, Ohlsson MA, Afonso AM, Ebell MH. External Validation of Two Classification and Regression Tree Models to Predict the Outcome of Inpatient Cardiopulmonary Resuscitation. *J Viral Hepat*. 2017; 32(5):333–8. Epub 2016/10/01. <https://doi.org/10.1177/0885066616686924> PMID: 28049389.
 17. Langarizadeh M, Moghbeli F. Applying Naive Bayesian Networks to Disease Prediction: a Systematic Review. *Acta informatica medica: AIM: journal of the Society for Medical Informatics of Bosnia & Herzegovina: casopis Društva za medicinsku informatiku BiH*. 2016; 24(5):364–9. Epub 2017/01/13. <https://doi.org/10.5455/aim.2016.24.364-369> PMID: 28077895
 18. Marucci-Wellman HR, Lehto MR, Corns HL. A practical tool for public health surveillance: Semi-automated coding of short injury narratives from large administrative databases using Naive Bayes algorithms. *Accid Anal Prev*. 2015; 84:165–76. Epub 2015/09/29. <https://doi.org/10.1016/j.aap.2015.06.014> PMID: 26412196.
 19. Lucas P. Bayesian analysis, pattern analysis, and data mining in health care. *Current opinion in critical care*. 2004; 10(5):399–403. Epub 2004/09/24. PMID: 15385759.
 20. Bellazzi R, Zupan B. Predictive data mining in clinical medicine: current issues and guidelines. *Int J Med Inform*. 2008; 77(2):81–97. Epub 2006/12/26. <https://doi.org/10.1016/j.ijmedinf.2006.11.006> PMID: 17188928.
 21. Miasnikof P, Giannakeas V, Gomes M, Aleksandrowicz L, Shestopaloff AY, Alam D, et al. Naive Bayes classifiers for verbal autopsies: comparison to physician-based classification for 21,000 child and adult deaths. *BMC Med*. 2015; 13:286. Epub 2015/11/27. <https://doi.org/10.1186/s12916-015-0521-2> PMID: 26607695
 22. Manning T, Sleator RD, Walsh P. Biologically inspired intelligent decision making: a commentary on the use of artificial neural networks in bioinformatics. *Bioengineered*. 2014; 5(2):80–95. Epub 2013/12/18. <https://doi.org/10.4161/bioe.26997> PMID: 24335433
 23. Zhang Z. A gentle introduction to artificial neural networks. *Annals of translational medicine*. 2016; 4(19):370. Epub 2016/11/09. <https://doi.org/10.21037/atm.2016.06.20> PMID: 27826573
 24. Patel JL, Goyal RK. Applications of artificial neural networks in medical science. *Current clinical pharmacology*. 2007; 2(3):217–26. Epub 2008/08/12. PMID: 18690868.
 25. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988; 44(3):837–45. Epub 1988/09/01. PMID: 3203132.
 26. Harrell FE Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996; 15(4):361–87. Epub 1996/02/28. [https://doi.org/10.1002/\(SICI\)1097-0258\(19960229\)15:4<361::AID-SIM168>3.0.CO;2-4](https://doi.org/10.1002/(SICI)1097-0258(19960229)15:4<361::AID-SIM168>3.0.CO;2-4) PMID: 8668867.
 27. Discacciati A, Crippa A, Orsini N. Goodness of fit tools for dose-response meta-analysis of binary outcomes. *Research synthesis methods*. 2017; 8(2):149–60. Epub 2015/12/19. <https://doi.org/10.1002/jrsm.1194> PMID: 26679736
 28. GW B. Verification of forecasts expressed in terms of probability. *Mon Wea Rev*. 1950;(78):1–3.
 29. Lemon SC, Roy J, Clark MA, Friedmann PD, Rakowski W. Classification and regression tree analysis in public health: methodological review and comparison with logistic regression. *nn Behav Med*. 2003; 26(3):172–81. Epub 2003/12/04. https://doi.org/10.1207/S15324796ABM2603_02 PMID: 14644693.
 30. Chen S, Zhou S, Yin FF, Marks LB, Das SK. Investigation of the support vector machine algorithm to predict lung radiation-induced pneumonitis. *Med Phys*. 2007; 34(10):3808–14. Epub 2007/11/08. <https://doi.org/10.1118/1.2776669> PMID: 17985626

31. Orru G, Pettersson-Yeo W, Marquand AF, Sartori G, Mechelli A. Using Support Vector Machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review. *Neurosci Biobehav Rev.* 2012; 36(4):1140–52. Epub 2012/02/07. <https://doi.org/10.1016/j.neubiorev.2012.01.004> PMID: 22305994.
32. Junior JR, Welling LC, Schafranski M, Yeng LT, do Prado RR, Koterba E, et al. Prognostic model for patients with traumatic brain injuries and abnormal computed tomography scans. *Journal of clinical neuroscience: official journal of the Neurosurgical Society of Australasia.* 2017. Epub 2017/03/30. <https://doi.org/10.1016/j.jocn.2017.03.012> PMID: 28347684.
33. Kafaki SB, Alaedini K, Qorbani A, Asadian L, Haddadi K. Hyperglycemia: A Predictor of Death in Severe Head Injury Patients. *Clinical medicine insights Endocrinology and diabetes.* 2016; 9:43–6. Epub 2016/10/04. <https://doi.org/10.4137/CMED.S40330> PMID: 27695379
34. Marmarou A, Lu J, Butcher I, McHugh GS, Murray GD, Steyerberg EW, et al. Prognostic value of the Glasgow Coma Scale and pupil reactivity in traumatic brain injury assessed pre-hospital and on enrollment: an IMPACT analysis. *J Neurotrauma.* 2007; 24(2):270–80. Epub 2007/03/23. <https://doi.org/10.1089/neu.2006.0029> PMID: 17375991.
35. Lee JJ, Segar DJ, Morrison JF, Mangham WM, Lee S, Asaad WF. Subdural hematoma as a major determinant of short-term outcomes in traumatic brain injury. *J Neurosurg.* 2017:1–14. Epub 2017/02/12. <https://doi.org/10.3171/2016.5.jns.16255> PMID: 28186445.
36. DiRusso SM, Sullivan T, Holly C, Cuff SN, Savino J. An artificial neural network as a model for prediction of survival in trauma patients: validation for a regional trauma area. *J Trauma.* 2000; 49(2):212–20; discussion 20–3. Epub 2000/08/30. PMID: 10963531.
37. Chen CJ, Shi HY, Lee KT, Huang TY. In-hospital mortality prediction in patients receiving mechanical ventilation in Taiwan. *Am J Crit Care.* 2013; 22(6):506–13. Epub 2013/11/05. <https://doi.org/10.4037/ajcc2013950> PMID: 24186822.
38. Kim S, Kim W, Park RW. A Comparison of Intensive Care Unit Mortality Prediction Models through the Use of Data Mining Techniques. *Healthcare informatics research.* 2011; 17(4):232–43. Epub 2012/01/20. <https://doi.org/10.4258/hir.2011.17.4.232> PMID: 22259725
39. Knol MJ, Vandenbroucke JP, Scott P, Egger M. What do case-control studies estimate? Survey of methods and assumptions in published case-control research. *Am J Epidemiol.* 2008; 168(9):1073–81. Epub 2008/09/17. <https://doi.org/10.1093/aje/kwn217> PMID: 18794220.
40. Gu W, Vieira AR, Hoekstra RM, Griffin PM, Cole D. Use of random forest to estimate population attributable fractions from a case-control study of Salmonella enterica serotype Enteritidis infections. *Epidemiol Infect.* 2015; 143(13):2786–94. Epub 2015/02/13. <https://doi.org/10.1017/S095026881500014X> PMID: 25672399.
41. Zou J, Han Y, So SS. Overview of artificial neural networks. *Methods Mol Biol.* 2008; 458:15–23. Epub 2008/12/11. PMID: 19065803.
42. Becalick DC, Coats TJ. Comparison of artificial intelligence techniques with UKTRISS for estimating probability of survival after trauma. UK Trauma and Injury Severity Score. *J Trauma.* 2001; 51(1):123–33. Epub 2001/07/27. PMID: 11468479.
43. Mutch CA, Talbott JF, Gean A. Imaging Evaluation of Acute Traumatic Brain Injury. *Neurosurg Clin N Am.* 2016; 27(4):409–39. Epub 2016/09/18. <https://doi.org/10.1016/j.nec.2016.05.011> PMID: 27637393
44. Chiu HC, Ho TW, Lee KT, Chen HY, Ho WH. Mortality predicted accuracy for hepatocellular carcinoma patients with hepatic resection using artificial neural network. *TheScientificWorldJournal.* 2013; 2013:201976. Epub 2013/06/06. <https://doi.org/10.1155/2013/201976> PMID: 23737707
45. Hsieh CH, Hsu SY, Hsieh HY, Chen YC. Differences between the sexes in motorcycle-related injuries and fatalities at a Taiwanese level I trauma center. *Biomed J.* 2017; 40(2):113–20. Epub 2017/05/20. <https://doi.org/10.1016/j.bj.2016.10.005> PMID: 28521902.
46. Hsieh CH, Liu HT, Hsu SY, Hsieh HY, Chen YC. Motorcycle-related hospitalizations of the elderly. *Biomed J.* 2017; 40(2):121–8. Epub 2017/05/20. <https://doi.org/10.1016/j.bj.2016.10.006> PMID: 28521903.