A Comparison of Logistic Regression, CART, and Artificial Neural Network for Predicting Outcomes in Traumatic Brain Injury Patients

Vineet Kumar Kamal *, R. M. Pandey *

Vineet Kumar Kamal,
Ph.D. Scholar, Department of Biostatistics,
*All India Institute of Medical Sciences,
New Delhi -110029, India.
vineetstats@gmail.com
Mobile +91 9013970773
Contents of Today’s Presentation

- Introduction
- Motivation
- Methods
- Results
- Summary
Contents of Today’s Presentation

- Introduction
- Motivation
- Methods
- Results
- Summary
- References
Traumatic brain injury (TBI) is a significant public health problem and poses a leading cause of disability and mortality in all regions of the globe.

Predicted to surpass many diseases as a major cause of death and disability by the year 2020 (World Health Organization. Projections of Mortality, 2002).

Main cause of one third to one half of all trauma deaths and the leading cause of disability in people under 40, severely disabling 15-20 per 100,000 populations per year (Fleminger, 2005).

Asia has the highest percentage of TBI-related outcomes (Adnan & Puvanachandra, 2009).

Also, a leading cause of mortality, morbidity, disability, and socioeconomic losses in India as well as in other developing countries.

India and other developing countries are facing the major challenges of prevention, pre-hospital care and rehabilitation in their rapidly changing environments to reduce the burden of TBIs (Gururaj, 2002).
Introduction

Statistical modelling?

- In medicine, prognosis is central

- Like diagnosis and treatment, prognosis is a fundamental responsibility of all clinicians after a TBI (Holland & Shigaki, 1998; Junqué, Bruna, & Mataró, 1997)

- In TBI, there are many factors that may affect outcome

- Statistical modelling for prognostication, hypothesis generation and stratification of patients in research studies (Helmy, Timofeev, & Hutchinson, 2010)

- Accurate prognostication can help in justifiable transfer to neurosurgical specialist services as well as in early management of the individual patient and to advice patient’s relatives

- Intelligent application of statistical models can improve understanding of the pathology and treatment of TBI
Introduction

Model assumptions and predictions

- Better predictions if assumptions are met
- Some violation inherent in empirical data
- Evaluate predictions in new data

Evaluation of predictions

Calibration: observed vs predicted
- average of predictions correct?
- low and high predictions correct?
- Graphically, including deciles
  (links to Hosmer-Lemeshow goodness of fit test)
- Specific subgroups of patients

Discrimination
- distinguish low risk from high risk patients?
- AUC
Motivation

- A statistical model is a powerful tool for evaluating trauma care.

- Titterington et al., 1981 demonstrated that it was the choice of variables and the setting in which models were applied which is more important rather than the formulae.

- Logistic regression method is applied most often but recursive portioning with construction of prediction tree may be attractive to clinicians because of simple presentation.

- Neural networks are becoming more popular due to their flexibility to predict the outcome when the relationship between the variables is complex, multidimensional, and nonlinear.

- All these techniques might be used and explored parallel in the future.

- There is no study which has used a common data set to compare the prediction performance of Logistic Regression, Classification and Regression Tree, and Artificial Neural Network for predicting in-hospital mortality and unfavorable outcome at 6-months post admission in moderate and severe head injury patients.
Methods

Total Eligible Patients (admitted within 72 hours of injury)
Inclusion criteria: GCS≤ 12 at admission in emergency department & admitted to ICU

Model Prediction
Outcome: In-hospital mortality
Model development data set (n = 1466)
(from May 2010 to July 2012)
Model validation data set (n = 316)
(from August 2012 to February 2013)

Model Prediction
Outcome: 6-month Unfavorable Outcome
Model development data set (n = 1007)
(from May 2010 to July 2012)
Model validation data set (n = 269)
(from August 2012 to February 2013)
Source of data: Jai Prakash Narayan Apex Trauma Center (JPNATC), AIIMS, New Delhi (India) is the best integrated level 1 trauma centre in India
Development of Models: Binary Logistic Regression, Classification and Regression Tree (CART), and Neural Network

Outcome(s)
1. In-hospital mortality
2. Unfavorable outcome: death, persistent vegetative state and severe disability, or GOS=1, 2, 3 & favourable outcome: moderate disability, good recovery or GOS=4,5

Predictor variables: demographic (Age (yrs), Sex), clinical (Motor Score, Pupillary reactivity, Limb movement, Cause of injury, Major extracranial injury, Duration before admission (hrs)), secondary insult (Hypotension), CT reports (degree of MLS, SDH, EDH, basal cistern effaced, tSAH/IVH).

Comparison of Models -> Performances of Models: Discrimination (Sensitivity, Specificity, Positive Predictive Value, Negative Predictive Value, Area Under ROC) and Calibration (Hoshmer-Lemeshow (H-L) goodness of fit test), and Overall performance (Brier Score)
Recommendations for developing and validating prognostic models in traumatic brain injury (Source: Mushkudiani et al., 2008)

Study population

• Large sample size (N>500)
• Reflects the inherent heterogeneity
• (in terms of injury type and severity) of the disease
• Representative for current practice

Predictors

• Plausible, based on previous research and expert opinion
• Precisely defined
• Measurable with little observer variability
• Readily obtainable

Outcome

• Assessed at a fixed time point
• Relevant to the disease (e.g., mortality/Glasgow Outcome Scale/neuropsychological measures/quality of life)
• Precisely defined
• Measurable with little observer variability

Model development

• Valid handling of missing predictor values, for example, with statistical imputation
• Use of appropriate statistical techniques for selection of predictors and estimation of prognostic effects
• Presentation in a readily applicable format

Model validation

• Internal validation with efficient procedures, for example, with bootstrapping
• External validation on patients different in time and/or place
• Performance assessment with sensible and interpretable measures, evaluating calibration and discrimination aspects
Logistic regression

- Statistical model for the analysis of binary responses
- In terms of the probability that \( Y = 1 \) given \( X \), the values of the predictors:

\[
P(Y = 1 | X) = \left[ 1 + \exp(-X\beta) \right]^{-1}
\]

where, \( X\beta \) stands for \( \beta_0 + \beta_1 X_1 + \beta_2 X_2 \ldots + \beta_k X_k \) (Harrell, 2006)

If \( X, X_2, \ldots, X_n \) denote \( n \) predictor variables, \( Y \) denotes the hospital mortality \( (Y = 1) \), and \( p \) denotes the probability of hospital mortality (i.e., the probability that \( Y = 1 \)), the following equation describes the relationship between the predictor variables and \( p \):

\[
\log \left( \frac{p}{1-p} \right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 \ldots + \beta_n X_n,
\]

Where, \( \beta_0 \) is a constant and \( \beta_1, \beta_2, \ldots, \beta_n \) are the regression coefficients of the predictor variables \( X, X, \ldots, X_n \).
Classification and Regression Trees (CART)

- Classification and regression tree (CART) is a non-parametric developed by Brieman et al. (1984), which uses so-called learning sample - a set of historical data with pre-assigned classes for all observations - to construct so-called decision tree, representing a classification system or predictive model.

- The basic idea here is to derive a tree that consists of a series of binary decisions that lead to patient classification (Breiman et al. 1984).
The CART graphic indicates the degree of increased homogeneity induced by each split. Trees can then be pruned back to produce a classification rule that makes clinical sense and is fairly easy to remember.

Let \( s \) be a split at node \( t \). Then, the goodness of split is defined as decrease in impurity measured by

\[
\Delta \bar{t}(s, t) = \bar{t}(t) - p_L[\bar{t}(t_L)] - p_R[\bar{t}(t_R)]
\]

Where \( s = \) a particular split
\( p_L = \) the proportion of the cases at node \( t \) that go into the left child node, \( t_L \),
\( p_R = \) the proportion of the cases at node \( t \) that go into the left child node, \( t_R \),
\( \bar{t}(t_L) = \) impurity of the left child node, and
\( \bar{t}(t_R) = \) impurity of the right child node.

The complexity of a tree is measured by the number of its terminal nodes, and the resubstitution misclassification rate is an accuracy measure that always improves as trees gets larger. Breiman, Friedman, Olshen and Stone (1984) suggest the following cost complexity measure for any tree,

\[
R_\alpha(T) = R(T) + \alpha |T|
\]

Where, \( R(T) \) is the resubstitution estimate of the misclassification rate of a tree, \( T \), and \( |T| \) is the number of terminal nodes of the tree, \( R_\alpha(T) \) is the cost–complexity measure, for a tree, \( T \). for each \( \alpha \geq 0 \) (Breiman et al., 1984).

For each value of \( \alpha \), CART searches for the subtree that minimizes the cost-complexibility measure.
Advantages of this method

 It gives a rule that is intelligible to clinicians and can be judged by its clinical criteria.

 In case of data sets having both a large number of cases and a large number of variables,

 resistant to outliers (Dan Steinberg and Phillip Colla, 1995).

 can be used to identify important variables and complex interactions between predictors which may be difficult or impossible to uncover using traditional multivariate techniques.

 this method can also be used for relatively simple tasks, such as the imputation of missing values (Harrell, 2006).

Disadvantage is that, when applied to continuous covariates it looses information due to the fact that it dichotomizes the selected variable at each split.
Neural Networks

- A nonlinear mathematical model for predicting or classifying system performance (i.e., system output) inspired by the structure and function of human biological neural networks.
- Based on algorithms that are patterned after the structure of the human brain and is developed, and derived to have a function similar to the human brain by memorizing and learning various tasks and behaving accordingly. Neural networks have the ability to “learn” mathematical relationships between a series of input (independent variables) and the corresponding output (dependent variable or outcome) variables.

Artificial neural network structure

$x_1, x_2, \ldots, x_n$ are input units, $w_1, w_2, \ldots, w_n$ are connection weights, $\sum$ is product of input units and connection weights, $f$ is transfer function (step, sigmoid and hyperbolic tangent sigmoid function), $Y$ is output, Bias is the complexity restriction that the neural network architecture imposes on the degree of fitting accurately the target.
Neural Networks

This method attempts to outperform the logistic regression approach by adopting models that vary from complex to extremely complex (Hinton 1992).

Advantages

- Great name
- Sometimes does better than logistic regression.

Disadvantages

Method is essentially a black box. You need a computer to apply it and it is very difficult to gain intuitive insight into what it is doing.
Contents of Today’s Presentation

- Introduction
- Motivation
- Methods
- Results
- Summary
Performance measures in prospective (external) validation for in-hospital mortality

<table>
<thead>
<tr>
<th>Model</th>
<th>LR</th>
<th>CART</th>
<th>ANN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n = 316</td>
<td>n = 316</td>
<td>n = 316</td>
</tr>
<tr>
<td>Predicted</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Yes</td>
<td>68 (TP)</td>
<td>26 (FP)</td>
<td>83 (TP)</td>
</tr>
<tr>
<td>No</td>
<td>37 (FN)</td>
<td>185 (TN)</td>
<td>22 (FN)</td>
</tr>
<tr>
<td>Sensitivity (95% CI)</td>
<td>64.8 (55.3, 73.2)</td>
<td>79.1 (70.3, 85.7)</td>
<td>69.5 (60.2, 77.5)</td>
</tr>
<tr>
<td>Specificity (95% CI)</td>
<td>87.7 (82.6, 91.5)</td>
<td>69.2 (62.7, 75.0)</td>
<td>86.7 (81.5, 90.7)</td>
</tr>
<tr>
<td>Positive predictive value (95% CI)</td>
<td>72.3 (62.6, 80.4)</td>
<td>56.1 (48.0, 63.8)</td>
<td>72.3 (62.9, 80.1)</td>
</tr>
<tr>
<td>Negative predictive value (95% CI)</td>
<td>83.3 (77.9, 87.7)</td>
<td>87.0 (81.0, 91.2)</td>
<td>85.1 (79.7, 89.3)</td>
</tr>
<tr>
<td>Diagnostic accuracy (95% CI)</td>
<td>80.0 (75.3, 84.1)</td>
<td>72.5 (67.3, 77.1)</td>
<td>81.0 (76.3, 85.0)</td>
</tr>
<tr>
<td>Area under ROC curve (95% CI)</td>
<td>0.86 (0.81, 0.90)</td>
<td>0.80 (0.75, 0.86)</td>
<td>0.86 (0.82, 0.91)</td>
</tr>
<tr>
<td>Hosmer-Lemeshow p-value</td>
<td>0.401</td>
<td>0.005</td>
<td>0.055</td>
</tr>
<tr>
<td>Brier Score</td>
<td>0.137</td>
<td>0.153</td>
<td>0.140</td>
</tr>
</tbody>
</table>
Comparison of Logistic regression, CART and Neural Network for model

Difference between areas under the ROC curves ($AUC_d = AUC_{row} - AUC_{column}$), and p-values

<table>
<thead>
<tr>
<th></th>
<th>CART</th>
<th>NN</th>
</tr>
</thead>
<tbody>
<tr>
<td>LR</td>
<td>$AUC_d = 0.055$</td>
<td>$AUC_d = -0.005$</td>
</tr>
<tr>
<td></td>
<td>$p = 0.003$</td>
<td>$p = 0.503$</td>
</tr>
<tr>
<td>CART</td>
<td>$AUC_d = -0.060$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$p = 0.002$</td>
<td></td>
</tr>
</tbody>
</table>
# Ranking of variables according to relative importance in models for outcome in-hospital mortality

<table>
<thead>
<tr>
<th>LR</th>
<th>CART</th>
<th>ANN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable (normalized importance %)</td>
<td>Variable (normalized importance %)</td>
<td>Variable (normalized importance %)</td>
</tr>
<tr>
<td>1. Age (100.0)</td>
<td>1. Hypotension (100)</td>
<td>1. Age (100.0)</td>
</tr>
<tr>
<td>2. Motor Score (100.0)</td>
<td>2. Motor score (95.2)</td>
<td>2. Hypotension (66.9)</td>
</tr>
<tr>
<td>3. Pupillary Reactivity (100.0)</td>
<td>3. Pupillary reactivity (89.2)</td>
<td>3. Motor score (65.3)</td>
</tr>
<tr>
<td>4. Hypotension (100.0)</td>
<td>4. Limb movement (42.2)</td>
<td>4. Pupillary reactivity (45.9)</td>
</tr>
<tr>
<td>5. SDH (32.0)</td>
<td>5. Age (18.3)</td>
<td>5. DBA (44.3)</td>
</tr>
<tr>
<td>6. tSAH/IVH (97.0)</td>
<td>6. Midline shift (5.5)</td>
<td>6. Basal cistern effaced (38.6)</td>
</tr>
<tr>
<td>7. Cause of injury (61.5)</td>
<td>7. DBA (5.1)</td>
<td>7. Degree of MLS (29.2)</td>
</tr>
<tr>
<td>8. Limb Movement (80.5)</td>
<td>8. tSAH/IVH (1.7)</td>
<td>8. Cause of injury (23.2)</td>
</tr>
<tr>
<td>9. Basal cistern effaced (85.5)</td>
<td>9. MEI (1.2)</td>
<td>9. Sex (21.8)</td>
</tr>
<tr>
<td>10. Midline Shift (7.0)</td>
<td>10. Cause of injury (0.1)</td>
<td>10. Limb movement (20.8)</td>
</tr>
<tr>
<td>11. EDH (8.0)</td>
<td>11. Basal cistern effaced (0.0)</td>
<td>11. tSAH/IVH (17.8)</td>
</tr>
<tr>
<td>12. Sex (13.0)</td>
<td>12. EDH (0.0)</td>
<td>12. MEI (10.3)</td>
</tr>
<tr>
<td>13. MEI (15.5)</td>
<td>13. SDH (0.0)</td>
<td>13. EDH (8.5)</td>
</tr>
<tr>
<td>14. Sex (0.0)</td>
<td>14. Sex (0.0)</td>
<td>14. SDH (7.0)</td>
</tr>
</tbody>
</table>

DBA: Duration before admission (hrs); MEI: Major extracranial injury; EDH: Epidural hematoma; SDH: Subdural hematoma
Performance measures in prospective (external) validation for unfavorable outcome at 6-months post trauma

<table>
<thead>
<tr>
<th>outcome: Unfavourable outcome</th>
<th>Models</th>
<th>LR</th>
<th>CART</th>
<th>ANN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Yes</td>
<td>146 (TP)</td>
<td>25 (FP)</td>
<td>120 (TP)</td>
<td>17 (FP)</td>
</tr>
<tr>
<td>No</td>
<td>22 (FN)</td>
<td>76 (TN)</td>
<td>48 (FN)</td>
<td>84 (TN)</td>
</tr>
</tbody>
</table>

Diagnostic Characteristics

<table>
<thead>
<tr>
<th>Diagnostic Characteristics</th>
<th>LR</th>
<th>CART</th>
<th>ANN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity (95% CI)</td>
<td>87.0 (81.0, 91.2)</td>
<td>71.4 (64.2, 77.7)</td>
<td>85.7 (79.6, 90.2)</td>
</tr>
<tr>
<td>Specificity (95% CI)</td>
<td>75.3 (66.0, 82.4)</td>
<td>83.2 (74.7, 89.2)</td>
<td>83.2 (74.7, 89.2)</td>
</tr>
<tr>
<td>Positive predictive value (95% CI)</td>
<td>85.4 (79.3, 90.0)</td>
<td>87.6 (81.0, 92.1)</td>
<td>89.4 (83.7, 93.3)</td>
</tr>
<tr>
<td>Negative predictive value (95% CI)</td>
<td>77.6 (68.3, 84.7)</td>
<td>63.6 (55.2, 71.4)</td>
<td>77.8 (69.1, 84.6)</td>
</tr>
<tr>
<td>Diagnostic Accuracy</td>
<td>82.5 (77.5, 86.6)</td>
<td>75.8 (70.4, 80.6)</td>
<td>84.8 (80.0, 88.6)</td>
</tr>
<tr>
<td>Area under ROC (95% CI)</td>
<td>0.91 (0.87, 0.94)</td>
<td>0.87 (0.83, 0.91)</td>
<td>0.91 (0.87, 0.94)</td>
</tr>
<tr>
<td>Hosmer-Lemeshow p-value</td>
<td>0.604</td>
<td>0.155</td>
<td>0.285</td>
</tr>
<tr>
<td>Brier Score</td>
<td>0.12</td>
<td>0.14</td>
<td>0.12</td>
</tr>
</tbody>
</table>
Comparison of Logistic regression, CART and Neural ANN

Difference between areas under the ROC curves ($AUC_d = AUC_{row} - AUC_{column}$), and p-values

<table>
<thead>
<tr>
<th></th>
<th>CART</th>
<th>NN</th>
</tr>
</thead>
<tbody>
<tr>
<td>LR</td>
<td>AUC$_d$ = 0.039</td>
<td>AUC$_d$ = 0.0009</td>
</tr>
<tr>
<td></td>
<td>p = 0.0004</td>
<td>p = 0.876</td>
</tr>
<tr>
<td>CART</td>
<td>AUC$_d$ = -0.038</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p = 0.003</td>
<td></td>
</tr>
</tbody>
</table>
### Ranking of variables according to relative importance in models for outcome unfavorable outcome at 6-months post admission

<table>
<thead>
<tr>
<th>LR</th>
<th>CART</th>
<th>ANN</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Variable (normalized importance %)</strong></td>
<td><strong>Variable (normalized importance %)</strong></td>
<td><strong>Variable (normalized importance %)</strong></td>
</tr>
<tr>
<td>1. Age (100.0)</td>
<td>1. Motor score (100.0)</td>
<td>1. Age (100.0)</td>
</tr>
<tr>
<td>2. Motor Score (100.0)</td>
<td>2. Hypotension (100)</td>
<td>2. Motor score (54.5)</td>
</tr>
<tr>
<td>3. Hypotension (100.0)</td>
<td>3. Pupillary reactivity (66.8)</td>
<td>3. Hypotension (51.5)</td>
</tr>
<tr>
<td>4. Pupillary Reactivity (99.5)</td>
<td>4. Limb movement (33.5)</td>
<td>4. Pupillary reactivity (41.3)</td>
</tr>
<tr>
<td>5. Limb Movement (80.5)</td>
<td>5. Basal cistern effaced (21.3)</td>
<td>5. DBA (25.9)</td>
</tr>
<tr>
<td>6. Basal cistern effaced (69.5)</td>
<td>6. Age (17.5)</td>
<td>6. tSAH/IVH (16.9)</td>
</tr>
<tr>
<td>7. tSAH/IVH (57.0)</td>
<td>7. Midline shift (14.8)</td>
<td>7. Degree of MLS (15.8)</td>
</tr>
<tr>
<td>8. Sex (51.0)</td>
<td>8. DBA (3.3)</td>
<td>8. Sex (15.8)</td>
</tr>
<tr>
<td>9. EDH (40.0)</td>
<td>9. tSAH/IVH (2.5)</td>
<td>9. EDH (13.2)</td>
</tr>
<tr>
<td>10. SDH (20.0)</td>
<td>10. MEA (1.9)</td>
<td>10. Basal cistern effaced (11.5)</td>
</tr>
<tr>
<td>11. Midline Shift (18.5)</td>
<td>11. Cause of injury (1.0)</td>
<td>11. Limb movement (11.0)</td>
</tr>
<tr>
<td>12. MEI (9.5)</td>
<td>12. SDH (0.3)</td>
<td>12. MEI (5.9)</td>
</tr>
<tr>
<td>13. EDH (0.0)</td>
<td>13. EDH (0.0)</td>
<td>13. SDH (4.9)</td>
</tr>
<tr>
<td>14. Sex (0.0)</td>
<td>14. Sex (0.0)</td>
<td>14. Cause of injury (4.1)</td>
</tr>
</tbody>
</table>

DBA: Duration before admission (hrs); MEI: Major extracranial injury; EDH: Epidural hematoma; SDH: Subdural hematoma
Contents of Today’s Presentation

- Introduction
- Motivation
- Methods
- Results
- Summary
Our study is the first of its kind in India, based on the largest study and sample size in the country, to development and validation of models for TBI patients.

Our models performance is good and these models are generalizable for predicting outcomes in new patients of moderate and severe TBI.

Logistic Regression and Artificial Neural Network seem to be having similar performance, but both these methods outperformed classification and regression tree (CART) method in terms of discrimination, calibration, and overall performances.

We recommend for the use of these models in predicting outcomes for severe and moderate TBI patients in India as well as other similar countries like India.
Thank you

Special thanks to ISCB for providing conference award for developing nations to present this work.