Original Article

Forecasting Psychological Symptoms in Mild Traumatic Brain Injury: A Nonparametric Approach

Arash Nademi [1], Elham Shafiei [2, *] and Reza Valizadeh [3]

[1]Department of Statistics, Ilam Branch, Islamic Azad University, Ilam, Iran
[2]Department of Psychiatry, Ilam University of Medical Sciences, Ilam, Iran
[3]Department of Psychiatry, Psychosocial Injuries Research Center, Ilam University of Medical Sciences, Ilam, Iran

[*] Corresponding author: Clinical Reasearch Development Unit, Mostafa Khomeini Hospital, Ilam University of Medical Science, Ilam, Iran. Tel/Fax: +98-8433330110, Email: eshafiei1524@gmail.com

Abstract

Background: Nowadays, forecasting models based on nonparametric models have been developed in many branches of science such as mathematics and economics. However, the relatively complicated structure of these models has made them less practical in medical sciences.

Objectives: In this study, we investigated the application of a nonparametric regression model to predict the psychological symptoms appearing six months after mild traumatic brain injury. We also made a comparison between the performance of nonparametric regression and artificial neural network models in predicting the psychological symptoms.

Methods: In a six-month period during 2015 - 2017, information of 100 mild traumatic patients was included in a prospective cohort study. The data were then divided randomly into two groups of training (n = 50) and testing (n = 50) for 100 times. In the training group, the focus was on 100 artificial neural network and nonparametric regression models. However, in the testing group, a comparison was made between the values obtained using the two final models. To compare the models, the ROC curve and the accuracy rate were finally applied.

Results: According to the obtained results, the nonparametric regression model showed an accuracy rate of 91.25% while the neural network model had an accuracy rate of 85.34%. In the experimental set for both neural network and nonparametric regression models, the area under the ROC curve appeared to be 81.51% and 85.73 %, respectively.

Conclusions: The nonparametric regression model appeared to be more powerful than the neural network model in predicting psychological symptoms.

Keywords: Neural Network, Statistics, Nonparametric, Forecasting, Brain Concussion, Behavioral Symptoms

1. Background

In recent years, there have been significant improvements worldwide in head trauma surgery, hospital services, and trauma care and recovery. However, mild traumatic brain injury (MTBI) patients are mostly faced with various problems related to their physical, psychological, and neuronal function. They may also experience psychological symptoms for weeks just after the injury (1-3). Although these problems can result in chronic disabilities, they are not paid enough attention and consequently receive no appropriate treatment, which may be due to the complexity of their mental, biological, and social status. Exploring TBI through an organized method is most difficult because almost odd clinical manifestations are created as a result of the distribution of brain injury (4). Accordingly, the majority of the experiments have focused on

exploring the nature and the effect of consequent physical problems following MTBI. However, exploring the expected psychological symptoms following TBI is not completely achieved yet. Therefore, developing statistical models to predict these psychological symptoms is of importance.

Determining the relationships between variables along with defining effective variables and predictors are the basic mission of modeling (5). Human health is the most concerned subject in the medical and epidemiological research (6) and the methods of modeling and prediction using parametric models seem to be limited and unpractical (7). In logistic regression, the assumption of independence of errors and variables is very vital (8). In the case of data complexity, the model's suppositions may be consequently far from accuracy. The prior knowledge of the functional form relationship makes the parametric

models' approach. According to this approach and in the case of knowledge accuracy, most data sets are modeled well by a parametric method. However, in the case of the selection of wrong functional forms, a large bias will be created as compared to competitive models (9). Thus, it seems that it is valuable to use methods that are capable of prediction with the lowest error and highest confidence. Nonparametric models have been among the best methods of prediction in recent decades (10). Unlike parametric models, nonparametric models make only mild assumptions about the data and are appropriate when there is no assumption about the distribution of data.

Over the last decade, increasing attention has been devoted to nonparametric models as a new technique for estimation and forecasting in different branches of science (11, 12). Nonparametric model analysis relaxes the restricted assumptions in the parametric models and enables one to explore the data more flexibly (13, 14).

Among the popular parametric models, the artificial neural networks (ANNs) models have been used frequently in predicting variables. These models are used for various purposes such as anticipation of cancer (15), prediction of mortality after gastric surgery (16), etc. The ANNs are defined as adaptive models that analyze data and they are inspired by the human brain's functioning processes (17).

2. Objectives

In this study, we aimed to predict the psychological symptoms after MTBI by nonparametric regression and neural network models.

3. Materials and Methods

3.1. Patients and Study Design

This prospective cohort study was performed on 100 MTBI patients (GCS: 13 to 15) who were hospitalized in a neurosurgery unit. The sample was composed of male and female patients aged 15 to 65 years. Before the study, we received approval from the Ethics Committee of the hospital and informed consent from the patients. Patients with a history of spinal cord injury, communicational problems, psychotic diseases, and mental retardation and those who were unwilling to cooperate were all excluded from the study.

In the first phase, the hospitalized patients were examined by a neurosurgeon. Then, if they were eligible for the study based on the defined criteria, their demographic and clinical data were recorded in questionnaires. Six months later, they were asked to attend a center for a consultation, neurological evaluation, and completing the brief

symptom inventory (BSI) questionnaire. For consistency in completing the tests, the participants were asked one by one and then an expert (MA in psychology) recorded their oral responses in the questionnaires. The psychologist was blinded to the results of neurological evaluation, organic brain pathology, and psychological assessment. This could reduce any outcome assessment bias or diagnostic suspicion bias. For those patients who passed during the follow-up period of six months and they still had not referred to any reason, during two week have been called. To increase the willingness of patients for participation in the study, they were told that the examinations would be done by a neurologist for free.

3.2. Sample Size

We selected the sample size based on the methodology of Sedgwick et al. (18). They noted a 95% confidence interval of the population prevalence that consists of the interval of 15% to 25%. The applied sample size was based on almost a 20% prevalence in patients. Therefore, because the patients' population was almost 550, we used 110 (20%) patients as the sample size. However, 10 patients were lost to follow-up. Finally, we had 100 patients (18).

In this study, 14 variables including age, sex, education level, marital and financial status, smoking habits, mental illness history in the family, being hospitalized in neurosurgery unit, history of trauma, underlying disease, psychological drug usage, alcoholism, anesthesia, and substance use were used to compare the performance of the two models in predicting psychological symptoms.

We used the BSI questionnaire, which is the short form of the SCI-90-R questionnaire. This questionnaire has 53 questions for evaluating the psychological symptoms and separating healthy individuals from others. Dragutis et al. first introduced this questionnaire in 1973. This questionnaire consists of nine different dimensions including psychoticism, paranoia, interpersonal sensitivity, somatization, depression, anxiety, obsessive-compulsive disorder, anger, and phobia. The measure has shown to be reliable and valid and has been used in research for examining MTBI. Internal consistency was proven by Cronbach's alpha coefficient of 0.86 (19).

When the questionnaires were completed, the score of each questionnaire was calculated. If the value of T-score was above 60 in any subscale, it was considered as a patient.

3.3. Statistical Analysis

3.3.1. ANN Modeling

MATLAB is the most popular software for developing ANNs. This supervised multilayer perceptron is formed

by input, hidden, and output layers. The nodes' number of input and output layers is based on the data structure. Nonetheless, the vital step in the architecture of the neural network is to find the optimum number of hidden nodes. The training set is to learn the patterns offered in the data and the validation set is responsible for evaluating the over-fitting. Thus, in the first part, the data were divided randomly into two sets of training (n = 50) and testing (n = 50). This means that this procedure was repeated 100 times. Through changing the number of the nodes in the hidden layer, we were to train a diversity of networks. Also, in the testing step, we used the root mean square (RMS) for comparing the performance of the network. Our architecture was fitted with 100 ANN models with 8 to 13 hidden nodes and at the end, we selected the network with minimum RMS. The back-propagation algorithm was then applied for training the networks in which a learning rate and momentum value in the range of 0.1 - 0.5 and 0.1 - 0.3 were considered, respectively, on the training set.

3.3.2. Nonparametric Regression Modeling

The nonparametric regression modeling is defined as a set of regression analysis resulting in the predictor, which is based on the information taken from the data. Therefore, it is not given a predetermined form. Unlike regression based on parametric models, the nonparametric regression involves larger sample sizes because both the model structure and the model estimates must be supplied by the data. The selection of nonparametric estimators is very important in fitting the data in the nonparametric regression for estimating the parameters. In this paper, in order to analyze the nonparametric regression, we applied one of the most popular nonparametric estimators called Nadaraya-Watson (NW) kernel estimate. According to NW kernel, the dependent variable is estimated from a limited set of data points by combining the data points' locations with a kernel function.

The procedure of splitting the data was just like that in the ANN model. We tried to fit 100 nonparametric regression models to the data. The MATLAB and SPSS software were also used for modeling the dependent variable. Then, the ROC curve (20) and RMS criterion for the testing set were both calculated so as to choose the best nonparametric regression model.

4. Results

The modeling process was done in the ANN by means of the training data set. We fitted and applied different three-layer ANN models on six structures with 813 and 10 neurons in the middle and hidden layers, respectively. Also,

in the same case, we faced a momentum of 0.8, the learning rate of 0.04, and back-propagation learning algorithm, which was the best model for data prediction with an RMS of 0.1104. The accuracy rate of the model was 85.34% (Table 1).

Table 1. The Best Neural Network Model for Observations

Architecture (Input/Middle/Output)	RMS	Accuracy Rate
(14/8/1)	0.1216	0.802
(14/9/1)	0.1181	0.816
(14/10/1)[a]	0.1104	0.853
(14/11/1)	0.1154	0.822
(14/12/1)	0.1246	0.801
(14/13/1)	0.1337	0.795

[a]The most appropriate architecture after network training based on RMS error and accuracy rate.

Afterward, the nonparametric regression model was designed and fitted to the data and then the most suitable model was selected. The RMS criterion was 0.1003 for the best model.

The area under the ROC curve is used as one of the diagnostic criteria to compare the models. The values of 0 to 0.5 in the model convey a classification that has been made randomly while the total diagnosis capacity of the model is confirmed if the values are 0.5 to 1. In the experimental set, the areas under the ROC curve were 81.51% and 85.73% (Table 2) for neural network and nonparametric regression models, respectively.

Table 2. Results of Comparing 100 Pairs of Nonparametric Regression and Neural Network Models

Index	Neural Network	Nonparametric Regression
Area under the ROC curve	0.815 (0.781 - 0.886)[a]	0.857 (0.798 - 0.937)
Accuracy rate	85.34 (78.12 - 89.62)	91.25 (88.31 - 93.01)

[a]95% Confidence Interval.

The ROC curve for both models is illustrated in Figure 1. The Classification Accuracy Rate is another index for fitness that is defined as the ratio of cases classified correctly in each group. The ability for recognizing the models is shown by values of 0.5 - 1. As indicated in the above Table 2, the proper classification of the nonparametric regression model compared to the neural network model is guaranteed by the correct predicted proportion. Based on Table 2, the accuracy rate for the ultimate nonparametric regression model was 91.25% while it was 85.34% for the final neural network model.
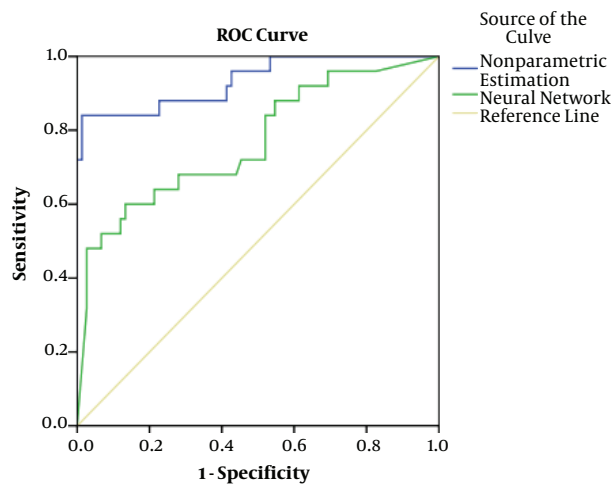
Figure 1. The ROC curve according to the final nonparametric regression and neural network models

5. Discussion

In the present study, we applied the nonparametric regression model based on the kernel estimation to predict the psychological symptoms during six months after MTBI. Also, we investigated the prediction accuracy of the nonparametric regression and neural network models for the first time. Our results showed that the nonparametric regression works better than its conventional alternative. Studies conducted on the statistical modeling of post-traumatic psychological symptoms were mostly to examine the effective factors on the psychological symptoms based on logistic regression (19, 21-24).

Most studies conducted on predicting the variables in medical sciences applied the neural network models. However, the practical application of ANNs seems to be faced with some problems. It is essential to know that designing a network is really hard and the related fundamental theory should be taken into account. Second, testing the relevance of input variables and selecting variables in non-linear methods just according to some fixed formal techniques are totally impossible (25). Third, as compared to conventional models, no etiologic explanation may be made for the weights calculated in the network (26). No mathematical relationship is firmed between the input variables and the target. Finally, this is obvious that learning ANN is really time-consuming computationally, which demands sophisticated software.

The nonparametric regression formulated function is not restricted and can focus and target at complicated patterns more flexibly than the neural network. For instance, when little information of the actual connection is available, pattern recognition is more strengthened by this spe-

cific (27). Moreover, as mentioned before, the nonparametric regression formulated function is not restricted and it is more workable and stable than the neural network to simulate complicated patterns. Another characteristic of nonparametric regression models is that they are able to find patterns even in the case of missing data (28). Incomplete noisy patterns are very powerfully searched by these models.

On the whole, when complex dependencies and interactions are concerned in the dataset, nonparametric regression models seem to be the best choice. On the contrary, since causal conclusion among variables is the basic mission of modeling and with regard to the determination of the effect of each variable on the response variable, it seems that the classic models such as logistic regression are useful enough.

In conclusion, the nonparametric regression model can be used appropriately for the psychological symptoms' prediction in trauma patients. Thus, the classification of trauma patients is possible through these predictions, which will contribute to recognizing and spotting health care resources required just based on patients' needs. Furthermore, since the linear and logistic methods are mostly used in the majority of predictive models for analysis, employing nonparametric regression models can help in planning and promoting programs that are more effective for screening individuals at risk of mental disorders.

Because there is some evidence about the impact of traumatic brain injury on the psychological symptoms of patients over time, we focused on the period of six months after trauma. One should note that there is no definitive period that can truly forecast this impact of brain injury and this is an open question that must be dealt with in future research. Our sample size was small such that we applied only 100 patients for our study. Though this sample size covers some scientific criteria, it would be better to select larger samples for getting a certain result.

Footnotes

Authors' Contribution: Arash Nademi developed and designed the study, collected the clinical data, performed the statistical analysis and interpretation of the data, and drafted the manuscript. Elham Shafei and Reza Valizadeh participated in the study conception and design, supervised the study, and revised the manuscript critically for

important intellectual contents. All authors read and approved the final manuscript.

Clinical Trial Registration Code: It is not declared.

Conflict of Interests: The authors declare no conflict of interest.

Ethical Approval: This study was approved by the Research Ethics Committee of Ilam University of Medical Sciences. The participants were assured of the confidentiality of their personal information.

Funding/Support: This study was supported by the Deputy of Research of Ilam University of Medical Sciences.

References

1. Lin MR, Chiu WT, Chen YJ, Yu WY, Huang SJ, Tsai MD. Longitudinal changes in the health-related quality of life during the first year after traumatic brain injury. *Arch Phys Med Rehabil*. 2010;**91**(3):474–80. doi: 10.1016/j.apmr.2009.10.031. [PubMed: 20298842].

2. De Silva MJ, Roberts I, Perel P, Edwards P, Kenward MG, Fernandes J, et al. Patient outcome after traumatic brain injury in high-, middle- and low-income countries: analysis of data on 8927 patients in 46 countries. *Int J Epidemiol*. 2009;**38**(2):452–8. doi: 10.1093/ije/dyn189. [PubMed: 18782898].

3. Garber BG, Rusu C, Zamorski MA. Deployment-related mild traumatic brain injury, mental health problems, and post-concussive symptoms in Canadian Armed Forces personnel. *BMC Psychiatry*. 2014;**14**:325. doi: 10.1186/s12888-014-0325-5. [PubMed: 25410348]. [PubMed Central: PMC4243369].

4. Draper K, Ponsford J, Schonberger M. Psychosocial and emotional outcomes 10 years following traumatic brain injury. *J Head Trauma Rehabil*. 2007;**22**(5):278–87. doi: 10.1097/01.HTR.0000290972.63753.a7. [PubMed: 17878769].

5. Bashiri M, Farshbaf Geranmayeh A. Tuning the parameters of an artificial neural network using central composite design and genetic algorithm. *Scientia Iranica*. 2011;**18**(6):1600–8. doi: 10.1016/j.scient.2011.08.031.

6. Wyatt JC, Altman DG. Commentary: Prognostic models: Clinically useful or quickly forgotten? *Bmj*. 1995;**311**(7019):1539–41. doi: 10.1136/bmj.311.7019.1539.

7. Buchner A, May M, Burger M, Bolenz C, Herrmann E, Fritsche HM, et al. Prediction of outcome in patients with urothelial carcinoma of the bladder following radical cystectomy using artificial neural networks. *Eur J Surg Oncol*. 2013;**39**(4):372–9. doi: 10.1016/j.ejso.2013.02.009. [PubMed: 23465180].

8. Agatonovic-Kustrin S, Beresford R. Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *J Pharm Biomed Anal*. 2000;**22**(5):717–27. doi: 10.1016/s0731-7085(99)00272-1. [PubMed: 10815714].

9. Huang M, Li R, Wang S. Nonparametric mixture of regression models. *J Am Stat Assoc*. 2013;**108**(503). doi: 10.1080/01621459.2013.772897. [PubMed: 24363475]. [PubMed Central: PMC3865811].

10. Cai TT, Low M, Ma Z. Adaptive confidence bands for nonparametric regression functions. *J Am Stat Assoc*. 2014;**109**(507):1054–70. doi: 10.1080/01621459.2013.879260. [PubMed: 26269661]. [PubMed Central: PMC4531001].

11. Li Y, Graubard BI, Korn EL. Application of nonparametric quantile regression to body mass index percentile curves from survey data. *Stat Med*. 2010;**29**(5):558–72. doi: 10.1002/sim.3810. [PubMed: 20013898]. [PubMed Central: PMC2822015].

12. Lin CY, Bondell H, Zhang HH, Zou H. Variable selection for nonparametric quantile regression via smoothing spline ANOVA. *Stat.*

13. 2013;**2**(1):255–68. doi: 10.1002/sta4.33. [PubMed: 24554792]. [PubMed Central: PMC3926212].

13. Ma L, Yan X. Examining the nonparametric effect of drivers' age in rear-end accidents through an additive logistic regression model. *Accid Anal Prev*. 2014;**67**:129–36. doi: 10.1016/j.aap.2014.02.021. [PubMed: 24642249].

14. Wang XF, Ye D. On nonparametric comparison of images and regression surfaces. *J Stat Plan Inference*. 2010;**140**(10):2875–84. doi: 10.1016/j.jspi.2010.03.011. [PubMed: 20543891]. [PubMed Central: PMC2882302].

15. Lv DJ, Zhang Y, Wang XY, Guo XM, Wang CY. [Application of artificial neural network to diagnosis of prostate cancer]. *Beijing Da Xue Xue Bao Yi Xue Ban*. 2009;**41**(4):469–73. Chinese. [PubMed: 19727241].

16. Amiri Z, Mohammad K, Mahmoudi M, Parsaeian M, Zeraati H. Assessing the effect of quantitative and qualitative predictors on gastric cancer individuals survival using hierarchical artificial neural network models. *Iran Red Crescent Med J*. 2013;**15**(1):42–8. doi: 10.5812/ircmj.4122. [PubMed: 23486933]. [PubMed Central: PMC3589778].

17. Andersson B, Andersson R, Ohlsson M, Nilsson J. Prediction of severe acute pancreatitis at admission to hospital using artificial neural networks. *Pancreatology*. 2011;**11**(3):328–35. doi: 10.1159/000327903. [PubMed: 21757970].

18. Sedgwick P. Sample size: How many participants are needed in a cohort study? *BMJ*. 2014;**349**:g6557. doi: 10.1136/bmj.g6557. [PubMed: 25361576].

19. Fakharian E, Omidi A, Shafiei E, Nademi A. Mental health status of patients with mild traumatic brain injury admitted to Shahid Beheshti Hospital of Kashan, Iran. *Arch Trauma Res*. 2015;**4**(1). e17629. doi: 10.5812/atr.17629. [PubMed: 25866741]. [PubMed Central: PMC4388991].

20. Liao P, Wu H, Yu T. ROC curve analysis in the presence of imperfect reference standards. *Stat Biosci*. 2017;**9**(1):91–104. doi: 10.1007/s12561-016-9159-7. [PubMed: 28694878]. [PubMed Central: PMC5501420].

21. Vassallo JL, Proctor-Weber Z, Lebowitz BK, Curtiss G, Vanderploeg RD. Psychiatric risk factors for traumatic brain injury. *Brain Inj*. 2007;**21**(6):567–73. doi: 10.1080/02699050701426832. [PubMed: 17577707].

22. Chong SL, Liu N, Barbier S, Ong ME. Predictive modeling in pediatric traumatic brain injury using machine learning. *BMC Med Res Methodol*. 2015;**15**:22. doi: 10.1186/s12874-015-0015-0. [PubMed: 25886156]. [PubMed Central: PMC4374377].

23. Booth-Kewley S, Schmied EA, Highfill-McRoy RM, Larson GE, Garland CF, Ziajko LA. Predictors of psychiatric disorders in combat veterans. *BMC Psychiatry*. 2013;**13**:130. doi: 10.1186/1471-244X-13-130. [PubMed: 23651663]. [PubMed Central: PMC3651311].

24. Perron BE, Howard MO. Prevalence and correlates of traumatic brain injury among delinquent youths. *Crim Behav Ment Health*. 2008;**18**(4):243–55. doi: 10.1002/cbm.702. [PubMed: 18803295]. [PubMed Central: PMC4112384].

25. Chokmani K, Ouarda TBMJ, Hamilton S, Ghedira MH, Gingras H. Comparison of ice-affected streamflow estimates computed using artificial neural networks and multiple regression techniques. *J Hydrol*. 2008;**349**(3-4):383–96. doi: 10.1016/j.jhydrol.2007.11.024.

26. Seixas JM, Faria J, Souza Filho JB, Vieira AF, Kritski A, Trajman A. Artificial neural network models to support the diagnosis of pleural tuberculosis in adult patients. *Int J Tuberc Lung Dis*. 2013;**17**(5):682–6. doi: 10.5588/ijtld.12.0829. [PubMed: 23575336].

27. Awate SP, Whitaker RT. Multiatlas segmentation as nonparametric regression. *IEEE Trans Med Imaging*. 2014;**33**(9):1803–17. doi: 10.1109/TMI.2014.2321281. [PubMed: 24802528]. [PubMed Central: PMC4440593].

28. Zhu W, Zhang H. A nonparametric regression method for multiple longitudinal phenotypes using multivariate adaptive splines. *Front Math China*. 2013;**8**(3):731–43. doi: 10.1007/s11464-012-0256-8. [PubMed: 25309585]. [PubMed Central: PMC4193387].