Artificial neural network in predicting craniocervical junction injury: an alternative approach to trauma patients
Fırat Bektaş, Cenker Eken, Secgin Soyuncu, İsa Kilicaslan and Yıldiray Cete

Objective The aim of this study is to determine the efficiency of artificial intelligence in detecting craniocervical junction injuries by using an artificial neural network (ANN) that may be applicable in future studies of different traumatic injuries.

Materials and methods Major head trauma patients with Glasgow Coma Scale ≤8 of all age groups who presented to the Emergency Department were included in the study. All patients underwent brain computerized tomography (CT), craniocervical junction CT, and cervical plain radiography. A feedforward with back propagation ANN and a stepwise forward logistic regression were performed to test the performances of all models.

Results A total of 127 patients fulfilling inclusion criteria were included in the study. The mean age of the study patients was 31 ± 17.7, 77.2% (n=98) of them were male, 13.4% of the patients (n=17) had craniocervical junction pathologies. About 64.7% (n=11) of these pathologies were detected only by CT; 23.5% (n=4) of them by both craniocervical CT and cervical plain radiography; and 11.8% (n=2) of them only by cervical plain radiography. A logistic regression model had a sensitivity of 11.8% and specificity of 99.1%. Positive predictive value was 66.7% and negative predictive value was 87.9%. Area under the curve for logistic regression model was 0.794 (P=0.000). ANN had a sensitivity of 82.4% and specificity of 100%. Positive predictive value was 100% and negative predictive value was 97.3%. Area under the curve for ANN model was 0.912 (P=0.000).

Conclusion ANN as an artificial intelligence application seems appropriate for detecting and excluding craniocervical junction injury but it should not replace craniocervical junction CT. However, these findings should lead us to test the applicability of ANN on hard-to-diagnose trauma patients or in constructing clinical decision rules. European Journal of Emergency Medicine 15:318–323 © 2008 Wolters Kluwer Health | Lippincott Williams & Wilkins.

Keywords: artificial neural network, craniocervical junction injury, trauma

Introduction It is essential to diagnose cervical spine injuries on initial assessment of patients who present to the Emergency Department (ED) after blunt trauma. Patients who are conscious, who do not have neck pain, cervical spine tenderness, who are not intoxicated, or have other distracting painful injuries, are at low risk for craniocervical junction injury (CCJI) [1]. However, it is rather difficult to rule out the possibility of cervical spine injury, especially in the upper cervical region, in patients who are unconscious or intubated [2–4]. Delays in diagnosing injuries of the upper cervical region (C1–C2) may cause permanent neurological deficits [5–8]. The probability of CCJI was 2–6% among patients with blunt trauma whereas it varies between 7 and 20% in patients with major traumatic brain injury [9,10]. The rate of delays in diagnosing the cervical injuries has been reported to be between 4.2 and 22.9% [11,12]. The cause of those delays was most frequently the wrong radiological interpretation owing to lack of physician experience, patient’s mental status changes [13], and low quality of the radiographs [14]. Therefore, some trauma centers use craniocervical junction computerized tomography (CT) as a routine test, especially in unconscious major trauma patients [5].

Artificial neural network (ANN) is a model for the application of artificial intelligence. ANN is an information-processing tool that is inspired by the structure and function of the human brain. The human central nervous system is constituted of neurons connected to each other, separated by synapses, and achieves information transfer via a series of action potentials [15]. By using these connections, neurons receive energy and after summing these energies, they send it to other neurons if the sum reaches a critical threshold. The brain learns by adjusting the number and strength of these connections. McCulloch and Pitts [16] first described ANNs as a method of information processing using a network of binary decision elements or ‘neurons’. Later, efforts were made to explain complex processes of the central nervous system [17]. A neural network is composed of a series of interconnecting
parallel nonlinear processing elements (nodes) with a limited number of inputs and outputs. The principle of ANN with a supervised learning algorithm can be described as a training process in which these neurons change their connection weights until the error between the predicted and actual outputs decreases to an acceptable level.

Although ANN has been applied to many medical problems, thus far it has not been used in any kind of trauma patients. The goal of this study is to determine the efficiency of artificial intelligence in detecting CCJIs by using ANN and thus serving as a stimulus for further studies of ANN as a diagnostic tool in trauma.

Materials and methods
This was a randomized, prospective, observational, and clinical study in a university hospital ED. The ED serves a population of approximately 1000000. The annual census is 50000. Major head trauma patients of all age groups presenting to the ED with Glasgow Coma Scale (GCS) ≤ 8 were included in the study. Patients with traumatic arrest, pregnant patients, hemodynamically unstable patients and those requiring immediate surgery, penetrating head trauma patients, and patients whose CT could not be taken were excluded from the study. The dataset consists of the following parameters:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>Coded (2 codes, i.e. male, female)</td>
</tr>
<tr>
<td>Age</td>
<td>Coded (2 codes, ≥ 65 or not)</td>
</tr>
<tr>
<td>Mean arterial pressure</td>
<td>Coded (as a continuous variable)</td>
</tr>
<tr>
<td>Pulse rate</td>
<td>Coded (2 codes, ≤ 100 or not)</td>
</tr>
<tr>
<td>Respiration rate</td>
<td>Coded (2 codes, ≤ 50 or not)</td>
</tr>
<tr>
<td>Glasgow coma scale</td>
<td>Coded (as an ordinal variable)</td>
</tr>
<tr>
<td>Revised trauma score</td>
<td>Coded (as an ordinal variable)</td>
</tr>
<tr>
<td>Motor vehicle accident</td>
<td>Coded (2 coded, 1 for yes and 0)</td>
</tr>
<tr>
<td>Pedestrian struck</td>
<td>Coded (2 coded, 1 for yes and 0)</td>
</tr>
<tr>
<td>Falls</td>
<td>Coded (2 coded, 1 for yes and 0)</td>
</tr>
<tr>
<td>Motorbike accident</td>
<td>Coded (2 coded, 1 for yes and 0)</td>
</tr>
<tr>
<td>Pathology on head computed</td>
<td>Coded (2 coded, 1 for yes and 0)</td>
</tr>
<tr>
<td>tomography (all pathologies</td>
<td>including cerebral edema,</td>
</tr>
<tr>
<td>including cerebral edema,</td>
<td>contusion, subarachnoid</td>
</tr>
<tr>
<td>contusion, subarachnoid</td>
<td>bleeding, etc.)</td>
</tr>
<tr>
<td>Alcohol intoxication</td>
<td>Coded (2 coded, 1 for yes and 0)</td>
</tr>
</tbody>
</table>

The data points depicted above, the brain and upper cervical region CT results and radiograph results were recorded on a study form. After initial patient stabilization, posterior–anterior and lateral cervical spine films were taken by a portable radiograph machine. Subsequently, all patients underwent head and upper cervical spine CT with 5-mm cross-sectional axial cuts down to C-2 using a Toshiba Xpress C6GT0008A CT scanner (Toshiba, Tokyo, Japan). Adding the upper cervical CT did not result in additional costs for the patients. CT results and plain radiographs were evaluated in terms of upper cervical region pathologies by two different radiologists who were blinded to the study. The patients with upper cervical regional pathology, who were diagnosed by CT and plain radiography, were followed up for at least 1 month in terms of neurological deficit.

Statistical analysis
The data were evaluated by SPSS 13.0 for Windows (SPSS Inc., Chicago, Illinois, USA). Continuous variables were expressed as mean ± SD. Categorical variables were expressed as rate and percentage. Univariate analysis for continuous variables for the two-group comparisons was performed using the Student’s t-test. Mann–Whitney U-test was preferred for ordinal variables. The comparison of two groups composed of categorical variables was carried out by χ² analysis.

Logistic regression
The SPSS 13.0 for Windows was also used for binary logistic regression analysis. As in ANN, 13 independent variables were assigned to predict craniocervical injury as the dependent variable. A stepwise forward logistic regression analysis was performed. We tested the fitness of the logistic regression model with the Hosmer–Lemeshow goodness-of-fit statistic.

Artificial neural network
A feedforward ANN with back propagation was performed by JMP (release 6.0; SAS, Cary, North Carolina, USA). As described above ANN works similarly to the human central nervous system. It is composed of three layers: input layer, hidden layer, and output layer. As a first step ANN gives weights to the each input (collected data) in the input layer and constitutes the hidden layer, which then connects input and output layers. By weighting the data in the hidden layer, an output is obtained using a probability function. To train an ANN the weights have to be adjusted according to the error between the predicted and actual outputs. This process is performed mostly by a back-propagation algorithm. The most important problem in ANN is that the ANN formulation may be too specific to the present dataset. This phenomenon is also known as overfitting, which can also be described as a problem when ANN learns the training set too accurately and yet cannot generalize when presented with a new test set. To prevent overfitting in a large dataset, it is divided into a training set in which ANN learns and a test set in which the performance of ANN is checked. However, for a small dataset a K-fold cross-validation model is suggested to avoid overfitting. The K-fold cross-validation method separates the data into K sets and assigning one of the K subsamples as the test set and the remaining (K–1 subsamples) as the training set. The cross-validation process is then repeated K times (the folds) with each of the K subsamples used exactly once as validation data. The K results from the folds then can be averaged to produce a single estimation. We performed six-fold
cross validation with two hidden units, 500 iterations and 20 tours. The overfit penalty was assigned as 0.001 and convergence criterion was chosen as 0.00001.

To compare the overall performances of all applications; sensitivity, specificity, positive predictive value, and negative predictive value of all applications were determined. In addition, a receiver operating characteristics (ROC) curve analysis was performed. \( P \geq 0.05 \) was accepted as significant.

**Results**

A total of 215 patients with multiple injuries and a GCS of 8 or lower were brought to the ED. About 127 of these patients fit our inclusion criteria and were entered into the study. The patient flowchart is shown in Fig. 1.

The mean age of the study patients was 31 ± 17.7, 77.2\% \((n = 98)\) of them were male. Of these patients about 40.2\% \((n = 51)\) were injured in motor vehicle accidents, 37.8\% of them \((n = 48)\) were pedestrians struck by a motor vehicle, 11\% of them \((n = 14)\) fell from a height, 9.4\% \((n = 12)\) were involved in motorcycle accidents, 1.6\% of the patients had other trauma etiologies. In all, 13.4\% of the patients \((n = 17)\) had craniocervical junction pathologies; 64.7\% \((n = 11)\) of these pathologies were detected only by CT, 23.5\% \((n = 4)\) of them by both craniocervical CT and cervical plain radiography, and 11.8\% \((n = 2)\) of them only by cervical plain radiography. No statistically significant variable was observed in determining CCJI in the univariate analysis. Table 1 depicts the univariate comparison of two groups.

**Logistic regression**

With a forward stepwise conditional method, a logistic regression model with 13 independent variables yielded two of 17 patients with craniocervical injury and 109 of 110 patients without injury. This translates to a sensitivity of 11.8\% and specificity of 99.1\%. Positive predictive value was 66.7\% and negative predictive value was 87.9\%. Area under the curve for logistic regression model was 0.794 \((P = 0.000)\). However, there was no significant independent variable predicting CCJI in the logistic regression model.

**Artificial neural network**

ANN identified 14 of 17 patients with craniocervical junction injury and all of the patients without injury. It had a sensitivity of 82.4\% and specificity of 100\%. Positive predictive value was 100\% and negative predictive value was 97.3\%. Area under the curve for the ANN model was 0.912 \((P = 0.000)\). Figure 2 shows the ANN diagram.

Table 2 shows the performances of two models and Fig. 3 depicts the comparison of logistic regression and ANN models with ROC analysis.

The frequency and the type of the fractures in the craniocervical region are shown in Table 3.

Various head CT abnormalities were detected in 109 patients (85.2\%). In contrast, 15 of 17 (88.2\%) patients with craniocervical junction pathology had additional pathologic findings on their head CT. The other two (11.8\%) patients had a normal head CT. The head CT findings of these patients are shown in Table 4.

Six of the 17 patients with craniocervical pathology died within 1 month. The mortality rate in this study was 35\%.

**Discussion**

ANN is an application of artificial intelligence such as genetic algorithm, fuzzy logic, etc. Although it enjoys widespread use in other industries, its application in...
medicine, thus far, has been limited. However, recently there has been an increasing trend in its utilization in different areas of medical science. Logistic regression has been the most popular analysis in medicine for detecting the hidden relationship between multiple variables. Clinical decision rules are usually formed using logistic regression, but ANN as a data mining technique might be a good alternative to logistic regression. Thus far, there have not been any studies in the medical literature using ANN in trauma care.

Cervical spine findings occur more often in major trauma patients who present to the ED unconscious or intubated [9,10]. It is especially difficult to exclude pathologies in the upper cervical anatomic regions [18]. For the initial assessment of the cervical spine of these patients, cervical plain radiographies are used [19]. Generally, if any fracture, dislocation, or suspected pathology is detected on plain film a cervical CT is performed [20]. Two main factors that may contribute in detecting the presence of cervical spine injuries on plain radiographs are: (i) the quality of the film and (ii) the experience of the physician who interprets the cervical plain radiograph [11,12]. Retrospective studies have shown that because of those factors the percentage of studies initially read as false negative is 23–33% [21,22]. It was a surprising result for us to find a much higher rate of 64.7% in our study. However, none of the patients with missed results developed neurological deficits of consequence at 1-week follow-up in the ICU.

An early and correct diagnosis of cervical spine injuries is very important to prevent neurological deficits that may develop later. The presence of cervical pathology was shown by CT only in 11 of 17 patients with upper cervical region pathologies, whereas simple plain radiography detected only two of 17 injuries. If a decision for these patients were made based purely on plain radiography, additional neurological pathologies would develop and the mortality and morbidity of these patients would have increased.
consequently increase. The logistic regression model identified only two of 17 patients but 109 of 110 patients without CCJI. It was thus more successful in excluding the patients without injury. Nonetheless, its negative predictive value was not powerful enough to exclude all patients (87.9%). ANN revealed 14 of 17 patients with CCJI but there was no patient without CCJI that ANN assigned as having injury. ANN only assigned three patients falsely with CCJI as not having CCJI. It has a negative predictive value of 97.3% and positive predictive value of 100% In addition, it also detected the two patients with CCJI that cranial CT did not.

Simple plain radiography may not be sufficient for the diagnosis of upper cervical region pathologies in high-risk patients who present to the ED with GCS ≤ 8. It may be necessary to evaluate these patients with a craniocervical junction CT for upper cervical region fractures, as a quarter of these pathologies are missed by plain radiography. Although we did not find any correlation between intracranial pathology and CCJI, upper cervical CT is strongly recommended for patients with traumatic intracranial bleeding, as craniocervical region pathologies often coexist with bleeding [23]. Similarly, there was no correlation between the trauma mechanism, demographical features of the patients, GCS score, revised trauma score, and craniocervical fractures in this study.

Despite CT’s superiority to plain radiography in diagnosing cervical spine fracture, it may be insufficient in diagnosing craniocervical region pathologies. In this study, the pathologies observed on plain radiography of two patients were interpreted as normal on CT by two radiologists. No further diagnostic methods could be applied as these patients deteriorated and died shortly after. In the study performed by Pech et al. [24], it has been demonstrated that pedicle and lateral condyle fractures are often missed when the clinical status of the patient was not known. Moreover, another disadvantage of CT is its inability to detect ligamentous injuries in the craniocervical region, as it cannot display the soft tissue properly [25].

The findings of this study demonstrate that ANN strongly excluded and detected patients with and without CCJI, but this does not mean that it is able to replace the craniocervical junction CT in trauma patients. Craniocervical junction CT should be obtained in all high-risk patients. However, these findings may be a source of inspiration for use of artificial intelligence models in developing clinical decision rules in difficult-to-diagnose trauma patients. For instance, they could be used in the decision of obtaining cranial CT in patients with minor head trauma or in selecting those patients required to take plain cervical graphs in blunt neck trauma. However, these are needed to be validated with further studies.

One of the critical points in artificial intelligence models is the selection of the variables. One can also constitute simple prediction rules both by ANN and logistic regression applications. After analyzing the significant variables in the univariate analysis or by using the likelihood ratios, it is possible to compose a prediction model by assigning significant variables to the ANN and logistic regression analysis. Furthermore, the independent variables that are statistically significant in a multivariate regression analysis can also be used to compose a prediction model. We preferred to use all of the variables probably related to CCJI for both models to prevent a possible bias.

Limitations
Although the CT detects craniocervical region injuries at a higher proportion than conventional plain radiography, we faced some restrictions and difficulties in this study. The greatest limitation of our study is the total number

---

### Table 3 Diagnosis of the occipital condyle and cervical vertebra fractures by plain radiography and CT, or only by CT

<table>
<thead>
<tr>
<th>Fracture location (total)</th>
<th>Number of patients with fracture (n=17)</th>
<th>Number of patients with fracture diagnosed by plain radiography (n=6) (35.3%)</th>
<th>Number of patients diagnosed by CT (n=15) (88.2%)</th>
<th>Number of patients diagnosed only by CT (n=11) (64.7%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Occipital condyle</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>C 1</td>
<td>8</td>
<td>3</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>C 2</td>
<td>7</td>
<td>3</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>C 1 and C 2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

CT, computed tomography.

### Table 4 The head computed tomography findings of all patients

<table>
<thead>
<tr>
<th>Pathology</th>
<th>Patients with craniocervical fracture, n=17</th>
<th>Patients without craniocervical fracture, n=110</th>
<th>Total, n=127</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epidural hematoma</td>
<td>1 (5.9%)</td>
<td>7 (6.4%)</td>
<td>8 (6.3%)</td>
</tr>
<tr>
<td>Subdural hematoma</td>
<td>3 (17.3%)</td>
<td>19 (17.3%)</td>
<td>22 (17.3%)</td>
</tr>
<tr>
<td>Depressed skull fracture</td>
<td>0</td>
<td>17 (15.3%)</td>
<td>17 (13.4%)</td>
</tr>
<tr>
<td>Subarachnoid bleeding</td>
<td>5 (29.4%)</td>
<td>51 (46.4%)</td>
<td>57 (44.1%)</td>
</tr>
<tr>
<td>Contusion</td>
<td>5 (29.4%)</td>
<td>51 (46.4%)</td>
<td>56 (44.1%)</td>
</tr>
<tr>
<td>Brain tumor</td>
<td>11 (64.7%)</td>
<td>78 (70.9%)</td>
<td>89 (70.1%)</td>
</tr>
<tr>
<td>Intraparenchymal bleeding</td>
<td>5 (29.4%)</td>
<td>58 (52.7%)</td>
<td>63 (49.8%)</td>
</tr>
</tbody>
</table>

Copyright © Lippincott Williams & Wilkins. Unauthorized reproduction of this article is prohibited.
of patients enrolled. This number is rather low when compared with other studies in the literature on this subject.

Another limitation of this study is not performing CT by taking cross-sections of 3-mm length in the axial plane as recommended by the Eastern Union of Traumatic Surgery guide [7, 8]. In this study, craniocervical CT was performed taking a 5-mm cross-section in the axial plane. This might have been the reason for the fact that the CT of those two patients with clear fractures on plain radiograph had no fracture detected on upper cervical spine CT.

Another point to underline is the overfitting problem in ANN models. Although we performed K-fold cross validation and two hidden nodes to avoid overfitting, testing of the ANN model in an unseen data performance should be better.

**Conclusion and future work**

ANN as an artificial intelligence application is a powerful tool in detecting and excluding CCJJI but it should not replace craniocervical junction CT. Craniocervical junction CT should be applied as a routine test for the patient population with a GCS score of 8 or less. Our findings should lead us to test the performance of ANN on other areas of trauma or in constructing clinical decision rules.

**Acknowledgements**

The authors thank Dr Stephan Rinnert for editing the syntax of the article and the residents and attending physicians in the ED at Akdeniz University Hospital. This study was supported by the Akdeniz University Research Foundation.

**References**


16 McCulloch WS, Pitts WA. A logical calculus of ideas immanent in the nervous system. Bull Math Biophys 1943; 5:115–133.


