Machine Learning Algorithms and Quantitative Electroencephalography Predictors for Outcome Prediction in Traumatic Brain Injury: A Systematic Review

NOR SAFIRA ELAINA MOHD NOOR1,2 AND HAIÐI IBRAHIM1, (Senior Member, IEEE)

1School of Electrical and Electronic Engineering, Engineering Campus, Universiti Sains Malaysia, Nibong Tebal 14300, Malaysia
2Department of Neurosciences, School of Medical Sciences, Health Campus, Universiti Sains Malaysia, Kota Bharu 16150, Malaysia

Corresponding author: Haidi Ibrahim (haidi_ibrahim@ieee.org)

This work was supported in part by the Ministry of Higher Education (MoHE) Malaysia, under Trans-Disciplinary Research Grant Scheme (TRGS), under Grant 203/PELECT/6768002, and in part by the Skim Latihan Akademik Bumiputra (SLAB).

ABSTRACT Recent developments in the field of machine learning (ML) have led to a renewed interest in the use of electroencephalography (EEG) to predict the outcome after traumatic brain injury (TBI). This systematic review aims to determine how previous studies have taken into consideration the important modeling issues for quantitative EEG (qEEG) predictors in developing prognostic models. A systematic search in the PubMed and Google Scholar databases was performed to identify all predictive models for the extended Glasgow outcome scale (GOSE) and Glasgow outcome scale (GOS) based on EEG data. Fourteen studies were identified that evaluated ML algorithms using qEEG predictors to predict outcome in patients with moderate to severe TBI. In each model, a maximum of five qEEG predictors were selected to determine the association between these parameters, and favorable or unfavorable predicted outcomes. The most common ML technique used was logistic regression, but the algorithms varied depending on the types and numbers of qEEG predictors selected in each model. The qEEG variability for the relative and absolute band powers were the most common qEEG predictors included in the models (46%) followed by total EEG power of all frequency bands (31%), EEG-reactivity (31%) and coherence (15%). Model performance was often quantified by the area under the receiving-operating characteristic curve (AUROC) rather than by accuracy rate. Various ML models have demonstrated great potential, especially using qEEG predictors, to predict outcome in patients with moderate to severe TBI.

INDEX TERMS Electroencephalography, machine learning algorithms, prediction methods, reviews.

I. INTRODUCTION

Traumatic brain injury (TBI) in definition is a medical condition caused by a unique combination of external or internal mechanical forces which directly gives an impact to the brain. TBI occurs from various mechanisms including road-traffic accidents, military blast injury, blunt object, falls and sport-related concussion [1]. TBI can be classified based on its seriousness. Glasgow Coma Scale (GCS) is a neurological scale that is used to classify the severity of TBI patients based on three categories (i.e., mild (GCS 13-15), moderate (GCS 9-12) and severe (GCS < 9)) [2]–[4]. TBI results in continuous and permanent effects not only to brain structure but also effects on morbidity, memory, emotions, attention and executive functions, personality, as well as social consequences that significantly alter the life and productivity [5], [6].

In clinical practice, numerous neuroimaging techniques are extensively used in diagnosis, prognosis and patient management of TBI-related brain dysfunction. It assists in the detection of injury for neurological evaluation, in the planning of the treatment, and in the prediction of outcomes [7], [8]. Generally, structural imaging techniques (e.g., computed tomography (CT) and magnetic resonance imaging (MRI)) play a role in critical diagnosis and management, while functional imaging techniques (e.g., functional MRI (fMRI),...
In investigating the timeline of TBI-related abnormalities and future recovery, the use of CT/MRI alone has been found inadequate, relatively because the sensitivity and specificity of these techniques cannot be obtained by using only one technique. Due to this, EEG or MEG would become a better experimental technique especially for evaluating the electrophysiological changes (i.e., slowing, missing or interrupted brain circuits, alterations of surviving structures) [9], [10]. MEG and EEG are particularly non-invasive electrophysiological methods and both are able to evaluate the cerebral cortex activities with superior temporal resolution (i.e., in milliseconds). However, EEG holds high promise for prognostic evaluations due to its ability to do diagnose, monitoring, and prognosis at different stages of brain injury including in critically ill patients [11]. EEG is the electrical wave of the brain that is recorded by using noninvasive electrodes positioned on the scalp, and it is relatively cheaper, portable, and easy to handle as compared to MEG that has limitations in general routine for clinical evaluation of TBI patients [12].

Recent developments in the field of machine learning (ML) have led to a renewed interest in the use of EEG for the purpose of accurately predicting the outcome by taking into account the sensitivity of this technique [13]–[18]. ML, which is a field of computer science, falls into a subdivision of artificial intelligence (AI) that enables algorithms to learn patterns in large, complex datasets without being explicitly programmed [19], [20]. ML models are generally grouped into three categories: (i) supervised; (ii) unsupervised; and (iii) semi-supervised learning. Thus far, ML techniques have the potential to enhance the capabilities of using electrophysiological data (i.e. EEG or MEG) by allowing the diagnosis and prognosis (i.e., prediction) at the single-subject and multi-subject levels [21]. Numerous studies have attempted to reveal the efficacy of using multivariate prediction models to systematically identify the most prominent quantitative EEG (qEEG) predictors that are robust to predict TBI outcomes [18], [22]–[28]. The qEEG is a method of analyzing the brain electrophysiological data using the modern mathematical and statistical analysis (e.g., Fourier transform, wavelet analysis) on traditional EEG recordings to extract quantitative patterns that may contribute to diagnostic information and/or cognitive deficiencies.

In addition to that, the research to date has been designed to determine the ability of the current ML models with EEG data to provide reliable prediction of outcome in achieving acceptable performance measure with relation to TBI outcome. For accurate outcome prediction, multiple predictors (i.e., age, clinical variables, GCS, imaging and functional data) need to be included in a prognostic model because single factors are inaccurate to provide sufficient predictive value to discriminate patient outcomes. In the work by Jennett et al. [29] and other studies published on prediction outcomes in TBI, the Glasgow Outcome Scale (GOS) has been accepted and used over the past forty years to evaluate the TBI patient outcomes [30].

Predictive models are statistical modeling that employs the ML techniques, which combine two or more information of patient clinical data to learn certain features from the training set, finally contributing to a predictable outcome in TBI patients. For a predictive model to be clinically useful, it must fulfill two requirements: it must be clinically valid and methodologically valid [31]. This systematic review aims to determine how previous studies have taken into consideration of four important modeling issues to better understand the potential of qEEG features for developing a predictive model. The modeling issues are (1) the study population, (2) choice of qEEG predictors and outcome, (3) model development, and (4) model validation.

II. METHODOLOGY

This systematic review was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidelines [32], [33]. A comprehensive search strategy was developed using the PubMed and Google Scholar databases without applying any specific date range to identify all relevant studies up to April 30, 2020.

A. SEARCH STRATEGY

The search was performed from January to April 2020, by creating search syntax with the following keywords in the PubMed (http://www.ncbi.nlm.nih.gov/pubmed) database: “machine learning”, “artificial intelligence”, “predict”, “traumatic brain injury” and “electroencephalography”. Related articles in English language published up to April 30, 2020 were evaluated. In the PubMed, we inserted the search terms by combining the Medical Subject Headings (MeSH) to find other relevant keywords for machine learning.

Then, we inserted the truncation for the word “predict*” to search variations or synonyms on the word stem. Finally, we entered the following search terms into the PubMed using the Boolean operator OR and AND: (((“machine learning”) OR “artificial intelligence”) OR predict*) AND “traumatic brain injury”) AND “electroencephalography”). The search results were 69 articles. Another search syntax combination of “electroencephalography”, “machine learning” OR “artificial intelligence” OR “predict” “traumatic brain injury” was inserted into the Google Scholar database, generated 499 results after excluding the review papers (see Figure 1).

B. STUDY SELECTION

Articles were included in the systematic review if the following conditions were met: (a) the article was written in English language; (b) original article in a peer-reviewed journal (c) the article reported on a method of ML model and EEG data for application in TBI outcome prediction; (d) the article evaluated the performance of the ML model technique used; (e) the article described the qEEG predictor variables.
including technique applied for features extraction of data; (f) study population of the article was limited to TBI group (i.e., mild, moderate, severe).

Articles were excluded if the following criteria were met: (a) the article did not report the original contribution to ML and EEG for outcome prediction in TBI (e.g., the paper reported on the application of ML and EEG for seizures outcome prediction); (b) the full-text article was not available; (c) the conference abstracts, review paper, poster presentation, letters to the editor; (d) the article reported in languages other than English; (e) the article involved non-human study; (f) the articles reported ML for TBI outcome prediction on other neuroimaging studies (i.e. MRI, fMRI, DTI, PET, MEG, ERP). The articles were excluded as the application of ML for TBI outcome prediction using these modalities had been reviewed elsewhere [15], [19], [34]–[38]. After screening the titles and abstracts and removing duplicate articles \((n = 384)\), 46 articles were selected for full-text review. No specific data range was applied in the search. The selection of full-text articles for inclusion was conducted by two authors (i.e., N. S. E. Mohd Noor and H. Ibrahim) independently for further analysis.

C. DATA EXTRACTION AND ANALYSIS APPROACH

For this systematic review, data extraction focused on four main aspects: (i) sample size; (ii) qEEG predictors variables; (iii) predictive model development; and (iv) predictive model validation. We performed qualitative and quantitative analysis of the studies included. Data extracted from each article were: (a) year of publication; (b) TBI groups (e.g., mild, moderate and severe); (c) number of TBI patients; (d) types of ML model used (e.g., supervised, unsupervised techniques); (e) qEEG input features; (f) size of training data; (g) size of testing data; (h) validation method.
III. RESULTS

A. SEARCH RESULTS

After the elimination of duplicates, a total of 568 publications in the PubMed and Google Scholar databases were identified. Following exclusion, 46 full-text articles were assessed for eligibility by title and abstract screening. Twelve studies described the modeling technique used for prediction of outcome in TBI but did not apply ML techniques. Ten studies applied ML algorithms to compare TBI patients from healthy subjects but did not report or evaluate the outcome of recovery. Ten studies evaluated the modeling technique using ML approaches for TBI outcome prediction but EEG data was not included as a predictor of outcome. A total of 14 studies has fully fulfilled the inclusion criteria for data extraction and qualitative synthesis (please refer to Figure 1).

D. DATA EXTRACTION AND ANALYSIS APPROACHES FOR QUANTITATIVE EEG PREDICTORS AND OUTCOME

Data extraction on qEEG predictors and outcomes concentrated on the number and type of qEEG predictors in each prognostic model construction. Data extracted from each article were: (i) number and type of selected qEEG predictors; (ii) types of EEG recording; (iii) method for extracting each qEEG predictors; and (iv) time of assessment of outcome. The outcome measure used in all studies was the GOS score as tabulated in Table 1.

TABLE 1. Details of GOS according to Jennet and Bond [39].

| GOS Score | Clinical Meaning | Description | Outcome |
|-----------|------------------|-------------|---------|
| 1         | Death            | Death due to brain damage | Poor    |
| 2         | Vegetative state | Unresponsive and unable to interact with environment for weeks or months after acute brain damage | Poor    |
| 3         | Severe disability| Dependent on daily support because of physical and/or mental causes | Poor    |
| 4         | Moderate disability| Able to live independently in daily life but may rely on a special equipment for aiding in everyday activities | Poor    |
| 5         | Good recovery    | Resumption of normal life, although there may be minor neurological and physiological deficits | Good    |

B. STUDY POPULATION

From the 14 studies, 85% of the predictive models were derived from a population that combined the older adults and children populations in the age range 5.5 to 85 years old. 14% were derived from an adult population (i.e., 44 - 67 years old). In relation to the severity of the TBI studied, ten out of fourteen (i.e., 71%) predictive models included severe TBI patients while four out of fourteen (i.e., 28%) models included moderate to severe TBI groups of patients. This systematic review has included the studies that presented a predictive model for patients with moderate to severe TBI because so far, no studies have reported on mild or moderate TBI patients as a targeted group for modeling the prognostic model. The predictive model or prognostic model term was used interchangeably. Eleven of the publications developed models on relatively small sample sizes (i.e., less than 100) and the largest study included 680 of severe TBI patients [40].

C. QUANTITATIVE EEG PREDICTORS AND OUTCOME

There were multiple qEEG features selected and tested for inclusion in the predictive models (see Table 2).

TABLE 2. The qEEG predictors considered in 14 predictive models in traumatic brain injury.

| Publication(s) | qEEG Predictor(s) Group |
|----------------|-------------------------|
| Tolonen et al. [41], Hebb et al. [42], Vespa et al. [43], Mikola et al. [44], Haveman et al. [45], van der Brink et al. [46] | Variability, Relative power, Absolute band power, PAV |
| Stefan et al [47] | Approximate entropy, Permutation entropy, Weighted symbolic mutual information |
| Stefan et al [47] | Micrstate analysis |
| Hack et al [40], Mikola et al [44], Haveman et al [45], Stefan et al [47] | Total EEG power |
| Stefan et al [47], Schorr et al [48] | Coherence analysis |
| Stefan et al [47] | Transfer entropy |
| Haveman et al [45] | SEP90% |
| van den Brink et al. [46] | Amplitude |
| Chen et al [17], van den Brink et al. [46] | Connectivity, Mean clustering in δ network, Modularity coefficient in δ network |
| Logi et al [49], Amanatini et al [50], Gulling et al [51] | EEG-R |
| Lee et al [52] | N2 stage sleep phase, Posterior dominant rhythm (PDR) |

Abbreviations: PAV (Percentage of Alpha Variability); EEG-R (EEG - reactivity); SEP90% (Spectral Edge Frequency at 90%).

1) qEEG PREDICTORS

The description of each of the qEEG features derived from the fourteen studies was summarized as follows:

1) **Absolute power per band:** The absolute power of each band was calculated by integrating all the power values
within each frequency band: delta \( \delta \) (i.e., 0.5 - 4 Hz); theta \( \theta \) (i.e., 4 - 8 Hz); alpha \( \alpha \) (i.e., 8 - 13 Hz); and beta \( \beta \) (i.e., 13 - 30 Hz).

2) **EEG total power:** The EEG total power is a summation of all power bands from 0.5Hz to 30 Hz.

3) **Relative power per band:** The relative power for each band was derived by expressing the absolute power in each frequency band as a percentage of the absolute power summed across four frequency bands. For example, the relative power over the two frequency bands from 0.5 to 30 Hz is given by Equation (1):

\[
\text{Relative power}(f_1, f_2) = \frac{P(f_1, f_2)}{P(0.5, 30)} \times 100\% \tag{1}
\]

where \( P \) refers to power, \( f_1 \) and \( f_2 \) represent the low and high frequencies, respectively [53].

4) **Variability per frequency band:** The variability in the power of each frequency band was estimated from the ratio of the median absolute deviation (MAD) to the median power [41], [45] in each frequency band and electrode position from each subject.

5) **SEF_{90%}**: The spectral edge frequency is the frequency below 90% of the total EEG power is located.

6) **Frequency band-specific amplitude:** The amplitude at each frequency band was expressed as percentage of the total EEG spectrum across EEG electrodes.

7) **Connectivity analysis:** For assessing the connectivity in the brain network, the magnitude squared coherence and the imaginary part of the coherence were computed for each patient from each EEG electrodes pair in the four frequency bands (i.e., \( \delta, \theta, \alpha, \beta \)). The brain connectivity between two groups of outcome (e.g., improved vs. unimproved) was determined through coherence, the imaginary part of coherence, weighted symbolic mutual information and symbolic transfer entropy [47], [54]–[56].

8) **Coherence:** The coherence was determined by calculating the mean of all magnitude-squared coherence of the signals in between all combinations of EEG channels [45], [47]. Coherence measures would indicate the functional relationship across brain region. The magnitude squared coherence of signals \( w(t) \) and \( r(t) \) for frequency \( f \) is defined as:

\[
C_{wr}(f) = \frac{|M_{wr}(f)|^2}{M_{wr}(f)M_{rr}(f)} \tag{2}
\]

where \( M_{wr}(f) \) and \( M_{rr}(f) \) are the power spectral densities of \( w(t) \) and \( r(t) \), respectively and \( M_{wr}(f) \) is their cross-spectral density [45], [47]. The imaginary part of coherence was considered to overcome the volume conduction effects (i.e., transmission of electrical signals from a primary current through brain tissue) on the real part of coherence. To correctly calculate connectivity across EEG scalp, the influence of volume conduction and differences between groups have to be considered.

9) **Entropy:** Entropy is an elementary measure of information theory [58] that is used to discriminate among complex system. The entropy was applied to EEG signals to measure the randomness of outputs of the complex system of neuronal networks underlying coma and consciousness.

10) **Approximate entropy (ApEn):** ApEn was introduced by Pincus as a measure of system complexity. Theoretically, ApEn measures the logarithmic likelihood of that at the next incremental comparisons, the patterns of data that are close to each other, remain close. Pincus [59], [60] described ApEn based on the following algorithm:

- **STEP 1:** Construct a segment of EEG data of \( N \) time samples \( w(1), w(2), \ldots, w(N) \).
- **STEP 2:** Fix \( m \) as integer \( r \) is a positive real number, where the value of \( m \) represents the vector length and \( r \) indicates a filtering level.
- **STEP 3:** Construct a series of vectors \( x(1), x(2), \ldots, x(N - m + 1) \) in \( R^m \) real \( m \)-dimensional space, which can be defined by \( x(i) = [w(i), w(i + 1), \ldots, w(i + m - 1)] \).
- **STEP 4:** Using the series of vectors \( x(1), x(2), \ldots, x(N - m + 1) \) to construct, for each \( i \), \( 1 \leq i \leq N - m + 1 \); additional quantity \( C_i^m \) can be calculated as:

\[
C_i^m(r) = \frac{N_r^{x(i)} \times x(i)}{N - m + 1} \tag{3}
\]

where \( N_r^{x(i)} \) is the number of \( x(i) \) such that \( d[x(i), x(j)] < r \); \( d[x, x^*] = \max_a |x(a) - w^*(a)| \); \( w(a) \) is the \( m \) scalar component of \( x \), \( d \) is the distance between the vectors of \( x(i) \) and \( x(j) \) at the maximum difference in their respective scalar component.

- **STEP 5:** Define:

\[
\phi^m(r) = \frac{\sum_{i=1}^{N-m+1} \log_{10}(C_i^m(r))}{N - m + 1} \tag{4}
\]

- **STEP 6:** Define approximate entropy (ApEn) as [45]:

\[
\text{ApEn}(N, m, r) = \phi^m(r) - \phi^{m+1}(r) \tag{5}
\]

where \( m \) and \( r \) are fixed as in **STEP 2**.

11) **Permutation entropy (PerEn):** As contrary to ApEn, permutation entropy (PerEn) used the symbolic transform, which the signal was represented by a sequence of discrete symbols. The probability density was evaluated to obtain the entropy. For example, a segment of EEG of \( N \) time samples, \( [w(1), w(2), \ldots, w(N)] \), then formed a set of sub-vectors \( [x(1), x(2), \ldots, x(N - m + 1)] \) where it could be constructed such that, \( x(i) = [w(i), w(i + \tau), \ldots, w(i + (m - 1) \tau)] \). Each \( x(i) \) was represented by a symbol (i.e., 1 and \( m! \)),
dependent on the order of amplitudes of the signal. Permutation entropy can be expressed as:

$$\text{PerEn} = -\sum_{i=1}^{m!} p_i \log_{10}(p_i)$$  \hspace{1cm} (6)

where \( p_i \) is the probability of the \( i \)-th symbol [47].

12) **Microstate analysis**: Microstate analysis is especially useful for classifying transitive brain state (e.g., Unresponsive Wakefulness Syndrome (UWS)/ Minimally Conscious State (MCS)) to probe the brain in different sleep stages and to understand the primary differences in brain functions of patient with different severities of coma. Microstate analysis was performed by analyzing the topographical maps of electrical potentials across the EEG electrodes as well as the temporal evolution of the selected topographical maps.

13) **Complex network analysis**: Complex network analysis was a part of connectivity, which employed the connectivity measures to represent the complex systems as network and extract the key information from topographical maps of these network. Complex network analysis represented the relationship of the structural-functional connectivity by defining functional connections from spatial-map of the brain. In EEG analysis, the networks have been created by representing the electrode position as nodes and forming the links between the nodes as functional connections. The graph-theoretical statistic (e.g., clustering coefficient, characteristic path length, modularity, participation coefficient and network level modular span) have been used to assess and compare the topology of these networks.

14) **EEG-Reactivity (EEG-R)**: EEG-R is defined as diffuse and temporary changes in EEG activity in response to external stimuli (e.g., from sensory, visual or auditory stimuli) and useful for prognostication in patients with impaired consciousness [11]. The EEG-R was assessed according to different categories: (1) present (amplitude and/or frequency changes); and (2) absent [49], [50].

15) **Normal EEG sleep stages**: EEG normal sleep stages is divided into 5 phases: wake, N1 (stage 1), N2 (stage 2), N3 (stage 3) and rapid eye movement (REM) sleep. N2 sleep transient refers to a transient stage-2 of normal EEG sleep phases. Posterior dominant rhythm (PDR) refers to the \( \alpha \) rhythm which occurs during relaxed wakefulness with eyes closed, attenuates with eyes opening [61]. Loss of specific continuous EEG (cEEG) features over time especially for bedridden TBI patients would be associated with outcome.

From our observation, qEEG variability for the relative and absolute band powers was the most common qEEG predictor included in the models (i.e., 46%) followed by total EEG power of all frequency bands between 0.5-20Hz (i.e., 31%), EEG-reactivity (EEG-R) (i.e., 31%) and also coherence for all possible combinations of EEG channels (i.e., 15%). In 8% of the models, other qEEG predictors were included (e.g., microstate,ApEn, transfer entropy (TE), connectivity, SEF\(_{90}\%), PerEn or complex network analysis).

One of the studies included the clinical variables (i.e., level of consciousness (LOC), LOC plus post-traumatic amnesia (PTA)) to compare the predictive performance of the models by adding qEEG predictors to increase the predictive power of the model [40]. EEG normal sleep characteristic (i.e., in the presence of PDR, N2 and delta-band (\( \delta \) ) activity) (i.e., 1 study) [52] and interactions between predictors (i.e., LOC variables, IMPACT variables, age, mean arterial blood pressure (MAP)) and qEEG parameters were investigated in two manuscripts [40], [52].

2) OUTCOME PREDICTION

Some variables needed transformation for further analysis. At the endpoint of the prediction model, the functional outcome was categorical variables (i.e., dichotomous GOS: favorable/unfavorable, or good/bad or positive/negative or awakening/death). Five studies used GOS as the primary output, but the time of outcome varied substantially; GOS at 12-month after injury (i.e., 1 study), GOS between 6 - to 12-month (i.e., 1 study), GOS at 6-month (i.e., 1 study), GOS at 17-month (i.e., 1 study) and also GOS at 30-day after injury (i.e., 1 study). In three studies, the GOSE had also been used as an output of the models (i.e., dichotomous: favorable/unfavorable, or good/bad or positive/ negative outcomes), in different time of assessment (i.e., at 3-month and 12-month after injury). Other categorizations of outcome variables were also made especially for predicting recovery outcome TBI patients in a vegetative state (i.e., dichotomous: UWS/MCS or improved/unimproved) (i.e., 3 studies). For the EEG-R, the outcome variables were dichotomous (present/absence), or (yes/no) or (reactive/unreactive) at 5-month of level of cognitive functioning scale (LCFS) (i.e., 1 study). Please refer to Table 3.

IV. PREDICTIVE MODEL DEVELOPMENT

Numerous ML techniques were used for developing the predictive models \( (n = 14) \) (please refer to Table 3). Most of the studies (i.e., \( n = 7 \)) used logistic regression (LR) models but others used random forest (RF) (i.e., \( n = 1 \)), generalized linear model (GLM) (i.e., \( n = 1 \)), linear discriminant analysis (LDA) (i.e., \( n = 2 \)), multivariate autoregressive model (i.e., \( n = 1 \)), support vector machine (SVM) (i.e., \( n = 1 \)), least absolute shrinkage and selection operator (LASSO) (i.e., \( n = 1 \)). Most of the studies (i.e., 95%) discussed about rationalization to include qEEG predictors in the prognostic model. The examination of the selected studies index revealed that two main types of qEEG feature extraction were used for model development of predictive TBI outcome: power spectral density (PSD) and fast Fourier transform (FFT) analyses.

In six studies (i.e., 43%), the FFT was applied to compute the amplitude at each frequency band.
### TABLE 3. Characteristics of 14 studies presenting predictive models for TBI outcome.

| No. | Authors & Year | Patients Group | Number of Patients | Input Features | ML Model | Training Set | Validation Method | Testing Set | Outcome Predicted | qEEG Predictors | Accuracy % | Sensitivity % | Specificity % | Area Under Curve | Time of Outcome |
|-----|----------------|----------------|--------------------|----------------|----------|--------------|-------------------|-------------|----------------|----------------|------------|-------------|--------------|----------------|-----------------|----------------|
| 1.  | Lee et al., 2019 [52] | Moderate to Severe | 152 TBI | eEEG | L.R | 106 | 46 | Unfavorable (GOSE 1 - 4) vs. Favorable (GOSE 5 - 8) | cEEG variables: (1) N2 sleep transients + (2) PDR + (3) δ activity | 24-hour cEEG | - | 94.0 | 49.0 | 0.77 | GOSE at 3-month |
|     |                |                |                    |                |          |              |                   |             |                |                |            |              |              |                |                 |                |
| 2.  | Haveman et al., 2019 [45] | Moderate to Severe | 57 TBI | eEEG | RF | 38 | 19 | Poor (GOSE 1 - 4) vs. Good (GOSE 5 - 8) | (1) 8 qEEG + (2) MAP + (3) age + (4) 9 IMPACT at 72 + 96 hour cEEG | Train set | - | 75.0 | 96.0 | 0.76 | GOSE at 12-month |
|     |                |                |                    |                |          |              |                   |             |                |                |            |              |              |                |                 |                |
| 3.  | 3. Stepfan et al., 2018 [47] | Severe | 39 TBI | EHG | GLM | 27 | 12 | Improved (i.e., MCS) vs. Unimproved (i.e., UWS) groups | Microstate duration in δ | - | - | - | 0.75±0.05 | 12 to 24 months |
|     |                |                |                    |                |          |              |                   |             |                |                |            |              |              |                |                 |                |
|     |                |                |                    |                |          |              |                   |             |                |                |            |              |              |                |                 |                |

VOLUME 8, 2020
TABLE 3. (Continued.) Characteristics of 14 studies presenting predictive models for TBI outcome.

| No. | Authors & Year | Patients Group | Number of Patients | Input Features | ML Model | Validation Method | Outcome Predicted | qEEG Predictors | Accuracy | Sensitivity | Specificity | Area Under Curve | Time of Outcome |
|-----|----------------|----------------|-------------------|---------------|-----------|------------------|-------------------|----------------|-----------|-------------|------------|-----------------|----------------|
| 7.  | Chen et al., 2017 [17] | Severe | 61 TBI | EEG, SVM | 4-fold CV | 18 | Positive outcome (GOSE > 2) vs. Negative outcome (GOSE < 2) | Microscale clustering coefficient in δ band | Modularity and clustering coefficient in δ band | 82.0 | 92.0 | 64.0 | 0.78 | GOSE at 12-month |
| 8.  | Schorr et al., 2016 [48] | Severe | 18 TBI | EEG, MVAR | 4-fold CV | Predicting recovery after severe TBI for improved (i.e., MCS) vs. unimproved (i.e., UWS) group | Parietal Coherence in δ | Parietal Coherence in δ | 73.0 | 75.0 | 0.75 | 12-month |
| 9.  | Mikola et al., 2015 [44] | Severe | 20 TBI | EEG, LR | 6 | Good outcome vs. Poor outcome | Mean NMAD_α | Mean TOT_δ | Mean NMAD_δ | Mean NMAD_θ | Mean NMAD_β | 64.0 | 77.0 | 0.71 |
| 10. | Logi et al., 2011 [49] | Severe | 60 TBI | EEG, LR | 14 | EEG-R reactive vs. unreactive | EEG-reactivity (Amplitude) | - | - | - | - | - | LCFS at 5-month |
| 11. | Amantini et al., 2005 [50] | Severe | 60 TBI | EEG, LR | 6 | EEG-R at Awakening (GOS 3 - 5) vs. Disability/Death (GOS 1 - 2) | EEG-R (Amplitude) | - | - | - | - | - | GOS at 12-month |
| 12. | Hebbet et al., 2002 [43] | Moderate to Severe | 51 TBI | EEG, LR | 1 | Favorable (GOS 4 - 5) vs. Unfavorable (GOS 1 - 2) | PAV threshold of mean 3-day PAV score | 0.10 | 100.0 | 20.0 | 100.0 | - | GOS at 1-month |
| 13. | Vespa et al., 2002 [43] | Moderate to Severe | 89 TBI | EEG, LR | 1 | Bad outcome (GOS 1 - 2); Good outcome (GOS 4 - 5) | PAV mean in 3 days + 6 clinical predictors | 64.0 | 55.0 | 69.0 | 0.68 | GOS at 1-month |
| 14. | Gutling et al., 1995 [51] | Severe | 158 TBI | EEG, Discriminant analysis | 1 | Good outcome (GR/MD); Bad outcome (SD/VS/D) | EEG-R | - | - | - | - | GOS at 17-month |

**Abbreviations:** ApEn (Approximate Entropy); BSR (Burst-Suppression Ratio); cEEG (Continuous EEG); CV (Cross-Validation); D (Death); DWT (Discrete Wavelet Transform); EEG (Electroencephalography); EEG-R (EEG-Reactivity); eMCS (Exit Minimally Conscious State); En (Entropy); GLM (Generalized Linear Model); GOS (Glasgow Outcome Scale); GOSE (Glasgow Outcome Scale Extended); GR (Good Recovery); hr (hour); IMPACT (International Mission for Prognosis and Analysis of Clinical Trials in TBI); LASSO (Least Absolute Shrinkage and Selection Operator); LCFS (Level of Cognitive Functioning Scale); LDA (Linear Discriminant Analysis); Lin. Regression (Linear Regression); LOC (Level of Consciousness); LOOCV (Leave One Out Cross-Validation); LR (Logistic Regression); max (maximum); MCS (Minimally Conscious State); MD (Moderate Disability); mTBI (Mild Traumatic Brain Injury); MVAR (Multi Variate Auto Regression); NMAD (Normalized Mean Absolute Difference); PSD (Power Spectral Density); PTA (Post Traumatic Amnesia); qEEG (Quantitative EEG); RF (Random Forest); sTBI (Severe TBI); SVM (Support Vector Machine); TOT (Total EEG power); UWS (Unresponsive Wakefulness Syndrome); VS (Vegetative State).

van den Brink et al. [46], used the FFT to create a metric of global frequency band-specific power across electrodes and across frequencies within four selected frequency bands. Five studies (i.e., 36%) used PSD analysis for computing the absolute power, total power, relative power and SEF90% within each frequency band: δ (0.5-4 Hz); θ (4-8Hz); α (8.13 Hz); and β (13-20 Hz). By applying this method, Mikola et al. [44] were able to extract 186 EEG features...
(e.g., normalized variability of the relative power in $\alpha$ and fast $\theta$, normalized variability of mean frequency, SEF$_{90\%}$, and total power in $\beta$, fast $\theta$ and $\alpha$ bands) for model training and testing (i.e., classification and prediction) using LDA classifier for predicting the positive or negative outcomes.

In two studies, coherence analysis [47], [48] was used to compute magnitude-squared coherence and the imaginary part of coherence of each pair of electrodes in the $\theta$, $\delta$, $\alpha$ and $\beta$ bands and two studies used connectivity analysis by applying two different approach in each study [17], [46]. A study employed Hilbert Transform analysis [46] to compute connectivity as complex analytic signal ($X$) over time ($t$) across scalp EEG electrodes and Chennu et al. [17] calculated debiased weighted phase lag index (dwPLI) value at the peak frequency of the signal oscillation across all the EEG channels to represent the connectivity between channel pairs. The authors constructed symmetric $173 \times 173$ dwPLI connectivity matrices of each selected frequency band (e.g., $\delta$, $\theta$, and $\alpha$).

Stefan et al. [47] calculated the global field power (GFP) by implementing a modified k-means clustering algorithm to extract the EEG topographic maps into four global microstates classes in different outcome groups (i.e., improved vs. unimproved). In the same study, the authors also determined the entropy of EEG signals to quantify the unpredictability neural networks due to consciousness. The authors reported that ApEn and PerEn provided valuable prognostic information of disorder of consciousness (DOC) due to severe head injury. Haveman et al. [45] trained an RF classifier to combine multiple individual classification tree using bootstrapped samples. With bootstrapped samples, the individual tree randomly selected each EEG feature and formed a split at each node. Based on their studies, modeling was performed by training an RF classifier with combinations of all qEEG features, age and mean of arterial pressure (MAP) to predict outcome at 24, 48, 72 and 96 hours after TBI. Prediction results were determined from the percentage vote of different classification tree [62].

Hack et al. [40] developed prognostic model using an LASSO classifier which used a regularized logistic regression model. In this approach, the authors used LASSO which consisted of a weighted combination of selected linear and non-linear EEG features together with selected clinical features. Chennu et al. [17] constructed SVM classifiers which trained two EEG predictors (i.e., $\delta$ modularity and clustering coefficients) to significantly predict future GOSE dichotomized outcomes in individual patients (accuracy = 82%).

V. PREDICTIVE MODEL VALIDATION

For model validation, fourteen studies performed split sample into training and testing sets. Six of these performed split-sample analysis using 70% for training and 30% for testing set or 67% for training and 33% model testing, respectively. In another eight studies, the split-sample analysis was not clearly reported. Concerning evaluation of the model performance, the discrimination was reported in the almost (8/14 studies, 93%) of the models through the area under curve (AUC). Accuracy rates were presented in four studies, sensitivity and specificity were reported in seven studies. Sensitivity, which is also called as True Positive Rate (TPR), describes the probability of measuring a positive value for a positive outcome. TPR is given in Equation (7).

$$TPR = \frac{TP}{TP + FN} = 1 - FNR$$ (7)

TP and FN denote true positive and false negative, respectively (i.e., TP: actual positive values correctly predicted as positive; FN: actual positive values mistakenly predicted as negative). FNR denotes false negative rate. The FPR can be calculated as (1 - specificity) or (1 - TNR). The specificity or TNR is given by Equation (8):  

$$TNR = \frac{TN}{TN + FP} = 1 - FPR$$ (8)

TN and FP denote true negative and false positive, respectively (i.e., TN: actual negative value correctly predicted negative; FP: actual negative values incorrectly predicted as positive).

In binary classification, the accuracy is used to evaluate the model performance. The accuracy rate is the proportion of patients or predictors with a certain outcome that is correctly predicted. The accuracy is determined by the following definition:

$$Accuracy = \frac{Number ~of ~correct ~predictions}{Total ~number ~of ~predictions} = \frac{TP + TN}{TP + TN + FP + FN}$$ (9)

The receiver operating characteristic curve (ROC) is a graphical representation of the performance of classifier or predictor as the classification threshold is altered. The information contained in the ROC curve is defined by the most popular index which is an AUC. The AUC provides a cumulative measure of performance across all possible classification thresholds [63], [64]. Since the AUC is a part of area in the unit square, its value will always be between 0 to 1.0. A perfect discrimination of the model has an AUC of 1.0, equivalent to a sensitivity and specificity of 100%.

3) PREDICTIVE MODEL VALIDATION PROCEDURES

Three studies performed classification using a leave-one-out cross validation (LOOCV) procedure in which each iteration used an only single sample for testing and the remaining for training. Two studies performed splitting up the sample into $k$ folds (i.e., 1-fold, 4-fold CV) [17], [43], using $k$-th fold as testing set and the residual ($k - 1$)-fold as training set [21]. A study used 10-fold stratified CV to avoid over-fitting and circulation analysis. In another eight studies, the cross-validation procedure was not clearly described. Table 4 reports the summary of key findings regarding predictive value of qEEG in predicting TBI outcomes.
### TABLE 4. Summary of findings regarding predictive value of EEG in predicting TBI outcomes.

| No. | Authors & Year | EEG Data Types | Number of Channels | qEEG Predictors | EEG Features Extraction | Key Findings | Conclusion |
|-----|----------------|----------------|-------------------|----------------|-------------------------|--------------|------------|
| 1.  | Lee et al., 2019 [52] | Continuous EEG | 13                | (1) Lack of PDR, (2) Lack of N2 sleep transient, (3) \( \delta \) activity | Visual analysis | **Predict unfavorable outcome:** LR got highest AUC (0.768), \( Se = 0.94 \) and \( Sp = 0.49 \) by adding three 24-hr cEEG features compared to IMPACT core model alone with AUC (0.649), \( Se = 0.79 \) and \( Sp = 0.44 \). **Predict favorable outcome:** After adding three 72-hr cEEG features + IMPACT predictors, the model predicted favorable outcome with AUC (0.689), \( Se = 0.96 \), \( Sp = 0.28 \). 139 from 152 (91%) patients had favorable outcome (good recovery) of 3-month GOS-E. | EEG features provide additional valuable predictors (prognostic information) in predicting outcome after severe TBI. Specific cEEG features over time are associated with functional outcome. |
| 2.  | Haveman et al., 2019 [45] | Continuous EEG | 19                | (1) Mean amplitude (total power), (2) SEP 90%, (3) Relative \( \alpha \)-power | PSD | **Predict poor outcome:** Best model (72-96 hr cEEG + IMPACT) used 19 features (8 EEG, MAP, age and 9 IMPACT predictors). Training set: (GOS1 - 2); AUC (50% CI)=0.94, \( Se = 0.75 \), \( Sp = 1.00 \). 29/38 (78%) patients had poor outcome. Testing set: (GOS1 - 2); AUC (50% CI)=0.81, \( Se = 0.75 \), \( Sp = 1.00 \). | Mean amplitude of EEG, age and MAP strongly contribute to the predictive ability of the poor outcome after TBI. Multi factorial RF models have potential to predict the outcome after TBI at different time intervals by combining all EEG features, age and MAP. EEG parameters (i.e., variability and absolute power) have positive correlation for predicting outcome in severe TBI. EEG parameters classify the groups of outcomes by providing highest AUC on its parameters for favorable outcome than unfavorable outcome group. |
| 3.  | Tolonen et al., 2018 [41] | Continuous EEG | 16                | 59 EEG parameters | PSD | 17/59 (28.8%) EEG parameters showed as the most significant EEG parameters for TBI outcome prediction. The best EEG parameters for outcome prediction: - Absolute \( \alpha \)-power (AUC = 0.87) - Variability of the relative fast \( \beta \) power (AUC = 0.84) - Variability of the relative \( \alpha \) power (AUC = 0.82) - Absolute slow \( \beta \) power (AUC = 0.84) | EEG parameters (i.e., variability and absolute power) have positive correlation for predicting outcome in severe TBI. EEG parameters classify the groups of outcomes by providing highest AUC on its parameters for favorable outcome than unfavorable outcome group. |
| 4.  | Stefan et al., 2018 [47] | Resting-state EEG | 39                | (1) Microstates in \( \delta \) and \( \theta \) bands, (2) ApEn in \( \phi \) band, (3) PerEn in \( \delta \) and \( \theta \) bands, (4) Power in \( \alpha \) and \( \delta \) frequency, (5) Connectivity (i.e., coherence in all EEG bands), (6) Complex network analysis (i.e., clustering coefficients \( \beta \) and \( \alpha \) coherence) | (1) OPR, (2) ApEn, (3) Multitaper method, (4) Coherence, (5) Graph-theoretical measure | EEG Biomarkers for Outcome Prediction (1) Microstate: Duration and frequency of microstate in \( \delta \) and \( \theta \) bands were the most successful EEG metrics for predicting coma outcome. (2) Entropy: ApEn in the \( \alpha \) band efficiently predicted outcome (AUC = 0.67 ± 5%), but permutation entropy in \( \delta \) (AUC = 0.71 ± 5%) and \( \theta \) (AUC = 0.83 ± 3%) performed better than ApEn. (3) Power in \( \alpha \) and \( \delta \) frequency: Power in \( \alpha \) and \( \delta \) performed better for dissociating improved vs. unimproved outcomes. (4) Connectivity: Coherence in all frequency bands was greater for patient with improved condition. The most prominent was coherence in \( \theta \) band yielding high classification accuracy (AUC = 0.78 ± 2%). wSMI & TB in \( \delta \) and \( \alpha \) exhibited the most notable prognostic power at predicting patient outcome. (5) Complex Network Analysis: Clustering coefficient in \( \beta \) and \( \alpha \) coherence performed best classification for discriminating patient into two outcomes (i.e., improved vs. unimproved). | Most EEG features presented were commonly applied in EEG analysis of consciousness (i.e., severe TBI). Application of ML techniques to build a model has significantly increased the usefulness of qEEG biomarkers by performing better than chance at predicting coma outcome. |
| 5.  | van den Brink et al., 2018 [46] | Task-free EEG | 8                 | (1) Amplitude, (2) Connectivity | (1) FFT, (2) Hilbert transform | Most successful EEG metrics to predict each patient’s chances of recovery: - \( \alpha \) amplitude yielded the highest classification accuracy (62%) by dissociating patients from MCS to eMCS outcome measures. - \( \theta, \alpha, \beta \) band connectivity was able to predict patient’s chance of recovery 3 months later with 75% accuracy. | Application of ML algorithms and EEG features would reveal the potential of connectivity and amplitude metrics as reliable prognostic markers. |
TABLE 4. (Continued.) Summary of findings regarding predictive value of EEG in predicting TBI outcomes.

| No. | Authors & Year | EEG Types | Data Collection | Number of Channels | qEEG Predictors | EEG Features Extraction | Key Findings | Conclusion |
|-----|----------------|-----------|-----------------|-------------------|----------------|------------------------|--------------|------------|
| 6.  | Hack et al., 2017 [40] | Eyes closed resting-state EEG | 5 | Total power | PSD | - The accuracy of TBI prediction model increased when adding EEG features to the LOC model with 83% improvement over LOC alone. - Adding EEG features as predictors had improved predictions outcomes of the model (LOC+EEG features) (AUC = 0.83) compared with LOC alone (AUC = 0.64) or LOC with amnesia (i.e., RGA/PTA) (AUC = 0.69). | qEEG acts as potential electro-physiological marker to prediction outperforms clinical observation of LOC and amnesia alone. |
| 7.  | Chennu et al., 2017 [17] | Resting-state EEG | 256 | Total power, dWPLI (connectivity) | PSD | Positive predictor of recovery Higher microscale clustering coefficient (AUC = 0.78) in δ network of patients with positive outcome. - EEG features: δ modulatory and clustering coefficient in δ network significantly were able to predict future GOSE in discriminating positive and negative outcomes (Acc: 82%, Se: 92%, Sp: 64%) | EEG network metrics (connectivity in δ band) confirm the valuable predictors of recovery which can complement demographic and clinical information. |
| 8.  | Schorr et al., 2016 [48] | Resting-state EEG | 256 | Coherence coefficient | Coherence | Partial coherence was predominant in prediction of consciousness recovery after 12-month compared to other brain region resulted in the highest sensitivity at 73%, specificity at 79% and AUC 0.75. Based on results, coherence in different frequency band showed strong discriminative performance for the differentiation of improved and unimproved group. | High sensitivity and specificity of the coherence at certain brain region and frequency bands are indicative of their usefulness as prognostic markers for a positive outcome. |
| 9.  | Mikola et al., 2015 [44] | Continuous EEG | 4 | 186 EEG features | PSD | 12/186 EEG features were found significantly correlated (p (AUC =0.5)) to the outcome of recovery of TBI patients. The best correlating EEG features included the normalized variability of the relative power in α and fast θ frequency bands, the normalized variability of mean frequency, spectral entropy and total power in β, fast θ and α bands. -92/20 severe TBI patients had been considered to have good outcome. | Certain EEG features (e.g., variability of the α band power) have been found to correlate to positive/good outcome. |
| 10. | Logi et al., 2011 [49] | Task related EEG | 16 | Amplitude of EEG-reactivity | FFT | Predicting consciousness recovery within 5-months following severe TBI EEG-reactivity (EEG-R) was a good predictor for recovery of consciousness but absence of EEG-R was not invariably associated with bad prognosis. | EEG-reactivity acts as a good positive predictor for the prognosis of recovery of consciousness in post-acute phase of TBI. |
| 11. | Amantini et al., 2011 [50] | Task related EEG | 8 | Amplitude of EEG-reactivity | FFT | EEG-R predicted awakening of consciousness at 1-year GOS with 30/36 patients had reactive EEG but 15/24 of them survived with severe disability. EEG-R predicted 1-year awakening from coma with AUC = 0.69. | Non-reactive EEG is associated with bad outcome. |
| 12. | Hebb et al., 2007 [42] | Continuous EEG | 16 | Percentage of Alpha Variability (PAV) | FFT | Predicting favorable recovery (GOS 4-5) vs. unfavorable (GOS 1-3) of 6-month outcome. - Positive association between the 5-day PAV, 7-day PAV patterns and 6-month clinical outcome of patients following TBI. - Results showed worsening in PAV scores to 0.1 associated with unfavorable outcome/death and high PAV scores to 0.20 associated with favorable recovery. (AUC = 0.76, Se = 0.87, Sp = 0.82). | PAV patterns provide positive association with 6-month outcomes (GOS 4-5) compared to PAV patterns of those having GOS 1-2. EEG predictors (i.e., PAV scores) appears to be the best predictor of 6-month clinical outcomes among the traditional clinical and anatomical predictors. |
4) DESCRIPTION OF PREDICTIVE MODEL VALIDATION

Lee et al. [52] developed a predictive model to predict functional outcome of unfavorable outcome (GOSE 1-4; death, unresponsive wakefulness or severe disability) and favorable outcome (GOSE 5-8; moderate disability or good recovery at 3-month) [52]. The authors developed the model from 152 moderates to severe TBI patients and validated the model in 46 patients from the same samples. The absence of PDR, absence of N2 sleep transients and predominant background activity during 24-hours cEEG were qEEG predictors of the model. They reported sensitivity and specificity for the unfavorable outcome of 94% and 49% with an AUC of 0.77. With the presence of these qEEG features within 72-hour, the model predicted good outcomes with a specificity of 96% and sensitivity of 28% with an AUC 0.76.

Haveman et al. [45] developed three prognostic models. In the first model, they used eight of qEEG features and in others, they added age and clinical variables (i.e., mean arterial pressure (MAP)) and nine IMPACT predictor variables on the second model and the third model used only IMPACT predictor for model comparison. The LOOCV was used for evaluating model performance in the training set. They developed the model in 38 patients included between 2013-2016 for the training set and 19 patients included between 2017-2018 for model validation/testing set. Finally, these three models were trained and evaluated to predict poor outcome as GOSE 1-2. They reported sensitivity, specificity, and AUC for predicting poor outcomes by comparing the model performance of each model in both training and validation set. Performance measures for prediction model were compared respectively; for training set: cEEG at 72 + 96hr (Se = 92%, Sp = 71%, AUC = 0.88), 72 + 9hr cEEG with IMPACT (Se = 75%, Sp = 100%, AUC = 0.94), IMPACT predictor (Se = 65%, Sp = 81%, AUC = 0.65), in validation set: cEEG at 72 + 96hr (Se = 63%, Sp = 83%, AUC = 0.75), 72 + 9hr cEEG with IMPACT (Se = 100%, Sp = 75%, AUC = 0.81), IMPACT predictor (Se = 88%, Sp = 73%, AUC = 0.84). From these findings, they demonstrated the contribution of qEEG predictors to enhance the predictive ability of poor outcomes and showed the potential of RF models to predict the poor outcomes. They reported the hierarchy contribution of qEEG features (i.e., eight qEEG predictors), age, MAP and other clinical factors (i.e., CT scan images, pupil reactivity, hypotension, hypoxia) in the prediction of good or poor outcome for selecting the best predictive model at 72 + 96hr after TBI in the easiest and user-friendly way.

Stefan et al. [47] developed a predictive model by fitting a generalized linear model (GLM) on training data and testing the model on test data. GLM model discriminated the patients into two classes (i.e., improved vs. unimproved) to determine the predictive power of the selected qEEG features (i.e., microstate, ApEn, power in the theta frequency, connectivity, complex network analysis). The 10-fold stratified cross-validation was implemented on the training set to avoid model from predicting noise and circular analysis. The authors revealed that coherence in the theta band got the highest classification accuracy with AUC_θ = 78 ± 2%, q < 0.0001 compared to alpha coherence (AUC_α = 62 ± 4%, q < 0.0001) and beta coherence (AUC_β = 67 ± 1%, q < 0.0001). For complex network analysis, they found that clustering coefficients from beta coherence (AUC_β = 82 ± 1%, q < 0.0001) and alpha coherence (AUC_α = 82 ± 2%, q < 0.0001) performed well at discriminating patients into the two outcome classes (UWS to MCS or death or a permanent DOC) at varied thresholds values.

Chennu et al. [17] built a predictive model using a supervised ML algorithm (i.e., SVM) trained with 4-fold cross-validation, in which each of the training and validation set contained homogenous proportions of each subject group to avoid any fold-specific effects. Seven qEEG features were calculated for three EEG selected frequency bands (i.e., δ, θ, α): (1) mean relative power over all channels; (2) clustering coefficients; (3) median connectivity across all paired channels; (4) modularity; (5) participation coefficients;
(6) characteristic path length; and (7) modular span, which resulted in a total of 21 metrics. They reported the ability of each qEEG metric to discriminate positive and negative outcomes of recovery by calculating the AUC from ROC analysis.

They reported the ability of each qEEG metric to discriminate positive and negative outcomes of recovery by calculating the AUC from ROC analysis. They reported that patients with positive outcomes got higher microscale clustering coefficients in their δ network connectivity and local topographical connectivity in δ band with an AUC = 0.78. They confirmed that the presence of δ modulatory and clustering coefficient in δ network had significantly predicted future GOSE in isolating the positive and negative outcomes of 82% accuracy, 92% sensitivity, and 64% specificity.

Vespa et al. [43] reported three different models. The outcomes were: predicting alive compared with the dead (i.e., GOS score at 5 vs. GOS score 1) and predicting good outcome compared with the bad outcome (GOS score at 4 or 5 vs. GOS score 1 or 2), respectively. For the validation of the two outcomes, they used 1-fold-cross-validation for each model in 89 TBI patients. The predictors were divided into two categories which were: (1) qEEG predictor (PAV mean in 3-day post-injury); and (2) six clinical predictors (GCS score, pupillary response to light, patient age, the result of computerized tomography and early hypotension). They reported the model discrimination of each model separately for model comparison.

For prediction of dichotomized GOS scores (score of 5 compared with scores 1), the combined six clinical predictors and PAV mean at 3-day predicting alive with dead provided highest AUC = 0.87, Accuracy = 78% Se = 83%, and Sp = 63% compared to standalones models (i.e., PAV mean 3-day (AUC = 0.82, Accuracy = 86%, Se = 93%, Sp = 63%) and six clinical predictors (AUC = 0.75, Accuracy = 75%, Se = 87%, Sp = 38%)). For prediction of dichotomized GOS scores (scores of 1 and 2 compared with scores of 4 and 5), the results revealed that the combined six clinical predictors + PAV mean 3-day for predicting good outcome got a 1-fold-cross-validation predictive accuracy of 73% (Se = 81%, Sp = 65%, AUC = 0.83) compared with individual models. Both models showed diminished in accuracy of 67% for clinical predictors model (Se = 80%, Sp = 55%, AUC = 0.76) and 63% accuracy of PAV mean 3-day model (Se = 63%, Sp = 62%, AUC = 0.72). The authors demonstrated that presence of qEEG predictors (i.e., PAV) had enhanced the sensitivity and specificity during forecasting the outcomes at 3-day post-injury compared to individually prognostic variables.

VI. DISCUSSIONS
This systematic review evaluates fourteen predictive models for use in a different cohort of TBI patients (i.e., moderate to severe) that considered qEEG data. In this systematic review, the authors developed a multifactorial predictive model as proof of concepts for using ML techniques including qEEG features for prediction of outcomes in moderate to severe TBI patients. Substantial limitations are scrutinized in the development of predictive models that limit their predictive performances.

A. STUDY POPULATION AND EEG PREDICTORS
Based on patient groups reported in fourteen publications of prognostic models for TBI patients, the models for moderate and severe TBI patients were frequently investigated rather than for mild and moderate TBI. As described in Rapp et al. [65], mild TBI prompted a dense adjustment of neuronal oscillation process that cannot be simply detected by standard magnetic resonance imaging (MRI) at any stage of injury. For examples: (i) shortening in the dendrite system of neurons may cause a decrease in the power of fast frequencies; (ii) fluctuations of firing properties of thalamic and cortical neurons may boost the θ frequency activity. Therefore, the EEG would become a promising tool in detecting the physiological changes of these processes [12] as well as for predicting the long-term outcomes of patients with TBI [66], [67].

Most of the studies (n = 14) focused on the development of the prognostic models in the chronic stage of TBI, however specific qEEG variables obtainable from mild and moderate TBI patients would possibly provide significant prognostic information in predicting outcomes after mild or moderate TBI. In addition to that, the association between qEEG features and predictive ability (i.e., predicting poor or good outcome) of prognostic model for these groups (i.e. mild and moderate) have not been clearly established.

Large number of sample sizes are necessary for reliable selection of predictors [35], [68], [69]. Limitation of a low number in sample size often gives an impact on the performance of the predictive model. Thus, manipulating data is a challenge to ML applications to determine the specific association of qEEG predictors to yield the most accurate prognosis in TBI outcomes [46]. In the present systematic review, the predictive models are developed on a relatively small number of patients (e.g., n = 16) but since a very little is known about the usefulness of any qEEG predictors for the prognosis of TBI patients, the available number of patients are still useful for providing a valuable predictive performance of each ML model.

Two studies included a large number of EEG predictors. Tolonen et al. [41] considered over 59 qEEG predictors, and Mikola et al. [44] included 186 qEEG predictors meanwhile the remaining studies used 3 to 5 qEEG predictors for training and model validation. In the twelve studies, the final models included the maximum five EEG predictors variables for each predictive model. Haveman et al. [45] demonstrated the positive association of including multiple qEEG parameters with TBI outcomes in RF model since no ML models based on multiple qEEG parameters were reported to predict patient outcome after moderate to severe TBI. In accordance with Tolonen et al. [41], the multiple qEEG parameters were grouped into five major classes (i.e., (i) absolute power;
(ii) relative power; (iii) asymmetry; (iv) variability; and (v) others: spectral entropy, burst suppression ratio) and the authors found that 17 of the 59 qEEG parameters were strongly associated with a good outcome. For example, the EEG absolute power showed a positive correlation to the outcome, as the patients who recovered, had stronger amplitude in their EEG than those who did not.

In other study, Mikola et al. [44] confirmed that 12 of the 186 qEEG predictors contributed high predictive value based on a calculation of their features probabilities of AUROC. Based on the above findings, these numbers of qEEG predictors would appear reasonable for developing predictive models without decreasing the model performance. In contrast, the predictive models which included five or seven clinical predictors outperformed models that included only three or four predictors [70]. Mushkudiani et al. [71] in a review suggested to use a limited set of predictors because including too many predictors would increase the risk of overfitting models and thus limiting the prognostic value of the model leading to poor discrimination or calibration.

B. PREDICTIVE MODEL DEVELOPMENT

In predictive model development, ML techniques ranging from linear regression to deep neural networks would be used to represent a powerful technology capable of effectively predicting outcomes after TBI to support clinical decision-making. In this systematic review, discriminant analysis, LDA, SVM, linear and logistic regression analyses, network analysis using the LASSO technique, classification tree using RF technique and generalized linear model (GLM) are considered for model development. Regression analysis is the most frequently used. Logistic regression analysis, which is a supervised learning algorithm is frequently used to calculate the percentage of correct or incorrect predictions according to dichotomized output (i.e. good or bad outcome) of the model.

Logistic regression relates a number (n) of qEEG predictors (M = {M1, ..., Mn}) to an outcome (K) by multiplying the characteristics with regression coefficients (β = {β1, ..., βn}). These regression coefficients represent the strength of the relation between qEEG predictors characteristic and the dichotomized GOSE or GOS outcome. The output of the regression model is dichotomized into: (i) good outcome - can take the value 1 with a probability P; (ii) bad outcome - can take the value 0 with a probability 1 − P. Thus, the relationship between the qEEG predictor and outcome variables is expressed by the logit transformation of P:

\[
\text{logit}[P(M)] = \log\left(\frac{P(M)}{1 - P(M)}\right) = \alpha + \beta_1 M_1 + \beta_2 M_2 + \cdots + \beta_n M_n \quad (10)
\]

Logistic regression model is assessed with maximum likelihood methods by assuming that the outcome K follows a binomial distribution [72], [73]. The advantages of logistic regression over discriminant analysis is that it requires less assumptions. For example, independent variables do not need to be normally distributed, linearly related or have same variance within-group. In addition to that, the regression models are more statistically robust, work based on analysis of relationship between variables and trends, and handle predictions on continuous as well as categorical variables. In contrast, most of the authors agreed that other ML approaches (e.g., decision tree (DT), k-nearest neighbor (k-NN), neural networks (NN)) outperformed logistic regression in prediction of various disease conditions [19], [20], [34], [35], [69], [74], but there was substantial heterogeneity in the modeling process, input and output variables and methods for validating predictive model performance.

Supervised ML methods like SVM, LDA, GLM RF, and logistic regression fall under classification algorithms which involve the process of finding a model or function (f) that helps in separating the data into multiple categorical classes (i.e., discrete values). In classification techniques, the algorithms learn from given data input (i.e. labeled data) and then use this learning to classify new observations. In contrast to logistic regression, SVM is flexible in representing sophisticated relationships but tends to overfit [68], [75]. SVM classifies data points by choosing the “best hyperplane” for separating a set of points into two groups of data. In SVM, a data point is given as a k-dimensional vector (a list of k numbers) and SVM separates these points with a (k − 1) dimensional hyperplane. SVM will generate many hyperplanes that might classify the data but SVM will only select the “best hyperplane” to represent the largest separation or margin between the two classes [76].

In their article, Stefan et al. [47] used a GLM classifier in training and testing data of qEEG predictors (e.g., power in α and δ) to determine the predictive power of two groups of outcomes (i.e. improved vs. unimproved). The predictive performance was determined by comparing the binary output of GLM to the actual labels of unimproved vs. improved represented by 0 and 1. GLM is the extension of the linear regression model which assumes that the outcome given by the input features follows a Gaussian distribution. The GLM is implemented from the basic formula of linear regression (see Equation (11)) for classification of continuous outcomes and logistic regression for classification of binary output.

\[
K = \beta_0 + \beta_1 M_1 + \cdots + \beta_n M_n + \varepsilon \quad (11)
\]

K is the outcome of an instance expressed by a weighted sum of its n features with an individual error ε which follows a Gaussian distribution.

We notice that several studies used the advantage of ML techniques (i.e. LASSO, LDA and RF) to identify or develop a more appropriate prognostic model to predict TBI outcomes. A study developed an RF model that combined multiple individual classification trees to randomly select qEEG features at each node [45]. The optimal number of trees (n = 100) was computed by observing the out-of-bag (OOB) error and setting the number of maximum nodes (n = 20). RF classifier generates a model with less bias and variance by fitting and averaging several decorrelation decision trees [66].

N. S. E. Mohd Noor, H. Ibrahim: ML Algorithms and Quantitative EEG Predictors for Outcome Prediction
Decorrelation means that at each decision node, the variables will select the next branch of each tree in a random subset of all available variables, to make sure that the individual trees are heterogeneous and have variability in the data. RF classifier is an ensemble of learning method and is preferred because of its ability to avoid overfitting of a single tree [19], [77]. However, a problem occurred when using multifactorial data; the RF models excluded observations with missing data, meaning that the other possible relevant predictors (i.e., clinical or electrophysiological data), which were incomplete in all patients, were not considered in modeling process [45].

Network analysis (i.e., LASSO) is a supervised ML algorithm under regression analysis. The LASSO method develops a classifier using a regularized logistic regression model with a LASSO penalty (L1 absolute value). LASSO includes a term to penalize model complexity and diminish less importance variables towards zero to produce a scattered model and more interpretable [66], [78]. Hack et al. [40] used the LASSO classifier to demonstrate the performance of an algorithm to develop three predictive models derived from different input features (i.e., LOC only, LOC + amnesia information (i.e., PTA) and LOC + qEEG features). qEEG-based classification algorithm demonstrated significant improvement (AUC = 83%) in predictive power compared with other two models.

Based on data presented in our systematic review, most of the studies included a comprehensive description of the chosen ML techniques (i.e., 85%) which is important for understanding and mimicking of the developed predictive model. In our viewpoint, any predictive ML model should have their codes accessible for anyone to examine and do some improvement, as well as to reduce the black-box properties of ML models resulting in more practical implementation predictive ML models in clinical practice.

C. PREDICTIVE MODEL VALIDATION

Predictive model performance and generalizability are important for evaluating each prognostic model performance. Models validation involves two steps: (i) internal validation; and (ii) external validation. In this systematic review, most of the studies conducted internal validation within the development performed by split-sample, cross-validation or bootstrap techniques compared to external validation [71]. The predictive model performance can be estimated by evaluating calibration and discriminant characteristics. No studies used calibration tests in this review. Most of the studies (i.e., \( n = 14 \)) reported discriminant tests in terms of accuracy rate (Acc), Sensitivity (Se), Specificity (Sp) and AUC.

The accuracy of the prognostic model is determined from the “testing set” that consists of observations of previously unseen by the model. From four prognostic models reporting accuracy, Chennu et al. [17] presented the connectivity in \( \delta \) modularity coefficient with excellent predictive accuracy above 85% and other qEEG predictors presented at least good accuracy above 70% [46] and two studies presented their fair accuracy below 60% in different qEEG predictors (e.g., PAV [42], [43]. In addition to that, most of the prognostic model (\( n = 10 \)) considered GOS or GOSE for favorable or unfavorable as an outcome for AUC > 0.69 which was considered as excellent discrimination. The AUC > 0.69 also demonstrated the strong associations of individual EEG predictors to predict the favorable or unfavorable outcomes.

The purpose of the prognostic model is to provide valid outcome predictions for new data patients. Prognostic models have been proposed as an alternative to improve power in long-term TBI outcome prediction [79]–[81] and assist in clinical audit [82], [83]. For this purpose, it is crucial to evaluate both internal and external validities of proposed prognostic models. Internal validation is important to ensure that the model can be implemented in a patient similar to the development series using a similar setting. While for the external validation, the model could practically be used to assess performance in new patients, for example, patients from different settings.

Internal validation using split-sample analysis (i.e., splitting data into one part of the study population for model training and another part for validation or testing) was reported in eight studies. With split-sample validation, the patient population is randomly separated into two groups (i.e., training and testing). In training datasets, the model is created from \( \frac{2}{3} \) of the data, and the model performance is evaluated in from \( \frac{1}{3} \) of the data. This is a classical procedure and has various disadvantages such as inefficiency and uncertainty [74].

However, cross-validation (CV) and bootstrapping methods are the most preferred and efficient for model validation. Cross-validation involves extensive split-sample into \( k \) parts, using \( k \)-th part as testing data and remaining \( (k – 1) \) parts as the training set. This process is repeated \( k \) times until all data have been used as a testing set once. For example, leave one out cross-validation (LOOCV) is the most extreme and computationally demanding because each iteration uses a single data for testing and the rest for training. Stefan et al. [47] used an option to split the dataset into ten parts, used one part for testing and nine parts for training, which were equivalent to ten-fold CV. They required ten iterations but it is recommended to repeat the whole procedure many times to obtain more reliable performance measures [21].

To the best of our knowledge, the result of the current systematic review indicates the substantial benefits of using qEEG parameters for predicting outcome in patients with TBI. With the application of ML models, several qEEG features will potentially provide useful prognostic information in prediction of outcome in patients with moderate to severe TBI. In the future, we suggest that the predictive power of the qEEG features should be compared with clinical predictive models such as IMPACT in TBI in order to find out the added value of adding qEEG variables to the models. For this purpose, a combination of ML models and qEEG features would reveal the efficiency of considering qEEG features as reliable prognostic markers in TBI outcome prediction.
VII. CONCLUSION
To conclude, from the present relatively small sample population and few qEEG variables, we found that ML techniques have great potential for improving TBI outcome prediction. The ML models (i.e., logistic regression, RF, LASSO, GLM) described in this review would benefit from further development. We believe that development of prognostic model based on combination of multiple qEEG features can be rendered more efficiently using an advanced ML algorithm (i.e., neural networks) by considering the heterogeneity of brain injury (i.e., mild, moderate and severe), which in turn can hopefully increase the predictive power of a model. Future studies should explore the hurdles in model creation, validation and consideration of applying these methods for routine clinical TBI decision making.

REFERENCES
[1] S. S. Shin, J. W. Bales, C. Edward Dixon, and M. Hwang, “Structural imaging of mild traumatic brain injury may not be enough: Overview of functional and metabolic imaging of mild traumatic brain injury,” Brain Imag. Behav., vol. 11, no. 2, pp. 591–610, Apr. 2017.
[2] M. C. T. Collaborators, P. Perel, M. Arango, and T. Clayton, “Predicting outcome after traumatic brain injury: Practical prognostic models based on large cohort of international patients,” BMJ, vol. 336, no. 7641, pp. 435–429, Feb. 2008.
[3] W. A. Knaus, E. A. Draper, D. P. Wagner, and J. E. Zimmerman, “APACHE II: A severity of disease classification system,” Crit. care Med., vol. 13, no. 10, pp. 818–829, 1985.
[4] J. R. Edlich, “An injury severity scale for comprehensive management of central nervous system trauma,” J. Amer. College Emergency Physicians, vol. 8, no. 2, pp. 64–67, Feb. 1979.
[5] L. M. Mottett, “Traumatic brain injury: Assessment, resuscitation and early management,” Brit. J. Anaesthesia, vol. 99, no. 1, pp. 18–31, Jul. 2007.
[6] Y. You, B. S. Liew, A. K. Rosman, B. I. Devi, and A. Agrawal, “Outcome measures for traumatic brain injury: Advances and clinical utility,” Neuropsychiatric Disease Treat., vol. 11, pp. 2355, Sep. 2015.
[7] E. A. Draper, D. P. Wagner, and J. E. Zimmerman, “APACHE II: A severity of disease classification system,” Crit. Care Med., vol. 13, no. 10, pp. 818–829, 1985.
[8] J. R. Edlich, “An injury severity scale for comprehensive management of central nervous system trauma,” J. Amer. College Emergency Physicians, vol. 8, no. 2, pp. 64–67, Feb. 1979.
[9] L. M. Mottett, “Traumatic brain injury: Assessment, resuscitation and early management,” Brit. J. Anaesthesia, vol. 99, no. 1, pp. 18–31, Jul. 2007.
[10] Y. You, B. S. Liew, A. K. Rosman, B. I. Devi, and A. Agrawal, “Outcome measures for traumatic brain injury: Advances and clinical utility,” Neuropsychiatric Disease Treat., vol. 11, pp. 2355, Sep. 2015.
[11] E. A. Draper, D. P. Wagner, and J. E. Zimmerman, “APACHE II: A severity of disease classification system,” Crit. Care Med., vol. 13, no. 10, pp. 818–829, 1985.
[12] J. R. Edlich, “An injury severity scale for comprehensive management of central nervous system trauma,” J. Amer. College Emergency Physicians, vol. 8, no. 2, pp. 64–67, Feb. 1979.
[13] L. M. Mottett, “Traumatic brain injury: Assessment, resuscitation and early management,” Brit. J. Anaesthesia, vol. 99, no. 1, pp. 18–31, Jul. 2007.
[14] Y. You, B. S. Liew, A. K. Rosman, B. I. Devi, and A. Agrawal, “Outcome measures for traumatic brain injury: Advances and clinical utility,” Neuropsychiatric Disease Treat., vol. 11, pp. 2355, Sep. 2015.
[15] E. Celtikci, “A systematic review on machine learning in neurosurgery: The future of decision-making in patient care,” Turk Neurosurg, vol. 28, no. 2, pp. 167–173, 2018.
[16] E. D. Bigler. “Systems biology, neuroimaging, neuropsychology, neuroconnectivity and traumatic brain injury,” Frontiers Syst. Neurosci., vol. 10, pp. 55, Aug. 2016.
[17] S. Chen, J. Annen, S. Wannez, A. Thibault, C. Chatelle, H. Cassol, G. Martens, C. Schnakers, O. Gossières, D. Menon, and S. Laureys, “Brain networks predict metabolism, diagnosis and prognosis at the bedside in disorders of consciousness,” Brain, vol. 140, no. 8, pp. 2120–2132, Aug. 2017.
[18] E. M. Palacios, E. L. Yuh, Y.-S. Chang, J. K. Yue, D. M. Schneyer, D. O. Okonkwo, A. B. Valadka, W. A. Gordon, A. I. R. Maas, M. A. Vassar, G. T. Manley, and P. McKhann, “Resting-state functional connectivity alterations associated with six-month outcomes in mild traumatic brain injury,” J. Neurotrauma, vol. 34, no. 8, pp. 1546–1557, Apr. 2017.
[19] J. T. Sanders, P. C. Staples, A. V. Karhade, M. M. Zaki, W. B. Gormley, M. L. D. Broekman, T. R. Smith, and O. Arnaout, “Machine learning and neurosurgical outcome prediction: A systematic review,” World Neurosurg., vol. 109, pp. 476–486, Jan. 2017.
[20] J. T. Sanders, O. Arnaout, A. V. Karhade, H. D. Hansenbrock, W. B. Gormley, M. L. D. Broekman, and T. R. Smith, “Natural and artificial intelligence in neurosurgery: A systematic review,” J. Neurosurg., vol. 83, no. 2, pp. 181–192, Aug. 2018.
[21] Q. Nohrhomme, R. Brecheisen, D. Lesenfants, G. Antonopoulos, and S. Laureys, “Look at my classifier’s result: Disentangling unresponsive (minimally) conscious patients,” Neuroimage, vol. 145, pp. 288–303, Jan. 2017.
[22] E. Tolland, D. Galanaud, V. Pelbarg, P. Sanchez-Pena, Y. Le Fur, L. Abdennour, P. Cozzone, S. Lehericy, J. Chiras, and L. Puybasset, “Experience of diffusion tensor imaging and 1H spectroscopy for outcome prediction in severe traumatic brain injury: Preliminary results,” Crit. Care Med., vol. 37, no. 4, pp. 1448–1455, 2009.
[23] P. J. D. Andrews, D. H. Sleeiman, A. McQuatt, V. Corbule, C. H. Jones, T. P. Howley, and C. S. A. Macmillan, “Predicting recovery in patients suffering from traumatic brain injury by using admission variables and physiological data: A comparison between decision tree analysis and logistic regression,” J. Neurosurg., vol. 97, no. 2, pp. 326–336, Aug. 2002.
[24] G. E. Strangman, T. M. O’Neil-Pirozzi, C. Supelana, R. Goldstein, D. I. Katz, and M. B. Glenn, “Regional brain morphometry predicts memory rehabilitation outcome after traumatic brain injury,” Frontiers Hum. Neurosci., vol. 4, p. 182, Oct. 2010.
[25] A. Irimia, M. C. Chambers, J. R. Alger, M. Filipppo, M. W. Prastawa, B. Wang, D. A. Hovda, G. Gerig, A. W. Toga, R. Kikinis, P. M. Vesa, and J. D. Van Horn, “Neuroimaging of structural pathology and connectomics in traumatic brain injury: Toward personalized outcome prediction,” NeuroImage: Clin., vol. 1, no. 1, pp. 1–17, 2012.
[26] A. Irimia and J. Van Horn, “Functional neuroimaging of traumatic brain injury: Advances and clinical utility,” Neuropsychiatric Disease Treat., vol. 11, pp. 2355, Sep. 2015.
[27] E. A. Draper, D. P. Wagner, and J. E. Zimmerman, “APACHE II: A severity of disease classification system,” Crit. Care Med., vol. 22, no. 1, p. 184, Aug. 2018.
[28] J. D. Lewine, S. Plis, A. Uluoa, C. William, M. Spitz, J. Foley, K. Paulson, J. Davis, N. Banger, T. Snyder, and L. Weaver, “Quantitative EEG biomarkers for mild traumatic brain injury,” J. Clin. Neurophysiol., vol. 36, no. 4, pp. 296–305, 2019.
[29] S. A. Christie, A. S. Conroy, R. A. Callcutt, A. E. Hubbard, and M. J. Cohen, “Dynamic multi-outcome prediction after injury: Applying adaptive machine learning for precision medicine in trauma,” PLoS ONE, vol. 14, no. 4, Apr. 2019, Art. no. e0213836.
[30] D. B. Douglas, T. Ro, T. Toffoli, B. Krachwuch, J. Muldermans, J. Gullo, A. Dubler, A. E. Anderson, P. K. Douglas, and M. Wintermark, “Neuroimaging of traumatic brain injury,” J. Med. Sci., vol. 7, no. 1, p. 2, 2019.
A. Amantini, A. Grippo, S. Fossi, C. Cesaretti, A. Piccoli, A. Peris, F. Logi, P. Pasqualetti, and F. Tomaiuolo, “Predict recovery of consciousness,” Neurology, vol. 45, no. 5, pp. 915–918, May 1995.

H. Lee, M. A. Mizrahi, J. A. Hartings, S. Sharma, L. Pahen, L. B. Ngwyenya, B. D. Moseley, M. Privitera, F. C. Tortella, and B. Foreman, “Continuous electroencephalography after moderate to severe traumatic brain injury,” (in English), Crit. Care Med., vol. 47, no. 4, pp. 574–582, Apr. 2019.

R. Yuvaraj, M. Murugappan, N. M. Ibrahim, M. Iqbal, K. Sundaraj, K. Mohamad, R. Palaniappan, E. Mesquita, and M. Satiyani, “On the analysis of EEG power, frequency and asymmetry in Parkinson’s disease during emotion processing,” Behav. Brain Functions, vol. 10, no. 1, p. 12, Feb. 2014.

R. Lehembrue, B. Marie-Aurelle, A. Vanhaudenhuyse, C. Chatelle, V. Cologan, Y. Leclercq, A. Sodu, B. Macq, S. Laureys, and Q. Noirhomme, “Reesting-state EEG study of comatose patients: A connectivity and frequency analysis to find differences between vegetative and minimally conscious states,” Front. Neurol., vol. 27, no. 1, p. 41, 2012.

J.-R. King, J. D. Sitt, F. Faugeras, B. Rohaut, I. El Karoui, L. Cohen, L. Naccache, and S. Dehaene, “Information sharing in the brain indexes consciousness in noncommunicative patients,” Current Biol., vol. 23, no. 19, pp. 1914–1919, Oct. 2013.

U. Lee, S. Blain-Moraes, and G. A. Mashour, “Assessing levels of consciousness with symbolic analysis,” Phil. Trans. Roy. Soc. A, Math., Phys. Eng. Sci., vol. 373, no. 2034, Feb. 2015, Art. no. 20140117.

S. Gudmundsson, T. P. Runnarson, S. Sigurdsson, G. Eiriksdottir, and K. Johnsen, “Reliability of quantitative EEG features,” Clin. Neurophysiol., vol. 118, no. 10, pp. 2162–2171, Oct. 2007.

W. M. Shannon, “A mathematical theory of communication,” Bell Syst. Tech. J., vol. 27, no. 3, pp. 379–423, Jul./Oct. 1948.

S. M. Pincus, “Approximate entropy as a measure of system complexity,” (in English), Proc. Nat. Acad. Sci. USA, vol. 88, no. 6, pp. 2977–2981, Mar. 1989.

S. M. Pincus, “Irregularity and asynchrony in biologic network signals,” (in English), Methods Enzymol., vol. 321, pp. 149–182, 2000.

L. Van Marcke, M. Schauta, K. A. Mottart, K. M. Donnelly, V. Arnedo, and A. C. Grant, “Quantitative analysis of the EEG posterior-dominant rhythm in healthy adolescents,” Clin. Neurophysiol., vol. 119, no. 8, pp. 1778–1781, Aug. 2008.

L. Breiman, “Random forests,” Mach. Learn., vol. 45, no. 1, pp. 5–32, 2001.

I. Hafizi-Rastani, H. Khalili, S. Paydar, and S. Pourrahmad, “Identifying important attributes for prognostic prediction in traumatic brain injury patients,” Methods Inf. Med., vol. 55, no. 05, pp. 440–449, 2016.

S. Pourrahmad, M. Azad, and S. Paydar, “Diagnosis of malignancy in thyroid tumors by multi-layer perceptron neural networks with different batch learning algorithms,” Global J. Health Sci., vol. 7, no. 6, pp. 46–54, Mar. 2015.

P. E. Rapp, D. O. Keyser, A. Albano, R. Hernandez, D. B. Gibson, R. A. Zambon, W. D. Hairston, J. D. Hughes, A. Krystal, and A. S. Nichols, “Traumatic brain injury detection using electrophysiological methods,” Frontiers Hum. Neurosci., vol. 9, p. 11, Feb. 2015.

I. Sánchez Fernández, A. J. Sansevere, M. Gama-Lein, H. Kapur, and T. Loddenkemper, “Machine learning for outcome prediction in electroencephalography (EEG)-monitored children in the intensive care unit,” J. Child Neurol., vol. 33, no. 8, pp. 546–553, Jul. 2018.

J. Schwimm and M. A. Schütze, Electroencephalographic Recordings in Traumatic Brain Injury, 1 ed. Amsterdam, The Netherlands: Elsevier, 2015, pp. 319–339.

Q. D. Bunchlak et al., “Machine learning applications to clinical decision support in neurosurgery: An artificial intelligence augmented systemic review,” Neurosurg. Rev., Vol. 2019, doi: 10.1007/s11060-019-01163-8.

J. B. Sanders, M. Z. Aki, V. Karhade, B. Chang, W. B. Gormley, M. L. Broekman, T. R. Smith, and O. Arnaout, “An introduction and overview of machine learning in neurosurgical care,” Acta Neurochirurgica, vol. 160, no. 1, pp. 29–38, Jan. 2018.

C. W. P. M. Hukkelhoven, E. W. Steyerberg, J. D. F. Habbema, E. Farace, A. Marmarou, G. D. Murray, L. F. Marshall, and A. I. R. Maas, “Predicting outcome after traumatic brain injury: Development and validation of a prognostic score based on admission characteristics,” J. Neurotrauma, vol. 22, no. 10, pp. 1025–1039, Oct. 2005.
N. S. E. Mohd Noor, H. Ibrahim: ML Algorithms and Quantitative EEG Predictors for Outcome Prediction

[71] N. A. Mushkudiani, C. W. P. M. Hukkelhoven, A. V. Hernández, G. D. Murray, S. C. Choi, A. I. R. Maas, and E. W. Steyerberg, “A systematic review finds methodological improvements necessary for prognostic models in determining traumatic brain injury outcomes,” J. Clin. Epidemiol., vol. 61, no. 4, pp. 331–343, Apr. 2008.

[72] F. Harrell, Jr., Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis. Cham, Switzerland: Springer, 2015.

[73] F. E. Harrell, Jr., Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis. Cham, Switzerland: Springer, 2015.

[74] J.-Z. Feng, Y. Wang, J. Peng, M.-W. Sun, J. Zeng, and H. Jiang, “Comparison between logistic regression and machine learning algorithms on survival prediction of traumatic brain injuries,” J. Crit. Care, vol. 54, pp. 110–116, Dec. 2019.

[75] S. J. Russell and P. Norvig, Artificial Intelligence: A Modern Approach. Malaysia, Kuala Lumpur: Pearson, 2016.

[76] T. Marwala, Handbook of Machine Learning. Singapore: World Scientific, 2018, p. 328.

[77] W. Chen, Y. Wang, G. Cao, G. Chen, and Q. Gu, “A random forest model based classification scheme for neonatal amplitude-integrated EEG,” Biomed. Eng. OnLine, vol. 13, no. 2, p. S4, 2014.

[78] L. S. Prichep, S. G. Dastidar, S. G. Dastidar, W. Koppes, J. Miller, T. Radman, B. O’Neil, R. Naunheim, and J. S. Huff, “Classification algorithms for the identification of structural injury in TBI using brain electrical activity,” (in English), Comput. Biol. Med., vol. 53, pp. 125–133, Oct. 2014.

[79] A. T. Hale, D. P. Stonko, L. Wang, M. K. Strother, and L. B. Chambless, “Machine learning analyses can differentiate meningioma grade by features on magnetic resonance imaging,” Neurosurgical Focus, vol. 45, no. 5, p. E4, Nov. 2018.

[80] A. T. Hale, D. P. Stonko, J. Lim, O. D. Guillamondegui, C. N. Shannon, and M. B. Patel, “Using an artificial neural network to predict traumatic brain injury,” J. Neurosurgery: Pediatrics, vol. 23, no. 2, pp. 219–226, 2018.

[81] A. T. Hale, D. P. Stonko, A. Brown, J. Lim, D. J. Voce, S. R. Gannon, T. M. Le, and C. N. Shannon, “Machine-learning analysis outperforms conventional statistical models and CT classification systems in predicting 6-month outcomes in pediatric patients sustaining traumatic brain injury,” Neurosurgical Focus, vol. 45, no. 5, pp. 1–7, 2018.

[82] D. V. Agoston and D. Langford, “Big Data in traumatic brain injury; Promise and challenges,” Concussion, vol. 2, no. 4, p. CNC44, 2017.

[83] T. van der Ploeg, D. Nieboer, and E. W. Steyerberg, “Modern modeling techniques had limited external validity in predicting mortality from traumatic brain injury,” J. Clin. Epidemiol., vol. 78, pp. 83–89, Oct. 2016.

[84] E. W. Steyerberg, M. J. Eijkemans, F. E. Harrell, and J. D. Habbema, “Prognostic modelling with logistic regression analysis: A comparison of selection and estimation methods in small data sets,” (in English), Statist. Med., vol. 19, no. 8, pp. 1059–1079, 2000.