MethodsX 13 (2024) 102765

Contents lists available at ScienceDirect

MethodsX

journal homepage: www.elsevier.com/locate/methodsx

Machine learning and brain-computer interface approaches in prognosis and individualized care strategies for individuals with amyotrophic lateral sclerosis: A systematic review

Stephanie Yen Nee Kew, Siew-Ying Mok, Choon-Hian Goh*

Department of Mechatronics and Biomedical Engineering, Lee Kong Chian Faculty of Engineering and Science, Universiti Tunku Abdul Rahman, 43000 Kajang, Selangor Darul Ehsan, Malaysia

REVIEW HIGHLIGHTS

- The use of machine learning in the application of prognosis of individuals with ALS.
- The use of brain-computer interface in providing individualized care strategies for individuals with ALS.
- Insights into current status, limitations, and future directions of machine learning application in ALS prognosis.

ARTICLE INFO

Method name:
Systematic review methodology

Keywords:
Motor neuron disease (MND)
Autonomic intelligence
Prognostication
Personalized
Intervention

ABSTRACT

Amyotrophic lateral sclerosis (ALS) characterized by progressive degeneration of motor neurons is a debilitating disease, posing substantial challenges in both prognosis and daily life assistance. However, with the advancement of machine learning (ML) which is renowned for tackling many real-world settings, it can offer unprecedented opportunities in prognostic studies and facilitate individuals with ALS in motor-imagery tasks. ML models, such as random forests (RF), have emerged as the most common and effective algorithms for predicting disease progression and survival time in ALS. The findings revealed that RF models had an excellent predictive performance for ALS, with a testing R2 of 0.524 and minimal treatment effects of 0.0717 for patient survival time. Despite significant limitations in sample size, with a maximum of 18 participants, which may not adequately reflect the population diversity being studied, ML approaches have been effectively applied to ALS datasets, and numerous prognostic models have been tested using neuroimaging data, longitudinal datasets, and core clinical variables. In many literatures, the constraints of ML models are seldom explicitly enunciated. Therefore, the main objective of this research is to provide a review of the most significant studies on the usage of ML models for analyzing ALS. This review covers a variation of ML algorithms involved in applications in ALS prognosis besides, leveraging ML to improve the efficacy of brain-computer interfaces (BCIs) for ALS individuals in later stages with restricted voluntary muscular control. The key future advances in individualized care and ALS prognosis may include the advancement of more personalized care aids that enable real-time input and ongoing validation of ML in diverse healthcare contexts.

* Corresponding author.
E-mail address: gohch@utar.edu.my (C.-H. Goh).

https://doi.org/10.1016/j.mex.2024.102765
Received 16 April 2024; Accepted 15 May 2024
Available online 25 May 2024

Specifications Table

Subject area:	Selected from MEX categories.
More specific subject area:	Machine learning, brain-computer interface (BCI)
Name of the reviewed methodology:	Systematic review methodology
Keywords:	Amyotrophic lateral sclerosis (ALS), machine learning (ML), prognosis, individualized, treatment
Resource availability:	Google Scholar, Springer, PubMed, Scopus, Web of Science, IEEE Xplore, and ScienceDirect
Review question:	How is machine learning used in the detecting and predicting the prognosis of individuals with ALS?
	How is brain-computer interface used in providing individualized care strategies for individuals with ALS?
	What is the current status, limitation, and future direction of machine learning application in ALS prognosis?

Background

Motor neuron diseases (MNDs) are a class of neurological illnesses distinguished by the increasing dysfunction of the motor neurons (MNs). Even by neuronal standards, MNs are massive cells with exceptionally long axons that may reach lengths of up to 1 m in an adult individual, making them distinct from other types of neurons. Based on the location, these motor neurons can be classified into two: lower motor neurons and upper motor neurons within the central nervous system. While the upper MNs reside in the motor cortex, the lower MNs are situated in the spinal cord and brainstem [28]. Motor neurons play a role in innervating the skeleton muscle fibres which govern muscle contraction. When degeneration of the motor nerve axon occurs, the innervated muscle fibres would cease to function properly.

In 2019, 1034,606 disability-adjusted motor neuron diseases (DALYs) were lost to motor neuron diseases which resulted in 39,081 fatalities worldwide [26]. According to 2016 Global Burden of Diseases (GBD) estimates, the all-age incidence rates for MNDs were 0.78 per 100,000 person-years [7]. However, among all MNDs, amyotrophic lateral sclerosis (ALS) is the most prevalent, with a mean incidence of 1.8 per 100,000 in North America and 2.8 per 100,000 in Europe [19]. The majority of the cases are sporadic, and characterized by hyperreflexia, muscle weakness, and stiffness. ALS, formerly referred to as Lou Gehrig's disease, is a progressive neurological disease affecting motor neurons, resulting in the degeneration of nerve cells in the brain and spinal cord that regulate breathing and voluntary muscle movement. The impairment of respiratory muscles in ALS individuals has often restricted survival to 2–5 years after illness onset in ALS, which frequently has a localized origin before spreading to various body parts [20]. Overall, the decreased survival rate in ALS individuals has been confirmed as a result of respiratory muscle deficiency, and different assistance methods have varying degrees of effectiveness or success rates in alleviating the effects of this condition.

To lessen the worldwide burden that ALS represents, the prognosis and management techniques need to be enhanced by providing the best treatment at the best possible time. Early prognosis is an extremely crucial yet difficult endeavor in the plight of an effective treatment for ALS. Due to the gradual onset of symptoms and the resemblance of ALS to other neurological disorders, better diagnostic definitions that allow relevant disease categorizations prognostically are required. It has come to an understanding that ALS has emerged as a heterogeneous clinical population and not a single disease entity [16]. These variables may lead to diagnostic delays which may last from 9 to 27 months. Additionally, the uncertainties during ALS detection might predispose an individual to unneeded interventions, resulting in an accelerated disease progression [6]. In the context of this, an artificial intelligence (AI) model can come in handy and make significant strides in prognosis, enabling precise disease stratification and shortening detection times.

Recently, there have been widespread applications of AI-based technologies in healthcare facilities and various AI algorithms have been used by researchers to detect vast sets of diseases. One of the applications that has benefitted from AI is the progressive prediction of bulbar palsy and polytopic paralysis via a neuro-machine learning model [34]. It is recognized that both bulbar palsy and polytopic paralysis are associated with ALS in one way or another. Bulbar palsy is often regarded as a subtype of ALS, and whenever ALS predominantly affects the bulbar area, the manifestation of bulbar palsy may occur. On the other hand, polytopic paralysis is a characteristic feature of the progressive stages of ALS, as it involves widespread motor impairments, similar to what is observed in ALS. Furthermore, AI-based technology has been applied in detecting and distinguishing between cerebral palsy and hereditary spastic paraplegia through the Bayesian additive regression tree (BART) approach using data from clinical gait analysis [4]. Numerous algorithms have demonstrated their value in improving the detection procedures for different medical pathologies. Therefore, the aim of this research is to provide a review of certain well-known AI algorithms and their existing clinical applications in prognosis, detecting, and assisting ALS.

Method details

In this review, a systematic and critical methodology was conducted, which allowed for the synthesis of findings from the selected articles. Through this approach, meaningful inferences about the efficiency of machine learning (ML) and brain-computer interface (BCI) in enhancing the quality of life for ALS individuals can be derived. Additionally, it helps to address the review questions and offers an insightful understanding of the present state of this research field. A summarised flowchart is illustrated in Fig. 1.

Literature search strategy

A comprehensive literature search was performed using databases, such as Google Scholar, Springer, PubMed, Scopus, Web of Science, IEEE Xplore, and ScienceDirect, to identify studies that report the use of machine learning (ML) and brain-computer interface (BCI) approaches in the prognosis of individuals with amyotrophic lateral sclerosis (ALS) and individualized care strategies for them.

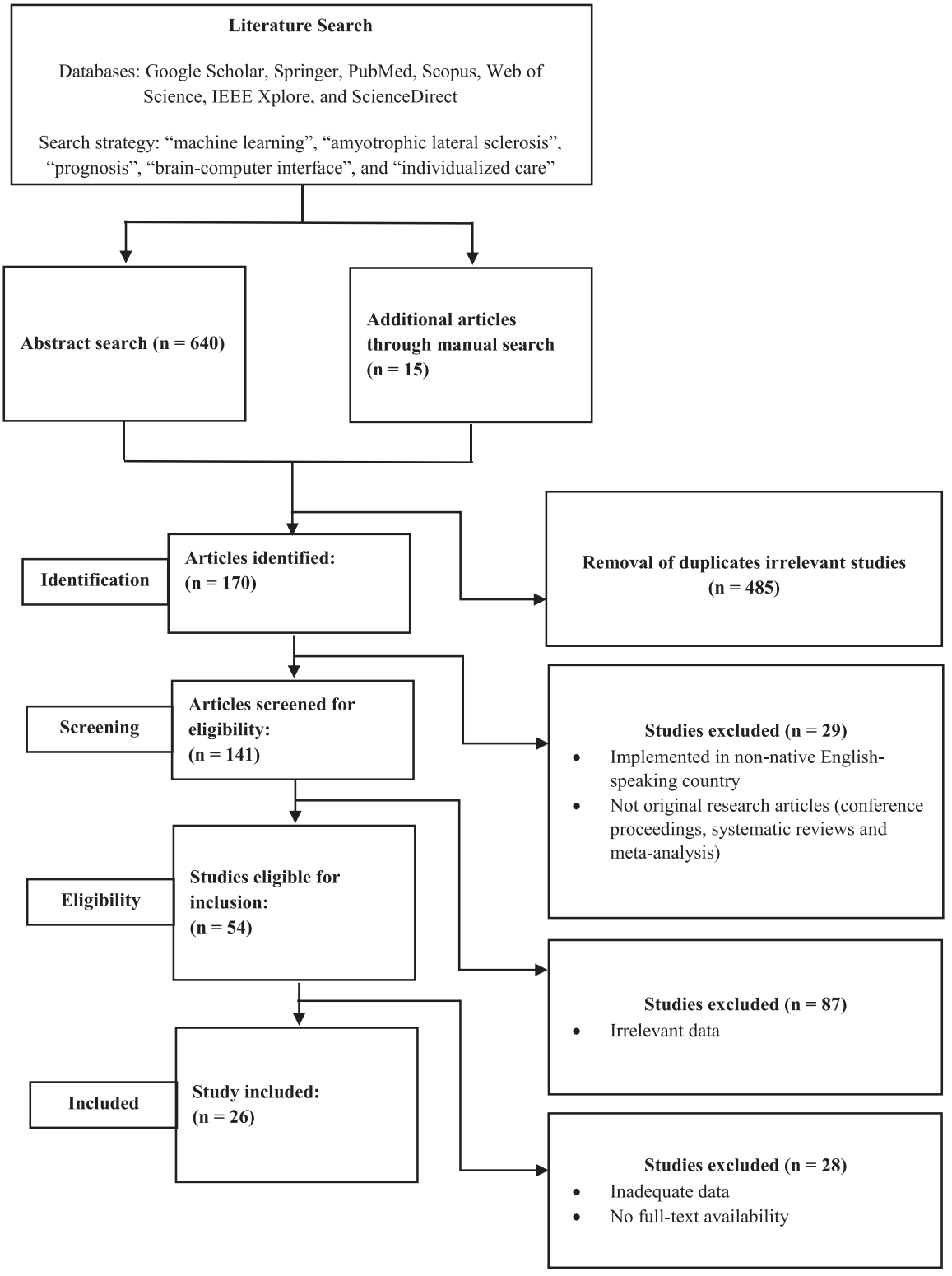

Fig. 1. Study search and selection flow.

A wide variety of search terms: “machine learning”, “amyotrophic lateral sclerosis”, “prognosis”, “brain-computer interface”, and “individualized care” were used in the search strategy to ensure thorough coverage of existing literature. To create a more precise and extensive search result, the Medical Subject Headings (MeSH) terms were employed. Apart from the preliminary database search, the reference lists of the retrieved articles were manually searched to identify additional articles. To determine the relevance and suitability of the articles, both titles and abstracts were carefully screened. Those that appeared relevant were retrieved for further evaluation.

Inclusion and exclusion criteria

Studies were only included when the searched articles focused on using ML and BCI approaches for ALS prognosis and personalized care. This covers articles discussing ALS disease prediction and progression, as well as advanced assistive devices aiding in motor imagery tasks and communication. Meanwhile, conference proceedings, systematic reviews, and meta-analyses are excluded. Only journal articles written in English and published within the last 10 years were included. Overall, this review did not require ethical approval because data were analyzed and synthesized from previously released articles. Any discrepancies were resolved by consensual judgment among reviewers.

Data extraction and analysis

A 2-tiered eligibility review was conducted by the reviewers. Initially, the article abstracts were screened and assessed; studies that did not meet all the predefined inclusion criteria or any exclusion criteria were excluded. This preliminary evaluation was carried out for quick exclusion of articles that are not in line with the objectives of this study. In cases where the abstract was insufficient to assess for eligibility, the full article was reviewed. Data from eligible journals were then synthesized, analyzed, and cross-checked. In general, relevant information was extracted from the chosen articles using a standardized form known as the data extraction form. Various data were extracted, including participant characteristics (sample size), datasets (origin and type), interventions (both ML and BCI approaches involved), and main findings related to prognosis and individualized care for ALS patients. Out of the original 655 studies identified, 485 studies were removed due to duplications and irrelevancy. The remaining 170 articles were evaluated; 54 met the eligibility criteria. Eventually, 28 studies were excluded since the data was inadequate or not available in full text. Hence, only 26 articles were included for analysis.

Results and discussion

Machine learning in amyotrophic lateral sclerosis prognosis

Early prognosis is extremely challenging due to variability in symptoms and progressive nature. It is also a challenge when conducting a direct performance of model performance in ML studies focused on ALS because of the factors of substantial differences in sample sizes, performance metrics, prediction goals, and study design across the numerous findings. For this, merely relying on clinical features or clinical data supplemented with other input data types to improve model performance is not well-supported by the evidence at hand. Nonetheless, machine learning holds great promise for identifying complex ALS-associated diseases. The key aim of prognosis and prediction for any kind of disease is to reduce the likelihood of an illness becoming more severe or deadly by optimizing interventions. While typical conventional statistical approaches have historically been used for predicting ALS prognosis [17,29,30], ML models can offer hitherto unheard-of prospects for discovering new prognostic indicators [1,2,12,18,25,27,32,38,41].

Prognosis can typically be conceptualized as a regression problem with a predefined duration of survival or a classification problem with specified categories. Frequently, the classification approach is classified as either survival (life expectancy) [27,32] disease phase (stage or progression of disease) [1], or functional decline (alterations in health) [18,24]. Meanwhile, as observed in some current studies, the regression approach utilized is geared toward predicting specific outcomes for instance, functional decline [12,37], survival rates [2,38] or respiratory function [15]. Numerous existing literature in this domain have showcased its performance in clinical settings, whether the prognostic modeling, is approached as a classification problem or regression problem. One of the most accurate classification techniques was observed at a 66 % accuracy rate [18] whereas, the regression method with the highest accuracy has a Root Mean Squared Error (RMSE) of 0.52 [12], both using RF algorithm. In terms of outcome prediction as a classification problem, the best performance was attained through the multivariate Royston-Parmar model, with 78 % accuracy. On the other hand, for outcome prediction as a regression problem on functional decline and survival, the result could be yielded by the innovative RF approach [33]. Table 1 summarised the studies applied machine learning in amyotrophic lateral sclerosis prognosis.

Prediction of disease progression in ALS individuals

There are present findings that have developed prognostic models by adopting Pooled Resource Open-Access ALS Clinical Trials (PRO-ACT) data [25,33]. As presented by Pancotti et al. [25], the combination of different architectures namely, feed-forward neural network (FFNN), and recurrent neural network (RNN) with the PRO-ACT datasets were employed to predict ALS disease trajectory. In this context, PRO-ACT dataset which is recognized as one of the most renowned repositories consisting of ALS clinical trial data was merely used. Nevertheless, owing to its sources collection or patient population, it might contain some inherent biases or limitations that could affect the model’s generalizability and accurate predictions. This is a result of temporal or spatial variability in PRO-ACT

Table 1
Overview of the research on machine learning in amyotrophic lateral sclerosis prognosis.

Reference	Dataset(s) origin	Dataset(s) type	ML approach	Application	Effectiveness
Beaulieu-Jones & Greene, [2]	Electronic Health Records (EHRs) from patient interactions with healthcare providers	Real-life data	Semi-supervised learning; Random forests (RF)	Improvement for classification; survival prediction in the context of ALS	The hidden layer of the deionizing autoencoder (DA) had a higher effectiveness basis for ALS patient survival prediction
Pfohl et al., [27]	Data from 801 deceased ALS patients	Clinical metrics	Generalized linear model (GLM); RF model	Identification of the "best" survival predictors; comprehension of how their predictive capability alters with disease progression	GLM demonstrated a mixing parameter of α of 0.01 and a regularization parameter λ of 6.629; RF model exhibited a testing R^2 of 0.524, suggesting a considerably strong predictive performance
van der Burgh, [38]	The dataset originated from 135 ALS patients	Real-life data	Deep learning; artificial neural networks (ANNs)	Prediction of the survival categories (short, medium, or long survivors)	Correct predicted survival category in 68.8 % of cases; deep learning prediction accuracy increased to 84.4 %
Schuster et al., [32]	The dataset originated from 60 ALS patients and 69 healthy controls	Real-life data	Binary logistic ridge regressions	Prediction of 18-month survival ALS patients	A combination of both MRI and clinical data had an improved accuracy of survival predictions, reaching up to 79.17 %
Pancotti, et al., [25]	Pooled Resource Open-Access ALS Clinical Trials (PRO-ACT) repository	Clinical trial dataset	Deep learning; feed-forward neural network (FFNN); recurrent neural network (RNN)	Prediction of disease progression, measured through the slope of the ALSFRS score between 3 and 12 months	Experimental ALSFRS slope possessed a weak Pearson Correlation Coefficient (PCC) of 0.2 with the time-to-death; log-rank test p-value was lower than 0.005, indicative of robustness for the observed survival differences
Seibold et al., [33]	Data derived from PRO-ACT database	Clinical trial data	Model-based RF	Prediction of treatment effects for patients with ALS	For the survival time, the minimum predicted treatment effect was 0.0717
Bandini et al., [1]	Data from 192 recordings from 64 patients with ALS	Real-life data	Feature selection and classification algorithm	Prediction of the disease progression	High accuracy of 87 % in differentiating differences between presymptomatic and symptomatic phases of bulbar decline
Ko et al., [18]	PRO-ACT repository	Real-life data	RF Classifier	Prediction of the progression of ALS	Accuracy of 66 % in the prediction of the ALS progression
Ong et al., [24]	PRO-ACT repository; de-identified records from ALS sufferers	Clinical trial data	Boosting algorithm	Prediction of ALS disease course (fast or slow progression)	Predicted functional decline class (fast or slow) had a fair accuracy (area under the curve (AUC) = 0.82))
Taylor et al., [37]	PRO-ACT ALS database	Clinical trial dataset	RF model	Prediction of ALS disease progression	RF model outperformed the pre-slope model at the initial time points; RF model performance was similar to a GLM model in the aspects of root-mean-square deviation; RF model outcome was the greatest when reaching the later time points
Hothorn & Jung, [12]	PRO-ACT database; 1197 ALS patients	Real-life data	RF algorithm	Prediction of disease progression in ALS patients	This method ranked third best in terms of prediction accuracy; the variance of the random slope (individual variability) was ~0.23
Jahandideh et al., [15]	PRO-ACT database; over 10,000 ALS patient records	Real-life; clinical trial data	Gradient boosting machine (GBM); RF model	Longitudinal prediction of vital capacity in ALS patients	All models showed similar accuracy for predicting support vector machine (SVM)
Huang, et al., [13]	PRO-ACT ALS database	Clinical trial data	RF regression model	Survival prediction; non-parametric ranking of patients' survival data	Higher accurate ranking predictions were observed on the PRO-ACT ALS dataset in comparison to the Cox proportional hazard model
Gomeni & Fava, [8]	PRO-ACT ALS database	Clinical trial data	Non-linear Weibull model	Prediction of disease progression; estimation of clusters of trajectories in ALS progression over 12 months of treatment	Identified two clusters of trajectories; 46 % of patients with slow disease progression; and 54 % of patients with fast disease progression

(*continued on next page*)

Table 1 (*continued*)

Reference	Dataset(s) origin	Dataset(s) type	ML approach	Application	Effectiveness
Westeneng et al., [41]	Data obtained from 14 specialized ALS centers	Clinical trial data	Multivariable Royston-Parmar model	Prediction of a composite survival endpoint	The model revealed remarkable external predictive accuracy (c statistic: 0.78) and calibration (calibration slope: 1.01)
Likhachov et al., (2021)	Vocal data from the Republican Research and Clinical Center of Neurology and Neurosurgery	Real-life data	Linear discriminant analysis (LDA)	Prediction of disease progression based on the acoustic analysis of the voice	Identified various parameters, for example, shimmer, jitter, degree of vibrato pathology, and many more, which can be pathological and serve as a sign of the progress of ALS
Imamura et al., [14]	Data from induced pluripotent stem cells (iPSCs) of both healthy control and ALS patients	Real-life data	Deep learning algorithm; convolutional neural network (CNN)	Classification of individuals into healthy control and ALS categories	Exceptional AUC of 0.97 for the classification between healthy control and ALS
Greco et al., [9]	692 blood data	Clinical dataset	SVM with Recursive feature elimination (SVM-RFE)	Differential prognosis of ALS	Average accuracy of 94 % using blood data, exclusively

repository due to differences in ALS disease progression rates across varying periods. In the study of Seibold et al., [33], the reliance on PRO-ACT data with a model-based RF approach to predict treatment effects in ALS has added credibility to its finding because of extensive datasets involved for validation but could be further strengthened by delving into a wider range of patient characteristics.

The most commonly employed prognostic model in ALS was identified to be RF to assess variable importance due to its predictive power, adaptability, and robustness in handling high-dimensional datasets, as featured in numerous existing bodies of work [2,13,15,27,37]. Moreover, it is proven among the best-performing techniques [2,13,27,37]. In training RF algorithms to predict the course of ALS disease progression, multiple analyses can be observed utilizing the electronic clinical record data including, medical history, and demographics as presented by Ko et al., [18]; Hothorn & Jung, [12]; and Taylor et al., [37]. The study by Ko et al., [18] implemented a cloud computing system with an RF classifier of Apache Mahout capable of rapid and precise prediction of ALS progress. The resulting model was able to predict the course of ALS with a prediction accuracy of 66 % but, this level of accuracy might be deemed as modest and may not be sufficiently reliable for clinical use. In this context, the inclusion of more comprehensive ALS individual records is crucial to enhance the predictive accuracy power of the model in several ways including, improved model generalization and better refined individual stratification.

Furthermore, the work by Hothorn & Jung, [12] suggested that in terms of prediction accuracy measured by the ALS functional rating scale (ALSFRS) with RF algorithm, where past disease progression would be an ideal marker of future disease risk and severity. Essentially, there is a close association between more rapid future progression of diseases and higher initial ALSFRS score fluctuation. Despite that, the ability of the model to predict the development of ALS over a longer time could be compromised by its short-term (3 months) prediction, which may not account for prolonged-term or more complicated disease trajectories. For a more comprehensive understanding of ALS disease progression patterns, incorporating time-series analysis methods or adjusting the architecture of the ML model that handles extended durations are some of the strategies for tackling the shortcomings of the first three-month interval of prediction. In essence, RF may offer a far superior prediction than the pre-slope model. The discovery by [37] explained that an RF model employing baseline data could potentially predict illness development accurately for both populations receiving medical attention at a tertiary care clinic and with the clinical trial dataset. A lot of referral centers generate predictive models making use of local datasets [32,38] rather than population-based data since they are more accessible. Furthermore, using datasets that are specific to the local context could provide a more representative of the disease progression patterns observed within the particular ALS individual cohort. However, the accessibility of population-based data is growing because of the regional [31], national [36], and international consortia registries [22].

The study by Imamura et al., [14] demonstrated the benefits of using deep convolutional neural network (CNN) in predicting ALS by making use of the induced pluripotent stem cells (iPSCs)-derived motor neuron images. The model sought to distinguish between ALS individuals with healthy controls through the image analysis of spinal motor neurons produced by iPSCs. As illustrated by its excellent area under the curve (AUC) of 0.97 results, this CNN algorithm performed exceptionally well, indicating high accuracy in the differentiation between different groups of motor neurons. Although promising, further validation would be helpful for a better transition of an algorithm from iPSC-derived neuronal images to actual clinical settings. To develop a comprehensive diagnostic solution, augmenting imaging results with clinical information, for example, genetic data, medical history, and diagnostic markers is considered necessary. Moreover, as described by Greco et al., [9], a support vector machine (SVM) integrated with a recursive feature elimination (RFE) algorithm with the usage of blood data to identify whether the ALS individual has a fast or slow disease progression. Through this architecture, the accuracy level can be maximized while simultaneously lowering the dimension of the dataset. With blood data alone, it was found that significant mean accuracy rates of as high as 94 % were attained in predicting ALS prognosis. While the study identified potential biomarkers via the use of blood samples, validation using other analytical techniques or longitudinal investigations could be useful to support the importance and practicality of these biomarkers in prognosis prediction.

Survival prediction in ALS individuals

Recent literature demonstrated examples of ML algorithms that exhibit potentiality in predicting and improving the survival outcomes of individuals affected by ALS [2,27]. A finding by Beaulieu-Jones & Greene, [2] revealed the use of a semi-supervised learning approach for the extraction of Electronic Health Record (EHR) phenotype with denoising autoencoders, allowing the model to leverage both labeled and unlabeled data. In this study, although the survival prediction in ALS clinical trial data was enhanced by the application of denoising autoencoders incorporated with random forests (RF) however since it dealt with only limited high-quality phenotype samples, further validation on more diverse and larger datasets is required otherwise, there may be a tendency of inaccuracies or biases during the phenotype extraction thereby, affecting the downstream analyses. Comparatively, the generalized linear regression model constructed by Pfohl et al., [27] too focuses on a large dataset of deceased ALS individuals, but the incorporation of various metrics for instance, demographic factors, muscle assessments, respiratory functions, and Amyotrophic Lateral Sclerosis Functional Rating Scale (ALSFRS-R) scores have made an extensive evaluation of survival predictors possible. While this study has identified the key predictors, it could also have investigated other factors that might have affected ALS survival span, for instance, environmental impacts or genetic markers to further improve the predictive model.

Often, prognosis in ALS is determined by clinical characteristics in prognostic models however, two recent studies use imaging measures to enrich their clinical data [32,38]. Based on van der Burgh, [38], artificial neural networks (ANNs) were trained to predict the survival time of ALS individuals by utilizing clinical characteristics and advanced magnetic resonance imaging (MRI) metrics. By employing ANN, the subtle correlations among various predictors can be captured for predicting ALS survival times with higher accuracy. Not only can this amalgamation with MRI metrics be able to improve the predictive accuracy of ALS also, it can lead to a more reliable prognostic assessment, providing a holistic view of the disease. Similarly, Schuster et al., [32] studied an 18-month ALS survival prediction via binary logistic ridge regression that makes use of the diffusion tensor white matter and surface-based morphometric parameters. The relevance of MRI metrics combined with clinical measures utilized in this study has the capability of providing insights into neurodegeneration patterns and potentially improving the predictive power with an accuracy of 79.17 %. However, the generalizability of the findings may be affected due to the small sample size where only 60 ALS individuals were involved. Moreover, the study's scope was constrained to predicting survival outcomes within an 18-month timeframe, which could be expanded for the exploration of additional nuances, patterns, or shifts observed in ALS disease progression. Table 2 summarised the studies on machine learning to assist ALS individuals.

Model performance has been assessed using a range of metrics. The mean absolute error, RMSE, and Pearson correlations between actual and projected values are commonly reported for regression techniques. In ALS, regression models including generalized linear models [13,27], non-linear Weibull models [8] and Royston-Parmar models [41] have also been widely implemented since their versatile framework enables capturing of the complex and multifaceted nature of ALS conditions. Despite their restrictive assumptions, regression models offer a lot of promise for use in therapeutic settings [41]. On top of that, RF regression can potentially be selected as the survival prediction model with the ability to outperform the Cox model, 40 times out of all 40 rounds of tests based on the finding of Huang et al., [13].

ML algorithms to assist ALS individuals

Thus far, ALS remains without a cure or effective treatment, however, machine learning has been harnessed in aiding individuals affected by ALS. It is essential for one to bear that ALS is characterized by progressive muscular paralysis indicative of motor neuron degeneration, and gradual spasticity that can frequently lead ALS individuals to enter into a locked-in condition where the subject is unable to move independently. Therefore, here comes the brain-computer interfaces (BCIs), sometimes referred to as the direct neural interface, which is helpful in such circumstances. This type of interface is envisioned to provide channels of communication that are independent of muscle activity by enabling the use of brain activity to operate an external assistive device. The BCI typically transforms brain activity into computer commands, allowing users to perform tasks like choosing letters to compose words on a digital keyboard. An ALS individual will then be able to restore their voice function by vocalizing these phrases artificially [39].

Numerous studies have been conducted, examining the application of BCI and its effectiveness for ALS individuals [3,5,10,11,21,23,35,40]. In the clinical study of Guy et al., [10], 20 ALS individuals were involved for the evaluation of the visual BCI device by integrating electroencephalography (EEG). These EEG signals are efficiently processed and analyzed in real time via machine learning algorithms. The ML-based algorithms are combined with both the flashing sequence optimization and the optimum stopping of flashes to improve information transmission efficiency. Through learning from user behavior over time, ML models has the ability to continuously adapt and enhance the accuracy of word predictions. The system was effective in that 65 % of participants selected the correct symbols with a 95 % accuracy. However, despite its effectiveness, this study involved a limited participant pool, necessitating a bigger sample size to validate and generalize the usability of the system.

Furthermore, Sorbello et al., [35] also proposed the usage of EEG-BCI with biofeedback factor (Bf) generated from users' intention, attention, and focus. In this context, the architecture incorporates ML algorithms in recognizing the mental states of the user and extracting patterns of interest from EEG. This information is further utilized to initiate customized responses in a humanoid robot, thereby creating a more responsive and personalized interaction. The outcome demonstrated a 96.67 % success of ALS subjects completing the tasks, with a Bf of 81.20 %. Among certain ALS patients during tasks, there is an exclusion of one ALS patient due to low general attention. Undoubtedly, this can affect overall engagement with the EEG-BCI system. Hence, to maintain focus throughout tasks, it is recommended to implement real-time feedback mechanisms within the system wherein, the users are alerted when their attention levels drop. In relation to this, a finding by Wang et al., [40] demonstrated the detection of the user's attention levels and

Table 2
Overview of research on ML algorithms to assist ALS individuals.

References	Sample size	ML optimization	Functions	Features	Main Findings
Guy et al., (2017)	20 patients with ALS	Efficiently analyze EEG signals; improve the information transmission rate; enhance word prediction accuracy	EEG acquisition; real-time processing; spelling tasks	Optimal stopping of flashes; word prediction	65 % selected >95 % of correct symbols; the average number of correct symbols selected per minute ranged from 3.6 (without word prediction) to 5.04 (with word prediction)
Sorbello et al. [35]	4 ALS patients in a near locked-in status; 4 healthy control subjects	Recognize the mental states of the user; extract desired patterns from EEG	Facilitate human-humanoid interaction; recognize users' mental states through biofeedback factor (Bf); elicit customized behaviors from a humanoid robot	Real-time recognition of users' mental states; Bf calculation	3 out of 4 ALS patients completed the task with 96.67 % success; healthy controls achieved 100 % success; ALS subjects had a higher Bf (81.20 %) compared to healthy controls (76.77 %)
Wang et al., [40]	10 participants	Recognize target objects; interpret user's intentions into actionable commands	Detection of user's attention levels; track user's gaze points; recognize target objects based on intention information	Single-channel EEG recorder; eye-tracking glasses; You Only Look Once v5 (YOLOv5) model	Experimental results showed a high success rate of 89.3 %
Okahara et al., [23]	3 late-stage ALS patients	Recognize distinguished neural patterns; allows users to maintain consistent control over BCI	Evaluation of ALS patient's ability to control neural prostheses	Steady-state visual evoked brain potentials; flickering green and blue light-emitting diodes	All individuals reliably operated the system with a median accuracy of 83.3 % over 27 months
Borgheai et al., [3]	6 participants of ALS	Enable selection of ideal channels and hemodynamic features from fNIRS data	Select optimal channels and analyze hemodynamic features	fNRIS-VM; P3S; statistical parametric mapping; linear discriminant analysis (LDA)	The fNIRS-VM protocol indicated a mean accuracy of 81.3 %±5.7 %; fNIRS-VM was far superior to the P3S (74.0 %±8.9 %) within short times (< 4 s); fNIRS-VM has greater accuracy (73.2 %±2.0 %) over the P3S (61.8 %±1.5 %) in longitudinal data, especially when the subjects were in the late locked-in state
Miao et al., [21]	18 ALS patients	Optimize BCI display approach; reduce errors caused by flashing adjacent stimuli	Validate an alternate ERP-based BCI display approach for communication	Conventional matrix speller paradigm (Matrix-P); new speller paradigm with peripherally distributed stimuli (Peripheral-P)	12 subjects demonstrated >80 % feedback accuracy during online performance; 7 of these subjects secured 90 % higher accuracy
de Oliveira Junior et al., (2018)	10 volunteers	Enhance real-time processing of brainwave signals	Communication between smart devices; assist ALS patients with predefined tasks	Digital assistant brainwave sensors; Internet of Things (IoT); Internet of Health Things (IoHT)	The experienced group had a mean of 66.67 % correct responses; the group without experience achieved a mean of 48.89 %
Hosni et al., [11]	8 patients with ALS	Evaluate the distinct spatial activation patterns	Quantification of subject-specific spatio-temporal characteristics of ALS patients' hemodynamic responses during motor-imagery tasks	Generalized linear model (GLM); linear support vector machine (SVM)	The average classification accuracy achieved was 85.4 %±9.8 %

gaze points with the aid of BCI using a single-channel EEG. By leveraging ML, the system is better able to recognize target objects and translate user's intentions into executable commands, attained by training on relevant data. In experimental trials, a high success rate of 89.30 % was obtained, which attested to its reliability and effectiveness as a useful assistive technology that caters to the specific needs of ALS individuals. Notwithstanding, given that the system depends on eye-tracking glasses and a single-channel EEG recorder, it may impact the accuracy of interpreting the user's intentions. Therefore, one plausible solution is to supplement the information

from eye-tracking glasses and EEG recorders with data acquired from additional modalities and sensors, for example, muscle signals or facial expressions.

The study of Okahara et al., [23] revealed the utilization of BCI system with steady-state visual evoked brain potentials to determine the neural prostheses in late-stage ALS individuals who progressed from locked-in syndrome (LIS) to a completely locked-in state (CLIS). With the aid of ML, the system adapted to each patient's specific neural patterns, giving them to maintain consistent control over the device for an extended duration. Results showed a median accuracy of 83.30 % was achieved throughout the course of 27 months, indicating the ability of all participants to reliably control the BCI system. Even though the 27-month evaluation duration is noteworthy, a more robust understanding of the sustained effectiveness can be accomplished by conducting longer-term studies with a higher number of subjects with a more diverse cohort. Meanwhile, Borgheai et al., [3] focus on the evaluation of ALS individuals, particularly, those in the later stages of the disease. The application of linear discriminant analysis (LDA), alongside correlation analysis and statistical parametric mapping, has potentially enabled the selection of ideal channels and hemodynamic features from the fNIRS data. The outcomes suggest an average accuracy of 81.30 %, outperforming the conventional EEG-based multi-trial P3Speller (P3S). Similar to Okahara et al., [23], it is necessary to focus on performing a longitudinal study where a more prolonged testing procedure should be involved, rather than relying merely on short assessment times (<4 s). This is particularly crucial in real-world scenarios, to gain a more accurate representation of the system's overall validity and effectiveness.

According to the finding in Miao et al., [21], an ERP-based BCI was used to validate a new speller paradigm featuring peripheral stimuli. The BCI display approach was optimized by the ML algorithms, presenting stimuli close to the periphery of the display thereby, reducing errors triggered by flashing adjacent stimuli. In this assessment, notably, a total of 12 subjects attained a feedback accuracy of greater than 80 % during online performance. However, to improve the overall system performance further, iterative testing and design methods that centered around user feedback can be conducted. In addition, the work of de Oliveira Junior et al., [5] presented a BCI system integrated with IoHT to help ALS individuals complete predefined activities by employing a personal digital assistant brainwave sensor. In this intervention, the recognition of eye blinking is to be transmitted to the ALSHelp programme and the real-time processing of brainwave signals was enhanced by ML algorithms. Comparing the experienced group to the inexperienced group, the former scored an average of 66.67 % of correct answers, whereas the latter obtained a mean of 48.89 % which confirmed the efficiency of the application. Although the initial tests were promising, gradual improvements may be facilitated by establishing a feedback loop where users can give their input on the system's functionality and performance.

Referring to the analyses conducted by Hosni et al., [11], the subject-specific spatio-temporal characteristics of ALS individuals were quantified using motor-imagery (MI)-based BCIs. Essentially, the use of a generalized linear model (GLM) in this study offered an approach to estimate and evaluate the unique spatial activation patterns for every ALS individual. The distinctive characteristics that distinguish one ALS patient from another were achieved through statistical optimization of channel sets, contributing to a nuanced comprehension of brain function during motor imagery tasks. Besides that, the linear support vector machine classifier was employed to determine the efficiency of ALS individuals in controlling a binary BCI. This finding demonstrated an average accuracy of 85.40 % and underscored the significance of data-driven approaches for BCI performance optimization. However, this finding is based on a relatively small sample size, involving only 8 ALS subjects, which may limit the generalizability of the outcome. Therefore, it is vital to broaden the cohort of ALS individuals, encompassing a diverse range of demographic characteristics and including patients at varying stages of disease progression.

Current status of ML application in ALS prognosis

ALS is a rapidly progressive neurodegenerative disease with currently limited prognostic and therapeutic options. In comparison to conventional methods in ALS prognosis, the ML models have a higher tendency to achieve a better performance when handling intricate, and high-dimensional datasets, adding to the growing corpus of literature that compares ML with traditional approaches [3]. Based on the existing literature, ML algorithms such as neural networks, SVM, and RF can potentially outperform conventional models. It is renowned that ML techniques have been applied in the analysis of various biomarkers which are often practiced in clinical settings such as imaging [14,32,38], clinical [25,33], and demographical data ([12,18]; and [27,37]). Hence, ML is known for its meaningful contributions to ALS prognosis and for enhancing ALS patients' quality of life. The successful main distinct prognosis identified can be classified into survival time and disease progression. After reviewing numerous studies, ML is demonstrated to be promising and could be advantageous in the development of a decision support system for the prognosis of ALS.

Limitations and future directions

The selection of the most appropriate ML algorithm for prognostic assessment should align with the specific research question and the underlying nature and features of the datasets being utilized. Significant factors to take into account are sample size, dataset balance, and number of variables. For example, adopting neural networks might result in overfitting problems in contexts when the population size is small. On the flip side, the SVM techniques have the advantage of decreased dimensionality, but careful kernel number calibration is required. In many studies, a possible constraint on the generalizability of the results is majorly due to the relatively small sample sizes that are employed. Furthermore, the lack of external validation is another limitation, as most findings solely rely on basic internal validation. Most of the current focus on specific biomarkers or features in ALS prognosis typically might have missing critical facets of the disease.

To increase the applicability and robustness of the established ML models, future initiatives should consider including ALS individuals from diverse demographic characteristics and cohorts that have different disease progression stages. Some of the prospective

directions may involve improving the utility of ML models in a clinical context by including validation across varied health settings. On top of that, to improve prognostic accuracy and gain a more comprehensive insight into ALS progression, a wider range of biomarkers and features should be included, such as exploring additional clinical variables, genetic markers, and neuroimaging data. In addition, further emphasis is needed on fostering a more personalized interaction between the ML technology and the ALS patients it seeks to assist by implementing a feedback loop that allows the users to provide real-time input, so timely adjustments can be made to the system.

Incorporating state-of-the-art techniques, including transfer learning, semi-supervised learning, deep learning, and modal-based RF algorithm, all can potentially be useful in further enhancing the performance, accuracy, interpretability, and clinical applicability of ML models for ALS prognostic assessment. For instance, transfer learning can be particularly helpful in dealing with limited ALS datasets, as it has the ability to improve performance on smaller or more specialized datasets by leveraging pre-trained models on larger datasets. Moreover, modelling complex relationships among data could benefit from employing graph neural networks, either brain connectivity networks or the widely known protein=protein interaction networks, to better understand the disease mechanisms occurring within ALS. Apart from boosting algorithms, approaches like multivariable Royston-Parmar models can present farther avenues for optimizing the overall design and analysis of various clinical trials for ALS.

Conclusion

The application of ML algorithms in prognosis and as an assistive tool for individuals with ALS offers significant opportunities for advancements. This review revealed random forests as the most popular ML model for predicting ALS prognosis, predominantly in the context of both survival time and disease progression. Across multiple ML techniques, variations were observed in the reported performance, accounting for complexities associated with disease progression itself and the diversity of the data involved. In the effort to enhance the applicability and utility of ML algorithms in prognostic studies, additional methodological work is required, particularly in terms of incorporating external validation and sample size expansion. Additionally, the integration of ML to BCI-powered assistive devices can help empower ALS patients to communicate and perform motor-imagery tasks with greater ease and precision. Despite several limitations, the current review highlights the tremendous potential of ML in revolutionizing prognosis and individualized ALS care, underlining the imperative of ongoing research to unlock its full capacity.

Supplementary material *and/or* additional information [OPTIONAL]

A graphical abstract was submitted as supplementary material along with my article.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Stephanie Yen Nee Kew: Conceptualization, Methodology, Data curation, Writing – original draft, Writing – review & editing. **Siew-Ying Mok:** Supervision, Writing – review & editing. **Choon-Hian Goh:** Supervision, Writing – review & editing.

Data availability

No data was used for the research described in the article.

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

[1] A. Bandini, J.R. Green, J. Wang, T.F. Campbell, L. Zinman, Y. Yunusova, Kinematic features of jaw and lips distinguish symptomatic from presymptomatic stages of bulbar decline in amyotrophic lateral sclerosis, J. Speech Lang. Hear. Res. 61 (5) (2018) 1118–1129.
[2] B.K. Beaulieu-Jones, C.S. Greene, Semi-supervised learning of the electronic health record for phenotype stratification, J. Biomed. Inform. 64 (2016) (2016) 168–178.
[3] S.B. Borgheai, J. McLinden, A.H. Zisk, S.I. Hosni, R.J. Deligani, M. Abtahi, … Y. Shahriari, Enhancing communication for people in late-stage ALS using an fNIRS-based BCI system, IEEE Trans. Neural Syst. Rehab. Eng. 28 (5) (2020) 1198–1207.
[4] M.A. Bruce, C.L. Kristen, S.K. Alan, K.M. Lynne, S.H. Michael, Discrimination between hereditary spastic paraplegia and cerebral palsy based on gait analysis data: a machine learning approach, Gait Posture 98 (2022) 34–38.
[5] W.G. de Oliveira Junior, J.M. de Oliveira, R. Munoz, V.C. de Albuquerque, A proposal for internet of smart home things based on BCI system to aid patients with amyotrophic lateral sclerosis, Neural Comput. Appl. 32 (2) (2020) 11007–11017.
[6] P. Francesca, S. Arianna, L. Giancarlo, P.E. Federica, Predictors of diagnostic delay in amyotrophic lateral sclerosis: a cohort study based on administrative and electronic medical records data, Amyotrophic Lateral Sclerosis Frontotemporal Degenerat. 20 (3–4) (2019) 176–185.
[7] GBD 2016 Motor Neuron Disease Collaborators, Global, regional, and national burden of motor neuron diseases 1990-2016: a systematic analysis for the global burden of disease study 2016, Lancet Neurol. 17 (2018) 1083–1097.

[8] R. Gomeni, M. Fava, Amyotrophic lateral sclerosis disease progression model, Amyotrophic Lateral Sclerosis Frontotemporal Degenerat. 15 (2013) 119–129.
[9] A. Greco, M.R. Chiesa, I.D. Prato, A.M. Romanelli, C. Dolciotti, G. Cavallini, … P. Bongioanni, Using blood data for the differential diagnosis and prognosis of motor neuron diseases: a new dataset for machine learning applications, Sci. Rep. 11 (3371) (2021).
[10] V. Guy, M.-H. Soriani, M. Bruno, T. Papadopoulo, C. Desnuelle, M. Clerc, Brain computer interface with the P300 speller: usability for disabled people with amyotrophic lateral sclerosis, Ann. Phys. Rehabil. Med. 61 (1) (2018) 5–11.
[11] S. Hosni, S. Borgheai, J. McLinden, Y. Shahriari, An fNIRS-based motor imagery BCI for ALS: a subject-specific data-driven approach, IEEE Trans. Neural Syst. Rehab. Eng. 28 (12) (2020) 3063–3073.
[12] T. Hothorn, H.H. Jung, RandomForest4Life: a random forest for predicting ALS disease progression, Amyotrophic Lateral Sclerosis Frontotemp. Degenerat. 15 (5–6) (2014) 444–452.
[13] Z. Huang, H. Zhang, J. Boss, S.A. Goutman, B. Mukherjee, I.D. Dinov, Y. Guan, Complete hazard ranking to analyze rightcensored data: an ALS survival study, PLoS Comput. Biol. 13 (12) (2017) e1005887.
[14] K. Imamura, Y. Yada, Y. Izumi, M. Morita, A. Kawata, T. Arisato, … H. Inoue, Prediction model of amyotrophic lateral sclerosis by deep learning with patient induced pluripotent stem cells, Ann. Neurol. 89 (6) (2021) 1226–1233.
[15] S. Jahandideh, A.A. Taylor, D. Beaulieu, M. Keymer, L. Meng, A. Bian, … D.L. Ennist, Longitudinal modeling to predict vital capacity in amyotrophic lateral sclerosis, Amyotrophic Lateral Sclerosis Frontotemp. Degenerat. 19 (3–4) (2018) 294–302.
[16] R. Jeffrey, S.J. Michael, Challenges in the understanding and treatment of amyotrophic lateral sclerosis/motor neuron disease, Neurotherapeutics. 12 (2) (2015) 317–325.
[17] J.A. Knibb, N. Keren, A. Kulka, P.N. Leigh, S. Martin, C.E. Shaw, … A. Al-Chalabi, A clinical tool for predicting survival in ALS, J. Neurol. Neurosurg. Psychiatry 87 (12) (2016) 1361–1367.
[18] K.D. Ko, T. El-Ghazawi, D. Kim, H. Morizono, Predicting the severity of motor neuron disease progression using electronic health record data with a cloud computing Big Data approach, IEEE Sympos. Comput. Intell. Bioinform. Comput. Biol. (2014) 1–14.
[19] G.I. Leslie, R.A. Guy, R. John, C.R. Neil, Clinical spectrum of amyotrophic lateral sclerosis (ALS), Cold Spring Harb. Perspect. Med. 7 (8) (2017) a024117.
[20] P. Masrori, P.V. Damme, Amyotrophic lateral sclerosis: a clinical review, Eur. J. Neurol. 27 (10) (2020) 1918–1929.
[21] Y. Miao, E. Yin, B.Z. Allison, Y. Zhang, Y. Chen, Y. Dong, … J. Jin, An ERP-based BCI with peripheral stimuli: validation with ALS patients, Cogn. Neurodyn. 14 (1) (2019) 21–33.
[22] H.-P. Müller, M.R. Turner, J. Grosskreutz, S. Abrahams, P. Bede, V. Govind, … J. Kassubek, A large-scale multicentre cerebral diffusion tensor imaging study in amyotrophic lateral sclerosis, J. Neurol.Neurosurg. Psychiatry 87 (6) (2016) 570–579.
[23] Y. Okahara, K. Takano, M. Nagao, K. Kondo, Y. Iwadate, N. Birbaumer, K. Kansaku, Long-term use of a neural prosthesis in progressive paralysis, Sci. Rep. 8 (2018) 16787.
[24] M.L. Ong, P.F. Tan, J.D. Holbrook, Predicting functional decline and survival in amyotrophic lateral sclerosis, PLoS ONE 12 (2017) e0174925.
[25] C. Pancotti, G. Birolo, C. Rollo, T. Sanavia, B.D. Camillo, U. Manera, … P. Fariselli, Deep learning methods to predict amyotrophic lateral sclerosis disease progression, Sci. Rep. 12 (2022) 1–10.
[26] J. Park, J. Kim, T. Song, The global burden of motor neuron disease: an analysis of the 2019 Global Burden of Disease Study, Front. Neurol. 13 (2022).
[27] S.R. Pfohl, R.B. Kim, G.S. Coan, C.S. Mitchell, Unraveling the complexity of amyotrophic lateral sclerosis survival prediction, Front. Neuroinform. 12 (2018) 36.
[28] A. Ragagnin, S. Shadfar, M. Vidal, M. Jamali, J. Atkin, Motor neuron susceptibility in ALS/FTD, Front. Neurosci. 13 (2019) 532.
[29] W. Reniers, M. Schrooten, K.G. Claeys, P. Tilkin, A. D'Hondt, R.D. Van, … D.P. Van, Prognostic value of clinical and electrodiagnostic parameters at time of diagnosis in patients with amyotrophic lateral sclerosis, Amyotrophic Lateral Sclerosis Frontotemp. Degenerat. 18 (5–6) (2017) 341–350.
[30] P. Rong, Y. Yunusova, J. Wang, J.R. Green, Predicting early bulbar decline in amyotrophic lateral sclerosis: a speech subsystem approach, Behav. Neurol. 2015 (2015) 183027.
[31] A. Rosenbohm, R.S. Peter, S. Erhardt, D. Lulé, D. Rothenbacher, A.C. Ludolph, N. Gabriele, Epidemiology of amyotrophic lateral sclerosis in Southern Germany, J. Neurol. 264 (4) (2017) 749–757.
[32] C. Schuster, O. Hardiman, P. Bede, Survival prediction in Amyotrophic lateral sclerosis based on MRI measures and clinical characteristics, BMC Neurol. 17 (73) (2017).
[33] H. Seibold, A. Zeileis, T. Hothorn, Individual treatment effect prediction for amyotrophic lateral sclerosis patients, Stat. Methods Med. Res. 27 (10) (2018) 3104–3125.
[34] G. Sekar, C. Sivakumar, J. Logeshwaran, NMLA: the smart detection of motor neuron disease and analyze the health impacts with neuro machine learning model, Neuroquantology. 20 (8) (2022) 892–899.
[35] R. Sorbello, S. Tramonte, M.E. Giardina, V.L. Bella, R. Spataro, B. Allison, … A. Chella, A human–humanoid interaction through the use of BCI for locked-in ALS patients using neuro-biological feedback fusion, IEEE Trans. Neural Syst. Rehabilit. Eng. 26 (2) (2018) 487–497.
[36] P. Talman, T. Duong, S. Vucic, S. Mathers, S. Venkatesh, R. Henderson, … M. Kiernan, Identification and outcomes of clinical phenotypes in amyotrophic lateral sclerosis/motor neuron disease: australian National Motor Neuron Disease observational cohort, BMJ Open 30 (6(9)) (2016) e012054.
[37] A.A. Taylor, C. Fournier, M. Polak, L. Wang, N. Zach, M. Keymer, … D.L. Ennist, Predicting disease progression in amyotrophic lateral sclerosis, Ann. Clin. Transl. Neurol. 3 (11) (2016) 866–875.
[38] H.K. van der Burgh, Deep learning predictions of survival based on MRI in amyotrophic lateral sclerosis, NeuroImage: Clinical 13 (2017) 361–369.
[39] C. Verbaarschot, D. Tump, A. Lutu, M. Borhanazad, J. Thielen, P.v. Broek, … P. Desain, A visual brain-computer interface as communication aid for patients with amyotrophic lateral sclerosis, Clinical Neurophysiology 132 (2021) (2021) 2404–2415.
[40] J. Wang, S. Xu, Y. Dai, S. Gao, An eye tracking and brain–computer interface-based human–environment interactive system for amyotrophic lateral sclerosis patients, IEEE Sens. J. 23 (20) (2023) 24095–24106.
[41] H.J. Westeneng, T.P. Debray, A.E. Visser, R.P. van Eijk, J.P. Rooney, A. Calvo, … L.H. van den Berg, Prognosis for patients with amyotrophic lateral sclerosis: development and validation of a personalised prediction model, Lancet Neurol. 17 (5) (2018) 423–433.