Abstract

The arrival of genotype-specific therapies in amyotrophic lateral sclerosis (ALS) signals the dawn of precision medicine in motor neuron diseases (MNDs). After decades of academic studies in ALS, we are now witnessing tangible clinical advances. An ever increasing number of well-designed descriptive studies have been published in recent years, characterizing typical disease-burden patterns in vivo and post mortem. Phenotype- and genotype-associated traits and “typical” propagation patterns have been described based on longitudinal clinical and biomarker data. The practical caveat of these studies is that they report “group-level”, stereotyped trajectories representative of ALS as a whole. In the clinical setting, however, “group-level” biomarker signatures have limited practical relevance and what matters is the meaningful interpretation of data from a single individual. The increasing availability of large normative data sets, national registries, extant academic data, consortium repositories, and emerging data platforms now permit the meaningful interpretation of individual biomarker profiles and allow the categorization of single patients into relevant diagnostic, phenotypic, and prognostic categories. A variety of machine learning (ML) strategies have been recently explored in MND to demonstrate the feasibility of interpreting data from a single patient. Despite the considerable clinical prospects of classification models, a number of pragmatic challenges need to be overcome to unleash the full potential of ML in ALS. Cohort size limitations, administrative hurdles, data harmonization challenges, regulatory differences, methodological obstacles, and financial implications and are just some of the barriers to readily implement ML in routine clinical practice. Despite these challenges, machine-learning strategies are likely to be firmly integrated in clinical decision-making and pharmacological trials in the near future.

Keywords: machine-learning; artificial intelligence; amyotrophic lateral sclerosis; primary lateral sclerosis; motor neuron disease; neuroimaging; biomarkers

1. Introduction

Biomarker development is a key goal of amyotrophic lateral sclerosis (ALS) research and a vast array of promising biomarkers have been evaluated including molecular, transcriptomic, and metabolic markers; panels of serum, urine, and cerebrospinal fluid (CSF) “wet” biomarkers; and positron emission tomography (PET) and magnetic resonance imaging (MRI) biomarkers [1]. Neuroimaging in ALS has been remarkably successful in capturing the substrate of phenotype-defining pathological change in vivo along the entire neuro-axis. ALS-associated imaging alterations have been described in the brain, spinal cord, plexi, and muscles with remarkable anatomical consistency between various studies. The archetypal imaging signature of ALS includes bilateral precentral gyrus atrophy, degeneration of descending corticospinal and corticobulbar tracts, degeneration of the corpus callosum and brainstem, changes in the lateral columns and the anterior horns of the spinal cord, and fatty infiltration of denervated muscles [2]. Based on statistical observations in large cohorts, fairly consistent disease-associated, genotype-specific, and phenotype-associated disease-burden patterns have been described [3]. To account for observations captured in various disease stages, a series of robust longitudinal studies have also been published, describing anatomical propagation patterns and generating disease-spread models in vivo [4]. Imaging metrics are often correlated with clinical data and linked to pathological stages. Presymptomatic changes have been captured decades before projected symptom onset in asymptomatic carriers of genetic variants, paving the way for viable presymptomatic interventions [5,6]. While the description of “typical” disease burden patterns, “stereotyped” disease propagation trajectories, and “representative” presymptomatic signatures are important academic milestones and help to generate novel biological hypotheses, they have remarkably limited clinical relevance to the diagnosis, monitoring, and management of individual patients. From a purely clinical perspective, “average” survival and “typical” progression rates have limited importance in the assessment of specific individuals. In clinics, we face individual patients who enquire about their own specific phenotype, their own survival prospects, their expected progression rate, likely prognosis, and likely response to therapy as opposed to overall
Fig. 1. The conceptual evolution of amyotrophic lateral sclerosis (ALS) research studies—the transition from group-level descriptive studies to single-patient data interpretation frameworks.

disease-, genotype-, or phenotype-associated averaged values. Accordingly, what is relevant in the clinic is the accurate and meaningful interpretation of single subject data profiles (Fig. 1). The quest, therefore, is the interpretation of the biomarker profile of a single patient when first met in clinic and the ambition is to accurately categorize that individual into a specific diagnostic group, phenotypic class, likely genotype for targeted testing, likely prognostic category, and likely response to specific therapies. There is a notion among ALS researchers that by the time a patient fulfills clinical criteria for ALS, a considerable disease burden has already been accrued, hindering effective pharmacological interventions. Longitudinal studies suggest that by the time a patient is formally diagnosed, the motor cortex, corticospinal, and corticobulbar tracts are already affected [7,8]; therefore, the expectation that a disease-modifying agent introduced at that point would result in perceptible functional gain may be naïve. The optimal therapeutic window, therefore, is likely to precede the point of meeting formal diagnostic criteria and we probably need to shift our attention to “suspected” patients not meeting diagnostic criteria and presymptomatic cohorts harboring ALS-associated genetic variants. From a biomarker perspective, the departure from large academic studies describing “group-level” observation to “single-patient” data interpretation frameworks is long-overdue.

2. Promising Machine Learning Initiatives across Biomarker Domains

2.1 Examples from Clinical Biomarker Data

Bulbar onset, comorbid cognitive, behavioral deficits, short symptom onset to diagnosis interval, early respiratory involvement, and low body mass index (BMI) have long been established as adverse prognostic indicators, but well-trained machine learning (ML) models promise accurate individual predictions [9–16]. ML has been successfully applied to data collected by wearable sensors generating functional insights that are superior to standard rating scales [17]. Gait features have been interpreted in a multi-class (Parkinson’s disease (PD), Huntington’s disease (HD), and ALS) environment combining Naïve Bayes and logistic regression approaches in an ensemble framework [18]. Concise panels of basic demographic and clinical variables have been repeatedly explored in ML models to predict functional disability [19], clinical subtypes [20], functional decline [21], rate of progression [13], caregiver quality of life [22], caregiver burden [23], and survival [16]. Other clinical features such as voice [24], facial movement [25], and electromyography (EMG) variables [26] have also been successfully integrated in ML models to distinguish subjects with ALS from controls. ML models are increasingly applied to rich epidemiology data sets, and interactions between clinical and environmental factors have also been investigated in a logistic regression model [27].

2.2 Examples from Imaging

The archetypal imaging features of ALS include primary motor cortex, corpus callosum, corticospinal tract, and brainstem degeneration [28–30], but selective basal ganglia, thalamic, hippocampal and cerebellar involvement [31–36] is increasingly accepted as part of the imaging signature of ALS. Despite initial focus on the primary motor cortex, the contribution of cerebellar and subcortical grey matter pathology to key clinical manifestations are increasingly recognized [37–39]. A wide range of imaging-derived
metrics have been explored in single-patient classification models [40,41], including diffusion data [42,43], morphometric data, functional MRI data [44], network dynamics parameters [45], functional near-infrared spectroscopy (fNIRS) variables [46], and combined panels of structural and diffusivity metrics [47–50]. In line with multi-class categorization efforts, MRI data have been increasingly utilized to distinguish specific phenotypes [51,52]. MRI-derived indices have also been evaluated in survival prediction [53,54]. More recently, vision transformer architectures were tested to distinguish subjects with ALS from controls combining spatial and frequency domain information to enhance model performance [55]. It is noteworthy that promising single-patient data interpretation has also been achieved using z-score-based contrasting of MRI metrics alone remains relatively challenging [53,54].

2.4 Examples from Genetics and Transcriptomics

Innovative ML strategies have been developed integrating functional genomics with genome-wide association study (GWAS) summary statistics to aid gene discovery [74]. ML models have been used to predict the pathogenicity of TARDBP and FUS gene variants [75] and also successfully applied to transcriptomic [76] and microRNA profiles [77]. The polygenic underpinnings of cognitive dysfunction in ALS have been recently explored by an innovative ML approach [78]. Advanced clustering methods have been applied to genetic [79], clinical and imaging [80] data sets capturing unique subpopulations with distinctive genetic, clinical, or radiological features.

3. Discussion

3.1 Shortcomings of Recent Studies

As demonstrated by the examples above, several methodologically diverse and promising pilot studies have been published recently. While all of these indicate the potential clinical role of ML in motor neuron disease (MND) and signal tangible future opportunities, a number of practical caveats hinder the implementation of these models in routine clinical practice (Table 1). The vast majority of recent ML initiatives in ALS were either single-center studies or merely trained and validated on national data sets. From a diagnostic perspective, strikingly few multi-class classification models were tested. The vast majority of biomarker and imaging studies in MND focus on ALS, and cohorts of Kennedy’s disease, PLS, post-polio syndrome, are spinal muscular atrophy patients are seldom included [81–83] despite overlapping clinical and imaging features [84]. Similarly, despite promising results, the accuracy of proposed models has not been convincingly tested on early-stage, suspected, or presymptomatic cohorts. Moreover, existing models rely either exclusively on clinical data, biomarker data, or imaging data and strikingly few studies have attempted to integrate inputs from a multitude of biomarker domains.

Similarly, procedures to account for missing data are often overlooked or inadequately addressed. Models developed for “real-life” clinical applications must accommodate for the fact that many patients do not have an entire array of comprehensive data sets encompassing clinical, CSF, serum, urine, and imaging inputs. Classification models to date have mostly categorized patients into diagnostic subgroups, phenotypic classes, survival prospect categories, and pathological stages, but unlike in other conditions, the potential of ML to predict response to therapy or likely genotypes have not been comprehensively explored to date. The acknowledgement and candid discussion of these limitations will likely help to shape future study designs and determine research priorities.

3.2 Future Directions

Model validation schemes in the future should include suspected cases, subjects with short disease duration, and presymptomatic gene carriers to compellingly demonstrate the discriminatory potential of a proposed model between patients with ALS and ALS mimics. While the majority of recent studies have implemented supervised ML models, the potential of robust unsupervised models needs to be
Table 1. Prospects and challenges of implementing effective machine-learning strategies in amyotrophic lateral sclerosis and other motor neuron diseases.

<table>
<thead>
<tr>
<th>Opportunities</th>
<th>Challenges</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early diagnosis of suspected cases</td>
<td>Data harmonization: scanners, immune assays etc.</td>
</tr>
<tr>
<td>Prognostic categorization</td>
<td>Data storage challenges: expense, maintenance, access, screening</td>
</tr>
<tr>
<td>Predicting response to specific therapies</td>
<td>Data regulations and regional regulatory differences: GDPR, European Union (EU), United States (USA), Australia etc.</td>
</tr>
<tr>
<td>Phenotypic classification</td>
<td>Cloud storage: expense, regulations</td>
</tr>
<tr>
<td>Identification of likely genotype for targeted testing</td>
<td>Computational demands: real-time versus post hoc analyses</td>
</tr>
<tr>
<td>Rate of functional decline predictions</td>
<td>Validation platforms and model adjustments</td>
</tr>
<tr>
<td>Classification into clinical and pathological disease stages</td>
<td>Financial implications: funding applications, industry collaborations, annual reports</td>
</tr>
<tr>
<td>Pre-enrolment patient stratification for clinical trials</td>
<td>Maintenance and curation of data repositories</td>
</tr>
<tr>
<td>Ascertainment of slow progressors and limited phenotypes</td>
<td>Inclusion bias in training data sets: cognitive, bulbar, NIV-dependent phenotypes may be underrepresented</td>
</tr>
<tr>
<td>Disability profile prediction</td>
<td>Defining access: only raw data contributors versus open-access platforms</td>
</tr>
<tr>
<td>Extra-motor expansion prediction (cognitive, behavioral, cerebellar features)</td>
<td>Weighted/balanced integration of multimodal data: wet biomarkers, clinical data, imaging etc.</td>
</tr>
<tr>
<td>Informing resource allocation (finances, PT, OT, modifications, adaptive strategies)</td>
<td>Legal framework and acknowledging the limitations of ML predictions and classification</td>
</tr>
<tr>
<td>Informing timing of interventions (PEG, NIV)</td>
<td>Data ownership questions</td>
</tr>
<tr>
<td>Phenocconversion prediction in asymptomatic/presymptomatic mutation carriers</td>
<td>Consenting for participation, right of withdrawing from training data sets</td>
</tr>
<tr>
<td>Discrimination of ALS from mimic syndromes and low-incidence phenotypes e.g., PLS, SBMA</td>
<td>Anonymization procedures and pseudonymization for cross-platform, multi-domain data</td>
</tr>
<tr>
<td>Identification of early ALS in FTD and cohorts with psychiatric conditions</td>
<td>Regulation of industry–academia collaborations</td>
</tr>
<tr>
<td>Clustering to identify subsets of patients with unique clinical, radiological, or biomarker profiles potentially harboring rare genetic variants</td>
<td>Intellectual property (IP) issues, IP ownership, commercialization issues: open-access versus subscription based access</td>
</tr>
<tr>
<td>Establishing the comparative diagnostic/monitoring/discriminatory sensitivity (hierarchy) of multiple markers across several domains; some of which may be cheaper yet superior</td>
<td>Imaging logistics: limited availability of PET, high-field MRI platforms and high scanning fees</td>
</tr>
<tr>
<td>Precision, objective tracking of disease burden over time in clinical trials instead relying on functional rating scales and indirect clinical measures</td>
<td>Wet biomarker logistics: LP, storage, transfer, freezing and cold-chain etc.</td>
</tr>
</tbody>
</table>

Abbreviations: ALS, amyotrophic lateral sclerosis; FTD, frontotemporal dementia; GDPR, General Data Protection Regulation; LP, lumbar puncture; MRI, Magnetic resonance imaging; NIV, non-invasive ventilation; OT, Occupational therapy; PEG, percutaneous endoscopic gastrostomy; PET, Positron emission tomography; PLS, primary lateral sclerosis; PT, physiotherapy; SBMA, spinal bulbar muscular atrophy; ML, machine learning.

explored in forthcoming studies. To avoid model overfitting to local data, it seems imperative to build, test, and validate models on multi-site international data sets. Instead of relying on single-domain data, such as clinical metrics, serum markers, or imaging indices in isolation, future models should attempt to integrate features from a multitude of modalities and biomarker platforms. Instead of defining features a priori and restricting analyses to predefined regions of interest in advance, formal variable importance analyses and ranking is important for streamlining models for the evaluation of the most relevant variables only. Binary classification initiatives have to be superseded by robust multi-class classification models to account for disease heterogeneity and common disease mimics to mirror real-life clinical dilemmas. Data harmonization efforts need to be doubled internationally and data transfer legislation needs to be simplified.

3.3 Cause for Optimism

Funding agencies, charities, and academic centers have long recognized the imperative of multi-site, cross-border collaborations, which are indispensable for effective model development and validation. Clinical trials sometimes make some of their raw data available after the con-
Fig. 2. The practical relevance of machine learning initiatives—potential clinical deliverables.

of successful multi-site data platforms. In parallel with increased international collaborations, a number of technological advances are also aiding data collection and real-time data interpretation such as the drop in the price of cloud solutions, the availability of high-performance computational facilities at academic centers, and wearable devices with wireless connection, etc. Despite current sample size limitations, data harmonization difficulties, and legislative challenges, ML methods will no doubt be firmly integrated in individualized diagnostic pipelines, clinical predictions, and assessment of treatment response in the near future (Fig. 2). Academic ML initiatives are likely to filter down to routine clinical practice and develop into viable clinical applications.

Author Contributions
ELT, JL and PB contributed equally to the conceptualisation and drafting of this manuscript. All authors contributed to editorial changes in the manuscript. All authors read and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate
Not applicable.

Acknowledgment
Not applicable.
**Funding**

This study is sponsored by the Health Research Board (HRB EIA-2017-019 & JPND-CoFund-2-2019-1), Spastic Paraplegia Foundation, (SPF), the Irish Institute of Clinical Neuroscience (IICN), Science Foundation Ireland (SFI SP20/SP/8953), and the EU Joint Programme – Neurodegenerative Disease Research (JPND).

**Conflict of Interest**

The authors declare no conflict of interest.

**References**


D’Hulst L, Van Weehege D, Chio A, Calvo A, Moglia C,


Hatano Y, Ishihara T, Onodera O. Accuracy of a machine learning method based on structural and locational information from AlphaFold2 for predicting the pathogenicity of TARDBP and FUS gene variants in ALS. BMC Bioinformatics. 2023; 24: 206.


