Amyotrophic Lateral Sclerosis (ALS) prediction model derived from plasma and CSF biomarkers

Radhika Khosla¹, Manjari Rain¹, Suresh Sharma², Akshay Anand¹*

¹ Department of Neurology, Neuroscience Research Lab, Post Graduate Institute of Medical Education and Research, Chandigarh, India, ² Department of Statistics, Panjab University, Chandigarh, India

Abstract

Amyotrophic Lateral Sclerosis (ALS) is a degenerative disorder of motor neurons which leads to complete loss of movement in patients. The only FDA approved drug Riluzole provides only symptomatic relief to patients. Early Diagnosis of the disease warrants the importance of diagnostic and prognostic models for predicting disease and disease progression respectively. In the present study we represent the predictive statistical model for ALS using plasma and CSF biomarkers. Forward stepwise (Binary likelihood) Logistic regression model is developed for prediction of ALS. The model has been shown to have excellent validity (94%) with good sensitivity (98%) and specificity (93%). The area under the ROC curve is 99.3%. Along with age and BMI, VEGF (Vascular Endothelial Growth Factor), VEGFR2 (Vascular Endothelial Growth Factor Receptor 2) and TDP43 (TAR DNA Binding Protein 43) in CSF and VEGFR2 and OPTN (Optineurin) in plasma are good predictors of ALS.

Introduction

Amyotrophic Lateral Sclerosis (ALS), a multi-system neurodegenerative disorder, is a rare motor neuron disease. The symptoms involve the degeneration of upper and lower motor neurons along with weak muscular strength, lost ability of movement and speech leading to total or partial paralysis. Talbott et al have reported the global prevalence of disease to be 6/100,000 individuals [1] with an approximate male: female ratio of disease incidence to be 1:3 [2]. Riluzole is the only known Food and Drug Administration (FDA) approved drug for ALS which gives only symptomatic relief to patients [3].

Diagnosis and prognosis of ALS is dependent upon clinical investigations. Various models have been proposed to predict the survival and prognosis of the disease [4–8]. These can also help in analysing the course of disease progression during clinical trials. Diagnosing ALS using clinical investigations can take a long time that leads to certain delay in starting the treatment of patients. Hence, diagnosing ALS at the earlier stages of the disease is immensely important.
Biomarkers are the measures that can provide significant information about the disease prediction or progression. In our previous study, out of a panel of six biomolecules including Vascular Endothelial Growth Factor (VEGF), VEGF receptor 2 (VEGFR2), Angiogenin (ANG), Optineurin (OPTN), Transactive response DNA binding protein 43 (TDP43) and Chemokine Ligand 2 (CCL2), five biomolecules were found to be significantly altered in plasma [9]. In another study, Cerebrospinal Fluid (CSF) from the same cohort (approx. half of the patients) was analysed for the same six molecules [10]. Three of the molecules, involved in the angiogenic and neuroprotective pathway, were found to be significantly altered. In the absence of a single biomarker for the disease diagnosis, analysing a panel of molecules in various biofluids simultaneously can have predictive value for ALS. The previous logistic regression model was proposed based on VEGF and CCL2 mRNA levels, serum levels of CCL2 and consumption of smoking and alcohol data with high sensitivity and specificity [11]. However, the model included fewer numbers of patients and only three biomolecules i.e VEGF, CCL2 and lipid hydroperoxides were studied.

We aimed to develop a predictive statistical model based on a new panel of six biomolecules analysed in Plasma and CSF of patients along with their socio-demographic characteristics of the patient population. The forward stepwise (binary likelihood) logistic regression model proposed in the present study can predict ALS with high sensitivity and specificity.

**Methods**

**Participants**

Total 239 participants (107 ALS and 132 controls) were recruited. All the participants provided informed consents. The study approval was provided by the Institutional ethical committee (IEC approval number PGI/IEC/2014/2249) of the Post Graduate Institute of Medical Education and Research, Chandigarh, India. ALS patients were recruited from the Neurology outpatient Department. Among all participants, biomarkers were estimated in 187 unhaemolysed plasma and 86 CSF samples. Socio-demographic data was also collected for variables such as gender, age, BMI, smoking, alcohol, diet, ALSFRS-R, disease onset and duration of the disease. The criterion for including samples in statistical analysis was that there is no missing value for any of the 21 variables. The logistic regression for developing the model was developed considering 23 ALS patients and 14 controls. The patients were diagnosed clinically and recruited on the basis of revised El Escorial criteria [11–13]. All the patients were found to be sporadic on the basis of family history. According to el Escorial criteria, the patients were categorised as definite/possible/probable ALS.

**Statistical analysis and modelling**

All the statistical tests were done using Statistical Product and service Solutions (SPSS v 23.0 SPSS Inc., Chicago, USA). Descriptive statistics was applied to analyse the distribution of data for various parameters. Binary Logistic regression model was applied for predicting risk of ALS based on the quantitative and qualitative data collected. Total 21 variables were tested including the proteins levels such as VEGF, VEGFR2, ANG, OPTN, TDP43 and CCL2 in plasma and CSF and socio-demographic details such as gender, age, BMI, smoking, alcohol, diet, ALSFRS-R, disease onset and duration of the disease. A forward stepwise (likelihood ratio) method was used for applying the model.
Results

Development of ALS predicting logistic regression model

Forward stepwise (likelihood ratio) binary logistic regression analysis was performed to compute the predicted risk \(P\) of ALS with the help of the following equation

\[
P = \frac{1}{1 + e^{-Y}}
\]

Where, \(Y\) is model score.

Before calculating \(Y\), Hosmer–Lemeshow goodness of fit statistic was applied to test whether the given data fits to the logistic model. Null hypothesis \(H_0\) indicating that the given data fits well to the logistic model was tested and chi square \((\chi^2) = 0.468\), degree of freedom \((df) = 8\) and \(p = 0.994\) suggests that the logistic model is adequately in agreement with the null hypothesis and fits the data.

Omnibus test of model coefficients also confirmed that forward stepwise (likelihood) binary logistic regression is highly appropriate analysis for generating predictive equation. Omnibus test yielded \(\chi^2 = 255.58\), \(df = 7\) and \(p < 0.001\).

The Wald test showed that out of the 21 predictors only 7 predictors were significant and can predict the risk of ALS. Following equation was obtained from the Beta values obtained by Wald test and are presented in \(Table 1\).

Model Score \((Y) = -57.04 + 0.151 \text{Age} - 0.243 \text{BMI} - 2501.477 \text{VEGFR2 plasma level} + 93.109 \text{OPTN plasma level} - 0.244 \text{VEGF CSF level} + 3.184 \text{VEGFR2 CSF level} - 0.244 \text{TDP43 CSF level.}

Adequacy of the logistic regression model was supported by -2 log-likelihood method with \(\chi^2 = 51.73\). Coefficient of determination \((R^2)\) was computed using Cox and Snell’s, and Nagelkerke’s \(R^2\), to check the association of variables in current logistic regression model. \(R^2\) close to 1 suggests strong association of selected independent variables with dependent variables. The present logistic regression model has Cox and Snell’s \(R^2 = 0.684\) and Nagelkerke’s \(R^2 = 0.912\).

Validity of logistic regression model

The Correct classification using logistic regression model of ALS was 94%. Sensitivity and specificity of the logistic regression model was 98% and 93%, respectively. Receiver operating characteristic (ROC) curve with 7 predictive variables revealed that the model for predicting ALS risk is an excellent model, as the area under the curve was 99.3% (Fig 1). As expected the ROC curve has low standard error of 0.003 with 95% confidence interval as 0.986–0.999 (Table 2).

Discussion

ALS is a motor neuron disease caused by degenerative changes in the motor neurons of spinal cord and cortical regions in brain. The degenerated neurons lead to impaired synaptic connections with muscles leading to paralysis in patients. In severe cases this may lead to respiratory failure causing fatality [14]. 10% of the cases have family history of ALS and are known as familial ALS cases. However, 90% of the cases are sporadic and occur because of mutations in varied number of genes. Most commonly associated cases are of C9ORF72 and SOD1 genes [15]. The variability in the pathophysiology of disease makes it a multi system degenerative disease or a multivariate disease. This multi system degeneration obscures the diagnosis, prognosis and treatment strategies even more. However, early prediction of the disease and predictive
prognostic patterns for individualised or cohorts of patients can help in finding effective treatment strategies for ALS and delaying ALS-related adversity and mortality.

Failing to find a single molecule or factor for diagnosis, opting for a panel of markers or some statistical models or equations can help in prediction of ALS. Such models can also help in analysing the prognosis of disease in patients. We have analysed such a panel of markers that are involved in pathways of pathology of disease. VEGF [16–18], VEGFR2 [19, 20] and ANG [21, 22] have been studied in respect to angiogenic pathways as well as in neuroprotective pathways. The dysregulated levels of angiogenic molecules can cause oxidative stress. Oxidative stress has been linked to neuronal degeneration in various studies [23, 24]. The soluble counterpart of another VEGF receptor (sVEGFR1) has been associated with ALS in previous lab studies. These molecules have been shown to be neuroprotective in various studies. Other

Table 1. Significant independent variables revealed by maximum likelihood method for logistic regression equation.

| Variables          | Beta (β) | Standard error | Wald   | Degree of freedom | p-value |
|--------------------|----------|----------------|--------|-------------------|---------|
| Age                | 0.151    | 0.040          | 14.601 | 1                 | <0.001  |
| BMI                | -0.243   | 0.106          | 5.296  | 1                 | 0.021   |
| VEGFR2 plasma level| -2501.477| 668.648         | 13.996 | 1                 | <0.001  |
| OPTN plasma level  | 93.109   | 33.889         | 7.548  | 1                 | 0.006   |
| VEGF CSF level     | -0.244   | 0.071          | 11.777 | 1                 | 0.001   |
| VEGFR2 CSF level   | 3.184    | 0.639          | 24.841 | 1                 | <0.001  |
| TDP43 CSF level    | -0.130   | 0.038          | 11.378 | 1                 | 0.001   |
| Constant           | -57.040  | 13.953         | 16.713 | 1                 | <0.001  |

Abbreviations: BMI Body Mass Index, VEGFR2 Vascular Endothelial Growth Factor Receptor 2, OPTN Optineurin, VEGF Vascular Endothelial Growth Factor, CSF Cerebrospinal Fluid, TDP 43 Transactive Response DNA Binding Protein 43.

https://doi.org/10.1371/journal.pone.0247025.001

Fig 1. Receiver operating characteristic (ROC) for the forward stepwise binary logistic regression model for predicting ALS risk.

https://doi.org/10.1371/journal.pone.0247025.g001
two molecules OPTN [25] and TDP43 [26] are associated with proteinopathy, which is a characteristic of ALS. Both molecules have been found to be accumulated in protein inclusions in the cytoplasm of neurons. Also TDP43 levels have been measured in CSF and serum of ALS patients as it is a major content of motor neuron inclusions. CCL2 is the main molecule of neuroinflammation pathway, which is also a characteristic feature of ALS [18, 27–29]. VEGF and CCL2 have also been shown to contribute significantly to the regression model developed by Gupta et al [11]. In this study, the molecules have been studied in plasma and CSF, both. Since there are theories that CSF is the fluid that may carry the pathogenic markers responsible for degeneration of motor neurons, measurement of proteins in CSF is important.

In combination with molecular markers simple socio-demographic factors such as age and BMI can also contribute significantly as predicting factors in logistic regression model. Also the protein levels in serum [30, 31] and SNPs [29, 32–34] can be seen in various candidate molecules and then the biomarker potential of these molecules can be explored using such regression models. Analysing the panel of markers in plasma and CSF both along with socio-demographic and clinical details can add more value to the model developed and improve the sensitivity and specificity of the model. The model has the potential of prediction of ALS even though other prognostic and survival prediction models have also been developed in the past years.

The model should be tested on larger cohorts to study the validity and predictability. Also, the markers should be analysed in CSF to blood and at cellular level (in the form of gene expression) to add to the validity of models.

### Conclusion

The proposed forward stepwise (binary likelihood) logistic regression model has shown high sensitivity and specificity. Also, the 99.3% area under the curve is indicative of the excellence of the model in predicting the risk of ALS. However, lesser sample size can be shortcoming of the predictive model. Developing such risk predicting models or combined models that can predict the risk of disease as well as survival using bigger cohorts of participants can be helpful in understanding the aetiology of ALS.

### Supporting information

S1 Data.

(XLSX)

### Acknowledgments

We acknowledge the volunteers who participated in the study.
Author Contributions
Conceptualization: Akshay Anand.
Data curation: Radhika Khosla.
Formal analysis: Radhika Khosla, Suresh Sharma.
Funding acquisition: Akshay Anand.
Investigation: Radhika Khosla.
Methodology: Radhika Khosla.
Project administration: Akshay Anand.
Supervision: Akshay Anand.
Validation: Suresh Sharma, Akshay Anand.
Visualization: Akshay Anand.
Writing – original draft: Radhika Khosla, Manjari Rain.
Writing – review & editing: Radhika Khosla, Manjari Rain, Akshay Anand.

References
1. Talbott E, Malek A, Lacomis D. The epidemiology of amyotrophic lateral sclerosis. In: Handbook of clinical neurology. Vol 138. Elsevier; 2016:225–238.
2. Manjaly ZR, Scott KM, Abhinav K, et al. The sex ratio in amyotrophic lateral sclerosis: a population based study. Amyotrophic Lateral Sclerosis. 2010; 11(5):439–442. https://doi.org/10.3109/17482961003610853 PMID: 20225930
3. Borras-Blasco J, Plaza-Macias I, Navarro-Ruiz A, Peris-Martí J, Anton-Canó A. Riluzole as a treatment for amyotrophic lateral sclerosis. Revista de neurología. 1998; 27(160):1021–1027. PMID: 9951030
4. Su XW, Simmons Z, Mitchell RM, Kong L, Stephens HE, Connor JR. Biomarker-based predictive models for prognosis in amyotrophic lateral sclerosis. JAMA neurology. 2013 Dec 1; 70(12):1505–11. https://doi.org/10.1001/jamaneurol.2013.4646 PMID: 24145899
5. Magnus T, Beck M, Giess R, Puls I, Naumann M, Toyka KV. Disease progression in amyotrophic lateral sclerosis: predictors of survival. Muscle & Nerve: Official Journal of the American Association of Electrodiagnostic Medicine. 2002 May; 29(5):709–14. https://doi.org/10.1002/mus.10090 PMID: 11994965
6. Taylor AA, Fournier C, Polak M, Wang L, Zach N, Keymer M, et al. Pooled Resource Open-Access ALS Clinical Trials Consortium. Predicting disease progression in amyotrophic lateral sclerosis. Annals of Clinical and Translational Neurology. 2016 Nov; 3(11):866–75. https://doi.org/10.1002/acn3.348 PMID: 27844032
7. Ong ML, Tan PF, Holbrook JD. Predicting functional decline and survival in amyotrophic lateral sclerosis. PLoS One. 2017 Apr 13; 12(4):e0174925. https://doi.org/10.1371/journal.pone.0174925 PMID: 28406915
8. Westeneng HJ, Debray TP, Visser AE, van Eijk RP, Rooney JP, Calvo A, et al. Prognosis for patients with amyotrophic lateral sclerosis: development and validation of a personalised prediction model. The Lancet Neurology. 2018 May 1; 17(5):423–33. https://doi.org/10.1016/S1474-4422(18)30089-9 PMID: 29598923
9. Modgil S, Khosla R, Tiwari A, Sharma K, Anand A. Association of Plasma biomarkers for Angiogenesis and Proteinopathy in Indian Amyotrophic Lateral Sclerosis Patients. Journal of Neurosciences in Rural Practice. 2020 Aug 20. https://doi.org/10.1055/s-0040-1714314 PMID: 33144793
10. Khosla R, Rain M, Chawathey S, Modgil S, Tyagi R, Thakur K, et al. Identifying putative cerebrospinal fluid biomarkers of amyotrophic lateral sclerosis in a north Indian population. Muscle & Nerve. 2020 Oct; 62(4):328–33.
11. Gupta PK, Prabhakar S, Sharma S, Anand A. A predictive model for amyotrophic lateral sclerosis (ALS) diagnosis. Journal of the neurological sciences. 2012 Jan 15; 312(1–2):68–72. https://doi.org/10.1016/j.jns.2011.08.021 PMID: 21907354
12. Brooks BR, Miller RG, Swash M, Munsat TL; World Federation of Neurology Research Group on Motor Neuron Diseases. El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis.
13. Agosta F, Al-Chalabi A, Filippi M, et al. The El Escorial criteria: strengths and weaknesses. Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration. 2015; 16(1–2):1–7. https://doi.org/10.3109/21678421.2014.964258 PMID: 25482030

14. Corcia P, Pradat PF, Salachas F, et al. Causes of death in a post-mortem series of ALS patients. Amyotrophic Lateral Sclerosis. 2008; 9(1):59–62. https://doi.org/10.1080/17482960701656940 PMID: 17924236

15. Renton AE, Chiò A, Traynor BJ. State of play in amyotrophic lateral sclerosis genetics. Nature neuroscience. 2014; 17(1):17. https://doi.org/10.1038/nn.3584 PMID: 24369373

16. Oosthuyse B, Moons L, Storkebaum E, et al. Deletion of the hypoxia-response element in the vascular endothelial growth factor promoter causes motor neuron degeneration. Nature genetics. 2001; 28(2):131. https://doi.org/10.1038/88842 PMID: 11381259

17. Lambrechts D, Storkebaum E, Morimoto M, et al. VEGF is a modifier of amyotrophic lateral sclerosis in mice and humans and protects motoneurons against ischemic death. Nature genetics. 2003; 34(4):383 https://doi.org/10.1038/ng1211 PMID: 12847526

18. Gupta PK, Prabhakar S, Abburi C, Sharma NK, Anand A. Vascular endothelial growth factor-A and chemokine ligand (CCL2) genes are upregulated in peripheral blood mononuclear cells in Indian amyotrophic lateral sclerosis patients. Journal of neuroinflammation. 2011 Dec 1; 8(1):114. https://doi.org/10.1186/1742-2044-8-114 PMID: 21906274

19. Brockington A, Wharton SB, Fernando M, et al. Expression of vascular endothelial growth factor and its receptors in the central nervous system in amyotrophic lateral sclerosis. Journal of Neuropathology & Experimental Neurology. 2006; 65(1):26–36. https://doi.org/10.1097/01.jnep.0000196134.51217.74 PMID: 16410746

20. Vijayalakshmi K, Ostwal P, Sumitha R, Shruthi S, Varghese AM, Mishra P, et al. Role of VEGF and VEGFR2 receptor in reversal of ALS-CSF induced degeneration of NSC-34 motor neuron cell line. Molecular neurobiology. 2015 Jun 1; 51(3):995–1007. https://doi.org/10.1007/s12035-014-8757-y PMID: 24880751

21. Crivello M, Oriordan SL, Woods I, Cannon S, Halang L, Coughlan KS, et al. Pleiotropic activity of systemically delivered angiogenin in the SOD1G93A mouse model. Neuropsychopharmacology. 2018; 133:503–11. https://doi.org/10.1016/j.neuropharm.2018.02.022 PMID: 29486168

22. Povyshev A, Shmarov M, Logunov D, Naroditsky B, Shulman I, Ogurcov S, et al. Post-spinal cord injury astrocyte-mediated functional recovery in rats after intraspinal injection of the recombinant adenoviral vectors Ad5-VEGF and Ad5-ANG. Journal of neurosurgery Spine. 2017; 27(1):105–15. https://doi.org/10.3171/2016.9.SPINE15959 PMID: 28452633

23. Anand A, Thakur K, Gupta PK. ALS and oxidative stress: the neurovascular scenario. Oxidative medicine and cellular longevity. 2013 Oct; 2013. https://doi.org/10.1155/2013/635831 PMID: 24367722

24. Polliari E, Goldsteins G, Bart G, Koistinaho J, Giniatullin R. The role of oxidative stress in degeneration of the neuromuscular junction in amyotrophic lateral sclerosis. Frontiers in cellular neuroscience. 2014 May 13; 8:131. https://doi.org/10.3389/fncel.2014.00131 PMID: 24860432

25. Akizuki M, Yamashita H, Uemura K, Maruyama H, Kawakami H, Ito H, et al. Optineurin suppression causes neuronal cell death via NF-κB pathway. Journal of neurochemistry. 2013; 126(6):699–704. https://doi.org/10.1111/jncc.12326 PMID: 23721573

26. Winton MJ, Igaz LM, Wong MM, Kwong LK, Trojanowski JQ, Lee VM-Y. Disturbance of nuclear and cytoplasmic Tar DNA binding protein (TDP-43) induces disease-like redistribution, sequestration and aggregate formation. Journal of Biological Chemistry. 2008. https://doi.org/10.1074/jbc.M800342200 PMID: 18305110

27. Nagata T, Nagano I, Shiote M, Narai H, Murakami T, Hayashi T, et al. Elevation of MCP-1 and MCP-1/VEGF ratio in cerebrospinal fluid of amyotrophic lateral sclerosis patients. Neurological research. 2007; 29(8):772–6. https://doi.org/10.1179/016164107X229795 PMID: 17672928

28. Henkel JS, Beers DR, Siklos L, Appel SH. The chemokine MCP-1 and the dendritic and myeloid cells it attracts are increased in the mSOD1 mouse model of ALS. Molecular and Cellular Neurosciences. 2006; 31(3):427–37. https://doi.org/10.1016/j.mcn.2005.10.016 PMID: 16357193

29. Anand A, Sharma NK, Gupta A, Prabhakar S, Sharma SK, Singh R, et al. Single nucleotide polymorphisms in MCP-1 and its receptor are associated with the risk of age related macular degeneration. PloS one. 2012 Nov 21; 7(11):e49905. https://doi.org/10.1371/journal.pone.0049905 PMID: 23185481

30. Sharma NK, Prabhakar S, Gupta A, Singh R, Gupta PK, Gupta PK, et al. New biomarker for neovascular age-related macular degeneration: eotaxin-2. DNA and cell biology. 2012 Nov 1; 31(11):1618–27. https://doi.org/10.1089/dna.2012.1786 PMID: 23025269
31. Anand A, Gupta PK, Sharma NK, Prabhakar S. Soluble VEGFR1 (sVEGFR1) as a novel marker of amyotrophic lateral sclerosis (ALS) in the North Indian ALS patients. European Journal of Neurology. 2012 May; 19(5):788–92. https://doi.org/10.1111/j.1468-1331.2011.03548.x PMID: 21978169

32. Sharma NK, Gupta A, Prabhakar S, Singh R, Bhatt AK, Anand A. CC chemokine receptor-3 as new target for age-related macular degeneration. Gene. 2013 Jul 1; 523(1):106–11. https://doi.org/10.1016/j.gene.2013.03.052 PMID: 23566847

33. Kamal Sharma N, Gupta A, Prabhakar S, Singh R, Sharma S, Anand A. Single nucleotide polymorphism and serum levels of VEGFR2 are associated with age related macular degeneration. Current neurovascular research. 2012 Nov 1; 9(4):256–65. https://doi.org/10.2174/156720212803530681 PMID: 23030506

34. Sharma NK, Gupta A, Prabhakar S, Singh R, Sharma SK, Chen W, et al. Association between CFH Y402H polymorphism and age related macular degeneration in North Indian cohort. PloS one. 2013 Jul 29; 8(7):e70193. https://doi.org/10.1371/journal.pone.0070193 PMID: 23922956