Clin. Lab. 2026;72:1-3

ORIGINAL ARTICLE

Identification of Glycolysis-Related Diagnostic Biomarkers for Amyotrophic Lateral Sclerosis Using Machine Learning

Ping Ye [1, 2], Binhan Guo [1, 2]

[1] *Department of Laboratory Medicine, West China Second University Hospital, Sichuan University, Chengdu, China*
[2] *Key Laboratory of Birth Defects and Related Diseases of Women and Children (Sichuan University), Ministry of Education, Chengdu, China*

ABSTRACT

Background: **Glycometabolism has been implicated in the pathogenesis of amyotrophic lateral sclerosis (ALS), yet the precise molecular mechanisms underlying this association remain poorly understood. The identification of reliable biomarkers for ALS diagnosis represents a critical unmet need in clinical practice, as early detection and intervention could significantly improve patient outcomes.**
Methods: **We employed a comprehensive analytical approach combining two-sample Mendelian randomization analysis to investigate the causal relationship between blood glucose levels and ALS. Additionally, we integrated differential expression analysis, multiple machine learning algorithms, and correlation analyses to identify potential diagnostic biomarkers for ALS. The machine learning framework utilized gradient boosting tree methodology to construct predictive models, with performance evaluation conducted through cross-validation procedures.**
Results: **Mendelian randomization analysis demonstrated a significant negative causal relationship between blood glucose levels and ALS risk. Through bioinformatic analysis and machine learning approaches, we successfully identified candidate genes and constructed a high-performance predictive model using gradient boosting tree methodology, achieving an average area under the curve (AUC) of 0.8782 in cross-validation. Validation studies utilizing both bulk and single-cell RNA sequencing datasets revealed that *COL5A1* and *VCAN* genes play significant roles in ALS pathogenesis, likely through their involvement in glycolytic pathways.**
Conclusions: **Our findings provide novel insights into the molecular mechanisms linking glycometabolism and ALS, while identifying potential diagnostic biomarkers for the disease. The identified genes, *COL5A1* and *VCAN*, represent promising targets for further investigation in ALS pathogenesis. However, the clinical translation of these findings requires validation through additional datasets and prospective clinical trials to establish their diagnostic utility and therapeutic potential.**
(Clin. Lab. 2026;72:1-3. DOI: 10.7754/Clin.Lab.2025.250766)

Correspondence:
Binhan Guo
Department of Laboratory Medicine
West China Second University Hospital
Sichuan University
Chengdu, 610041
China
Email: 2183983697@qq.com
ORCID: 0000-0002-1448-8022

Manuscript accepted August 14, 2025

Supplementary Data

Table S1.

Parameter	Fold1	Fold2	Fold3	Average
Accuracy	0.6207	0.6484	0.6599	0.6430
Bal_accuracy	0.5098	0.5869	0.6087	0.5685
Precision	0.6284	0.6754	0.6794	0.6611
Recall	0.9585	0.8380	0.8396	0.8787
F1	0.7591	0.7479	0.7511	0.7527
ROC_AUC	0.6766	0.6855	0.7145	0.6922
PR_AUC	0.7843	0.7645	0.7964	0.7817

Table S2.

Parameter	Fold1	Fold2	Fold3	Average
Accuracy	0.7874	0.7262	0.9078	0.8071
Bal_accuracy	0.7327	0.6795	0.8896	0.7672
Precision	0.7638	0.7373	0.8879	0.7963
Recall	0.9539	0.8704	0.9717	0.9320
F1	0.8484	0.7983	0.9279	0.8582
ROC_AUC	0.8685	0.7898	0.9762	0.8782
PR_AUC	0.9058	0.8555	0.9853	0.9155

Table S3.

Parameter	Fold1	Fold2	Fold3	Average
Accuracy	0.7184	0.7003	0.7723	0.7303
Bal_accuracy	0.6638	0.6481	0.7478	0.6865
Precision	0.7245	0.7154	0.7879	0.7426
Recall	0.8848	0.8611	0.8585	0.8681
F1	0.7967	0.7815	0.8217	0.8000
ROC_AUC	0.8034	0.8068	0.8380	0.8161
PR_AUC	0.8622	0.8849	0.8985	0.8819

Table S4.

Parameter	Fold1	Fold2	Fold3	Average
Accuracy	0.7989	0.7118	0.7695	0.7600
Bal_accuracy	0.7449	0.6318	0.7198	0.6989
Precision	0.7712	0.6946	0.7463	0.7374
Recall	0.9631	0.9583	0.9434	0.9550
F1	0.8566	0.8054	0.8333	0.8318
ROC_AUC	0.8760	0.8408	0.8456	0.8541
PR_AUC	0.9075	0.8968	0.8923	0.8989

Table 5.

Cell type	Genes
Radial glia	*PAX6*, *FABP7*
Early neuron	*DCX*, *TUBB3*
Neuron	*MAP2*
(Upper layer) neuron	*CUX1*, *SATB2*, *POU3F2*
Glutamergic excitatory neuron	*SLC17A5*, *SLC17A6*
Deep layer neuron	*BCL11B*, *FOXP2*
GABAergic inhibitory neuron	*GAD1*, *GAD2*
Glia/astrocyte	*GFAP*, *SLC1A3*, *APOE*
Oligodendrocyte precursor cell	*PLP1*, *MBP*
Endothelium	*PECAM1*, *PTPRB*
Choroid plexus	*HTR2C*, *CLIC6*

Table S6.

Parameter	Fold1	Fold2	Fold3	Average
Accuracy	0.6236	0.6225	0.6110	0.6190
Bal_accuracy	0.5000	0.5000	0.5000	0.5000
Precision	0.6236	0.6225	0.6110	0.6190
Recall	1.0000	1.0000	1.0000	1.0000
F1	0.7681	0.7673	0.7585	0.7647
ROC_AUC	0.7103	0.6740	0.6039	0.6627
PR_AUC	0.7866	0.7714	0.7254	0.7611