



OPEN ACCESS

EDITED BY

Wen-Xiong Chen,
Guangzhou Medical University, China

REVIEWED BY

Francesca Felicia Operto,
University of Salerno, Italy
Santasree Banerjee,
Jilin University, China
Zhanqi Hu,
Shenzhen Children's Hospital, China

*CORRESPONDENCE

Liyang Liang
✉ liangliy@mail.sysu.edu.cn
Wenting Tang
✉ tangwt@sysucc.org.cn

RECEIVED 21 March 2025

ACCEPTED 23 September 2025

PUBLISHED 07 October 2025

CITATION

Wu R, Luo X, He Z, Meng Z, Tang W and
Liang L (2025) Predicting the diagnostic
efficacy of trio-based whole exome
sequencing in children with low-function
autism spectrum disorders: a multicenter
study.
Front. Neurol. 16:1597588.
doi: 10.3389/fneur.2025.1597588

COPYRIGHT

© 2025 Wu, Luo, He, Meng, Tang and Liang.
This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other forums is
permitted, provided the original author(s) and
the copyright owner(s) are credited and that
the original publication in this journal is cited,
in accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Predicting the diagnostic efficacy of trio-based whole exome sequencing in children with low-function autism spectrum disorders: a multicenter study

Ruohao Wu^{1,2}, Xiangyang Luo^{1,2,3}, Zhanwen He^{1,2}, Zhe Meng^{1,2},
Wenting Tang^{4*} and Liyang Liang^{1,2*}

¹Department of Children's Neuro-endocrinology, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, Guangdong, China, ²Children's Medical Center, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, Guangdong, China, ³Weierkang Children's Rehabilitation Center, Guangzhou, Guangdong, China, ⁴Department of Research and Molecular Diagnostics, Sun Yat-sen University Cancer Center, Sun Yat-sen University, Guangzhou, Guangdong, China

Background: Although significant progress has been made in trio-based whole-exome sequencing (trio-WES) that enables the detection of exon-level variants, the diagnostic effectiveness of empirical and unselected use of trio-WES in children with low-function autism spectrum disorders (LF-ASDs) remains unsatisfactory. Thus, the identification of an appropriate approach for predicting the diagnostic efficacy of trio-WES at the pre-diagnosis stage is essential for implementing individualized diagnosis for children with LF-ASDs.

Methods: A total of 168 LF-ASDs patients who underwent trio-WES at Sun Yat-sen Memorial Hospital from September 2016 to December 2022 were enrolled as the training set. Additionally, 58 LF-ASDs patients who received trio-WES at Weierkang Children's Rehabilitation Center between January 2023 and December 2023 were recruited as an independent external validation set. Univariate and multivariate binary logistic analyses were performed on the training set to select phenotypic variables to establish a nomogram. The discriminative performance of the model was evaluated using receiver operating characteristic (ROC) curves and calibration curves. Furthermore, the nomogram was validated in external validation sets.

Results: Univariate and multivariate analyses identified independent trio-WES diagnosis-related predictive indicators, including severity of global developmental delay/intellectual disability, complexity of neurodevelopmental/neurological comorbid conditions, head circumference abnormalities, and brain malformations, in the training cohort and used to develop a nomogram. The nomogram showed excellent discrimination performance, with an area under curve (AUC) of the ROC in the training cohort of 0.868 (95% CI: 0.811–0.925), resulting in sensitivity, specificity, accuracy, precision, and F1 score values of 85.56, 82.05, 83.93, 84.62%, and 0.85, respectively. The model also exhibited strong prediction ability in the external validation set (AUC: 0.941, 95% CI: 0.880–0.998; sensitivity: 85.29%; specificity: 91.67%; accuracy: 87.93%; precision: 93.55%; and F1 score: 0.89). Moreover, the calibration curves demonstrated good agreement between the nomogram predictions and actual observations in both training and validation sets.

Conclusion: We developed an user-friendly and highly accurate model for predicting the diagnostic probability of trio-WES in LF-ASDs children, which could help implement an individualized diagnostic strategy for affected children and their families at the pre-diagnosis stage.

KEYWORDS

low-function autism spectrum disorders, neurodevelopmental/neurological comorbidities, trio-based whole-exome sequencing, diagnostic rate, phenotype-driven, nomogram, genetic abnormalities

Introduction

Autism spectrum disorders (ASDs) represent a genetically and clinically heterogeneous group of neurodevelopmental disorders characterized by dysfunctions in social communication/interaction and repetitive, stereotypic patterns of movements/behaviors typically manifesting within the first 2–3 years of life (1). With advancements in understanding of ASDs, their prevalence has increased significantly, now accounting for approximately 1–2% of children worldwide, according to the Network Organization of Autism and Developmental Disabilities Monitoring (2). As of 2016, the prevalence of ASDs in children and adolescents (ages 3–17) in the United States was 2.76%, rising to 3.49% by 2020 (3, 4). Recently, the World Health Organization reported that 1 in 100 children worldwide present with ASDs (5). The rapidly increasing prevalence of ASDs places enormous pressure on public health systems, social services, and economic burdens on families worldwide.

Low-function ASDs (LF-ASDs) represent a severe manifestation within the ASDs continuum, affecting nearly 42% of diagnosed with ASDs children (2). LF-ASDs can be defined as ASDs accompanied by varying severities of global developmental delay or intellectual disability (GDD/ID), marked by pronounced and easily observable deficits from an early age (often <18 months) (6). Previous studies indicate that identifiable neurodevelopmental/neurological comorbid conditions (NCCs), such as attention deficit hyperactivity disorder (ADHD) and epilepsy (EP) frequently co-occur in children with LF-ASDs (1). Moreover, there is a stronger correlation with LF-ASDs and significant structural and functional brain alterations, including abnormalities in brain volume growth trajectories and pronounced cortical connectivity disturbances, which can complicate diagnosis and treatment; consequently, LF-ASDs are often referred to as syndromic ASDs. Therapeutic interventions or disease management for ASDs and LF-ASDs are individualized and multidisciplinary, typically encompassing speech and language therapy, occupational therapy, cognitive behavioral therapy, and pharmacotherapy (e.g., risperidone and aripiprazole).

The etiology of ASDs is complex and multifaceted, stemming from a combination of genetic predispositions and environmental influences. The etiology of LF-ASDs encompasses the foundational genetic and environmental factors associated with ASDs but is statistically associated with a higher burden of pathogenic genetic variants; thus, genetic disturbances are still considered to play essential roles in the development of LF-ASDs (1). For instance, mutations or chromosomal abnormalities with larger effect sizes, such as fragile X syndrome and single-gene disorders like Rett syndrome (*MECP2* abnormalities), are strongly associated with the etiology of LF-ASDs (ASDs with GDD/ID). Recent

advancements in identifying the genetic components of LF-ASDs have accelerated, particularly due to the increased adoption and innovation of trio-based (parental-offspring model) whole-exome sequencing (trio-WES), the most common next-generation sequencing technology for detecting exon-level variants, including single-nucleotide variants (SNVs) and copy-number variants (CNVs) in clinical applications, making it possible to identify genetic components more frequently in many idiopathic LF-ASDs cases (7). Nonetheless, around half of children with LF-ASDs remain undiagnosed after receiving comprehensive trio-WES analyses, attributed to variants located outside exons (e.g., intronic, promoter, or enhancer-level variants); thus, the diagnostic yield of trio-WES for LF-ASDs remains unsatisfactory (1). Given the persistent challenges posed by the low diagnostic efficacy of trio-WES in LF-ASDs, it is imperative for pediatricians to develop practical tools for the early identification of children with LF-ASDs most likely be diagnosed by trio-WES, thereby facilitating timely evaluations of medical conditions. Additionally, due to the features of exon-level sequencing of trio-WES, it becomes essential for affected children and their families to use straightforward approaches to support their decision-making regarding the employment of trio-WES testing at the pre-diagnosis stage, ultimately aiding individualized family planning and reducing unnecessary financial and temporal costs.

Nomograms are powerful predictive tools, are widely used in forecasting the outcomes across various diseases due to their visualization, ease of use, objectivity, and accuracy (8). Nomograms have been widely used for predicting risks or outcomes of many pediatric neurodevelopmental disorders, including ADHD (9), infant neurodevelopmental delays (10), and teenager oppositional defiant disorder (11). However, to the best of our knowledge, no studies to date have reported the application of a nomogram for predicting the diagnostic efficacy of trio-WES in LF-ASDs children. Therefore, this multicenter study with independent external validation based on the phenotype-driven concept aims to generate the first user-friendly nomogram model for predicting the individualized diagnostic probability of trio-WES in children with LF-ASDs. Grounded in the phenotype-driven concept, which is essential in clinical genetics; this approach is used for Mendelian monogenic disorders to identify critical phenotypic characteristics that allow the identification of probands with a high probability of harboring relevant pathogenic genetic variants (12). Leveraging this concept, we used readily obtainable and objective phenotypic variables related to LF-ASDs and their associated complex NCCs, thereby establishing a predictive nomogram to assess the individualized diagnostic probability of employing trio-WES at the pre-diagnosis stage.

Materials and methods

Patients and subject selection criteria

As shown in [Figures 1A](#), a total of 560 individuals diagnosed with LF-ASDs were admitted to the tertiary Children's Medical Center of Sun Yat-sen Memorial Hospital (SYSMH) from September 2016 to December 2022. Following a comprehensive assessment involving clinical information, informed consent, and routine genetic screening (G-band karyotyping and fragile-X analysis) to exclude ineligible cases—such as children with unclear or incomplete clinical data (excluding 307 cases), those whose parents or guardians declined genetic testing or opted not to permit the use of their genetic results for publication (excluding 67 cases), and children presenting with apparent chromosomal disorders (e.g., Down syndrome or fragile-X syndrome), which rendered trio-WES inappropriate as a diagnostic strategy (excluding 18 cases), a total of 168 children with idiopathic LF-ASDs who had received trio-WES testing from SYSMH were ultimately in this retrospective study as training subjects. Moreover, 79 children diagnosed LF-ASDs were admitted to Weierkang Children's Rehabilitation Center (WCRC), a specialized pediatric neurorehabilitation center focusing on integrated diagnosis and treatment of neurodevelopmental disorders, from January 2023 to December 2023. After performing a series of screenings similar to those applied in the SYSMH group, 58 patients with idiopathic LF-ASDs ultimately qualified for enrollment, having undergone trio-WES testing between January 1, 2023, and December 31, 2023, thus serving as external validation subjects.

The present study defined LF-ASDs in accordance with the following: (I) A clinical diagnosis of ASDs made by professional pediatric psychiatrists based on the Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5) criteria for ASDs ([13](#)), supplemented by several main ASDs-related clinical assessment scales, including the Modified Checklist for Autism in Toddlers, the Clancy Autism Behavior Scale, the Autism Behavior Checklist, and the Childhood Autism Rating Scale ([14–16](#)). (II) Diverse severities of GDD/ID, where by clinical diagnostic criteria for GDD/ID were based on the DSM-5 criteria ([13](#)). GDD/ID severity was assessed mainly according to the Gesell Developmental Diagnosis Scale (GDDS) for infants under 3 year-old ([17](#)), and Wechsler Preschool and Primary Scale of Intelligence, IV Edition (WPPSI-IV) for children with 4 ~ 6 year-old. For subjects over 6 year-old, we used Wechsler Intelligence Scale for Children, IV Edition (WISC-IV) to assess their GDD/ID severity ([18](#)). (III) Exclusion of subjects with identified non-genetic causes such as hypoxic-ischemic encephalopathy, bilirubin encephalopathy and intrauterine infections, and positive findings from routine genetic screens (fragile-X analysis and G-band karyotyping) that indicate chromosomal disorders, such as fragile X syndrome or Down syndrome, deemed inappropriate for trio-WES in those conditions. (IV) Children with or without common NCCs, mainly ADHD and/or epilepsy, with diagnoses established by specialized child psychiatrists and neurologists following the criteria of the DSM-5 for ADHD ([13](#)) and the International League Against Epilepsy (ILAE) criteria for epilepsy ([19](#)). Due to the diagnostic challenges and lack of objective and Chinese version of assessment tools for other NCCs such as sleep disorder and anxiety disorder in young children, these conditions were excluded from current study.

Ethical compliance

The design of this multicenter study received approval from the Ethical Committee of Sun Yat-sen Memorial Hospital, Sun Yat-sen University (initiative affiliation approval number: SYSKY-2025-244-01). Written informed consent for genetic investigation and publication of genetic results was obtained from the parents or guardians of all 226 enrolled subjects.

Strategy for variant capture of trio-WES

The principles of the variant capture process and quality control systems for trio-WES have been described in previous studies ([20–22](#)), with details in the current study briefly outlined as follows: Genomic DNA was extracted from the whole blood of the proband and their parents using a commercial genomic extraction kit (Qiagen, Shanghai, China). The Illumina TruSeq Exome Kit (Illumina, San Diego, CA, United States) was used for DNA library construction and the generation of approximately 10GB of exome sequencing data per individual. GeneRanger (Xunyun Biotech, Shanghai, China) was employed for the exome sequencing data analysis. Subsequently, the Burrows-Wheeler aligner, Picard tool, and genome analysis tools were employed for read alignment, indel region realignment, base quality recalibration, variant capture, and calling/transformation based on the Genome Aggregation Database (gnomAD). The variant quality control system was set to a coverage depth of greater than 10, along with a minor allele frequency of less than 0.05%.

Pathogenicity criteria for trio-WES-identified exon-level variants and enrolled case grouping

The pathogenicity of trio-WES-identified SNVs was rated according to the 2015 American College of Medical Genetics (ACMG) guidelines for SNV interpretation ([23](#)), categorizing detected SNVs into “pathogenic/likely pathogenic” and “benign/uncertain significance” SNVs. As per prior research ([24](#)), gnomAD and in-house SNV population frequency databases were used to assess SNV allele frequency. *In silico* pathogenic predictions for identified missense, frameshift, nonsense, and deletion variants were conducted using online versions of Mutation Taster,¹ Protein Variation Effect Analyzer,² Polymorphism Phenotyping version 2,³ and Sorting Intolerant From Tolerant.⁴ *In silico* prediction for the splice variant was conducted using online version of Combined Annotation Dependent Depletion.⁵ Additionally, the Human Genomic Mutation Database and PubMed were consulted to determine whether identified variants had been previously documented, while Online Mendelian Inheritance in Man⁶

1 <https://mutationtaster.org/>

2 PROVEAN; <http://provean.jcvi.org/index.php>

3 Polyphen-2; <http://genetics.bwh.harvard.edu/pph2/>

4 SIFT; <https://sift.bii.a-star.edu.sg/>

5 CADD; <https://cadd.gs.washington.edu/>

6 OMIM; <https://omim.org/>

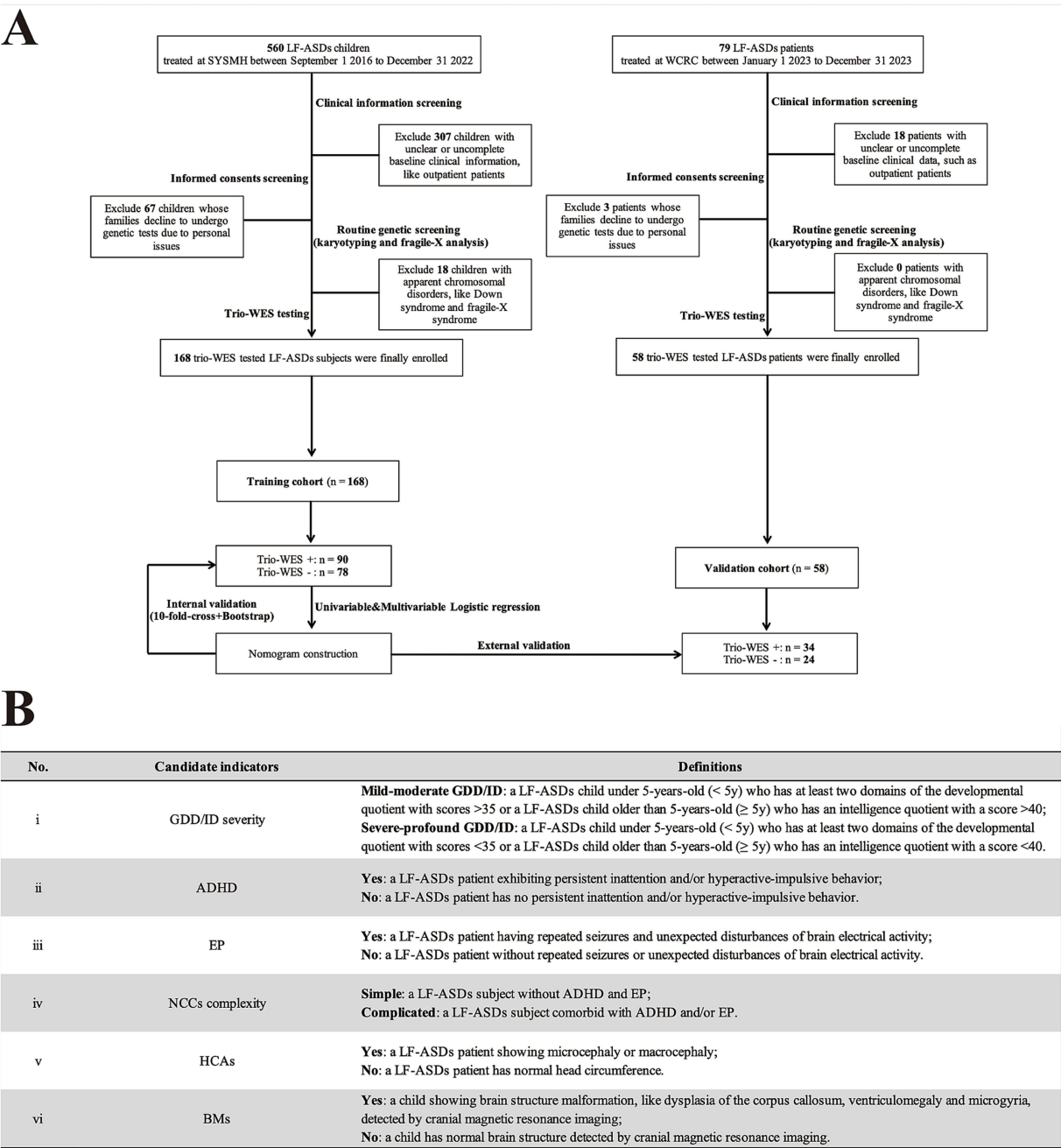


FIGURE 1 Flowchart and candidate variables description of this multicenter study. **(A)** Flowchart of this research. **(B)** A summary of candidate phenotypic indicators and their corresponding definitions in current research. LF-ASDs, low-function autism spectrum disorders; SYSMH, Sun Yat-Sen Memorial Hospital; WCRC, Weierkang Children's Rehabilitation Center; trio-WES, trio-based whole-exome sequencing. NCCs, neurodevelopmental/neurological comorbid conditions; HCAs, head circumference abnormalities; BMs, brain malformations. In **A** "+/-" indicates enrolled LF-SADs patients who underwent trio-WES and received a positive genetic diagnosis (having "likely pathogenic" or "pathogenic" variants) or a negative genetic diagnosis (having "benign" or "uncertain significance" variants), respectively, according to the guidelines of the American College of Medical Genetics. In **B**, Gesell Developmental Diagnosis Scale with developmental quotient score (cut-off value: 35) was used to assess GDD severity (mild-moderate or severe-profound GDD) for patients under 5-year-old. While, Wechsler Intelligence Scale with intelligence quotient score (cut-off value: 40) was used to evaluate ID severity (mild-moderate or severe-profound ID) for patients older than 5-year-old.

database was employed to obtain genotype–phenotype profiles linked to identified SNVs.

The pathogenicity of trio-WES-detected CNVs was assessed based on the 2019 ACMG guidelines for postnatal CNV interpretation (25), employing previously documented methods (26). Identified CNVs were manually interpreted and categorized into “pathogenic/likely pathogenic” or “benign/uncertain significance” by two or more experienced clinical geneticists adhering to ACMG guidelines.

Based on the pathogenicity assessments of these trio-WES-identified SNVs and CNVs mentioned above, enrolled subjects were categorized into LF-ASDs children with positive genetic diagnoses (+, harboring pathogenic/likely pathogenic SNVs or CNVs from their trio-WES testing reports) and LF-ASDs children with negative genetic diagnoses (–, harboring benign/uncertain significance SNVs or CNVs from their trio-WES testing reports).

Candidate variable collection and interpretation of collected indicators

Demographic and phenotypic factors of all enrolled subjects were collected from hospital medical records. These included: (I) demographic characteristics including sex, admission date, and age at which trio-WES was performed; and (II) candidate phenotypic factors: GDD/ID severity, ADHD, epilepsy, complexity of NCCs, head circumference abnormalities (HCAs), and brain malformations (BMs). A summary table detailing these phenotypic variables and their corresponding definitions, is provided in Figure 1B.

Model development and internal/external validation of model performance

For model development, independent phenotypic indicators were first screened through univariate and multivariate binary logistic regression alongside collinearity diagnostic analyses in the training set. Specifically, during the univariate and multivariate logistic analyses of the training set, indicators with clinical significance and significantly significant differences ($p < 0.05$) were identified between children with positive and negative genetic diagnoses via trio-WES. Then, collinearity diagnostic analyses were performed to determine the presence of significant collinearity among the screened indicators. Tolerance and variance inflation factor (VIF) metrics were used to evaluate the severity of collinearity. A tolerance value exceeding 0.5 and a VIF below 5 for each screened variable indicated no significant collinearity, thereby permitting the selection of these variables as independent for establishing the logistic regression model (27). Finally, we generated a nomogram using R packages to visualize the constructed logistic regression model.

For internal validation of model performance, the receiver operating characteristic (ROC) curve and the area under curve (AUC) of the ROC were initially used to evaluate the discriminative performance of the model in the training cohort. Subsequently, calibration curves coupled with the Hosmer-Lemeshow test were applied to assess the goodness-of-fit between predicted and observed data. A p value from the Hosmer-Lemeshow test < 0.05 indicated that the dotted line (representing model-predicted data) significantly differed from the solid line (representing actual observed data) in

calibration curve, demonstrating poor model fit; conversely, a p value > 0.05 implied good model fit. Additionally, the clinical applicability of the nomogram was evaluated through decision curve analysis (DCA) and clinical impact curve (CIC). Furthermore, we used two methods to assess consistency and mitigate overfitting bias: the 10-fold cross-validation and bootstrap resampling (with 1,000 bootstrap resamples). The 10-fold cross-validation approach is a common and robust resampling technique used to assess the consistence performance and internal stability of predictive model, and involves partitioning the dataset into 10 mutually exclusive and approximately equal-sized folds. During iterative training, 9 folds are used as the training set while the remaining fold serves as the validation set. This process is repeated across all folds. The established performance metric, concordance index (C-index), is calculated to examine the consistency (generalization) and stability of predictive model. Conversely, the bootstrap sampling method, a classical internal validation method, draws from the original training dataset with replacement and undergoes 1,000 repetitions. C-index values greater than 0.7 from both methods indicated the nomogram had good reliability (28).

For external validation of model performance, the optimal cutoff value was first set based on the maximal Youden index value corresponding to the optimal values of sensitivity and specificity of the model in the training set. Cases in the external validation set were then classified into “nomogram-predicted positive diagnostic cases” and “nomogram-predicted negative diagnostic cases” based on these optimal cutoff values. The AUC value, calibration curves with the Hosmer-Lemeshow test, and DCA/CIC, were subsequently used to validate the discriminative performance, consistency, and clinical benefits of the nomogram in the external validation set. Finally, we calculated the model sensitivity, specificity, accuracy, precision, and F1 scores for the training and external validation sets, and the results were visualized using Sankey plots.

Statistical analysis

Microsoft Excel software was used for data entry, while all statistical analyses were conducted using R.⁷ As referenced in previous studies using the R (8, 27, 29, 30), the following R packages were used for statistical analysis and data visualization: ggplot2,⁸ “foreign,” “rms,” “rmda,” “caret,” “tidyverse,” and “ggDCA.” p value < 0.05 were considered statistically significant.

Results

Clinical details of the enrolled subjects

In total, 168 and 58 children with unexplained LF-ASDs were enrolled in the training and external validation cohorts, respectively. Comparisons of the baseline demographics and phenotypic features between the two cohorts are shown in Table 1. Detailed data regarding the genotypes and phenotypes of subjects in the training and

⁷ version 4.4.2, <http://www.R-project.org/>

TABLE 1 Comparison of baseline demographics and phenotypic features between training and validation cohorts of trio-WES tested LF-ASDs children.

Demographics or characteristics	Indicators	Training cohort	Validation cohort	t/χ^2 value	p -value
Demographic parameters					
Case number (n)		168	58		
Sex [n (%)]	Female	43 (25.6%)	16 (27.6%)	0.089	0.766
	Male	125 (74.4%)	42 (72.4%)		
Admission Periods (MM/YY ~ MM/YY)		Sep/2016 ~ Dec/2022	Jan/2023 ~ Dec/2023		
Age of receiving trio-WES [Mean \pm SD/(yrs)]		4.7 \pm 3.4	5.3 \pm 2.6	1.324	0.188
Patient sources		SYSMH	WCRC		
Phenotypic features					
GDD/ID severity [n (%)]	Mild-moderate	92 (54.8%)	28 (48.3%)	0.728	0.393
	Severe-profound	76 (45.2%)	30 (51.7%)		
ADHD [n (%)]	Yes	34 (20.2%)	18 (31.0%)	2.837	0.092
	No	134 (79.8%)	40 (69.0%)		
EP [n (%)]	Yes	27 (16.1%)	11 (19.0%)	0.258	0.611
	No	141 (83.9%)	47 (81.0%)		
NCCs complexity [n (%)]	Simple	124 (73.8%)	30 (51.7%)	9.687	0.002
	Complicated	44 (26.2%)	28 (48.3%)	0.065	0.798
HCAs [n (%)]	Yes	52 (31.0%)	19 (32.8%)		
	No	116 (69.0%)	39 (67.2%)		
BMs [n (%)]	Yes	70 (41.7%)	15 (25.9%)	4.590	0.032
	No	98 (58.3%)	43 (74.1%)		
Outcomes					
trio-WES-based diagnostic status [n (%)]	Positive	90 (53.6%)	34 (58.6%)	0.444	0.505
	Negative	78 (46.4%)	24 (41.4%)		

LF-ASDs, low-function autism spectrum disorders; NCCs, neurodevelopmental/neurological comorbid conditions; trio-WES, trio-based whole exome sequencing SYSMH, Sun Yat-Sen Memorial Hospital; WCRC, Weierkang Children's Rehabilitation Center; GDD/ID, global developmental delay/intellectual disability; ADHD, attention deficit hyperactivity disorder; EP, epilepsy; HCAs, head circumference abnormalities; BMs, brain malformations. In NCCs complexity row, "simple" means patients only existing LF-ASDs without ADHD and (or) EP; "complicated" means LF-ASDs patients coexisting one or two type of other NCCs (ADHD and EP).

validation cohorts are summarized in [Supplementary Files 1–4](#), respectively. Among the 168 enrolled subjects in the training cohort, 90 (53.6%) had a genetic diagnosis via trio-WES, whereas 78 (46.4%) did not have a genetic diagnosis via trio-WES. Moreover, 58.6% (34/58) of the individuals included in the external validation cohort received a genetic diagnosis via trio-WES.

Independent predictive variable screening and logistic regression model establishment

As shown in [Table 2](#), the univariate logistic analysis revealed that five phenotypic indicators (GDD/ID severity, ADHD, NCC complexity, HCAs, and BMs) were potentially associated with a positive trio-WES diagnosis.

Following the univariate logistic analysis, the five candidate indicators were incorporated into a multivariate logistic regression model. As shown in [Table 2](#), the multivariate analysis results indicated

that GDD/ID severity (OR: 11.264; 95% CI: 4.788–26.495, $p < 0.001$), NCC complexity (2.671; 1.055–6.764, $p < 0.05$), HCAs (2.801; 1.070–7.332, $p < 0.05$), and BMs (3.701; 1.601–8.558, $p < 0.01$) were independently associated with a higher diagnostic efficacy when applying trio-WES in patients with LF-ASDs. In contrast, having ADHD was not independently associated with a higher possibility of receiving a genetic diagnosis by trio-WES in LF-ASDs children.

Collinearity diagnosis performed on the four candidate indicators (GDD/ID severity, NCC complexity, HCAs, and BMs) showed no significant evidence of collinearity, as the tolerances and the VIFs for each phenotypic factor were all >0.5 and <5 , respectively ([Table 3](#)).

Finally, based on the four-variable binary logistic regression β values and the intercept term, a regression model was established to predict the diagnostic efficacy of applying trio-WES in children with LF-ASDs. The corresponding formula for predicting the probability (P) of an individual with LF-ASDs being diagnosed by trio-WES is as follows: $\text{Logit}(P) = 2.422(\beta_1) \times \text{GDD/ID severity (severe-profound: 1; mild-moderate: 0)} + 0.983(\beta_2) \times \text{NCCs complexity (complicated: 1; simple: 0)} + 1.030(\beta_3) \times \text{HCAs (yes: 1; no: 0)} + 1.309(\beta_4) \times \text{BMs (yes: 1; no: 0)} - 1.898$ (intercept term).

TABLE 2 Univariate and multivariate logistic regression for predicting diagnostic efficacy of using trio-WES in 168 LF-ASDs children in training cohort.

Univariate logistic analysis			Multivariate logistic analysis		
Candidate indicators	OR (95% CI)	<i>p</i> value	Candidate indicators	OR (95% CI)	<i>p</i> value
Age of receiving trio-WES	0.981 (0.898–1.072)	0.675			
GDD/ID severity	15.836 (7.209–34.788)	< 0.001 ***	GDD/ID severity	11.264 (4.788–26.495)	< 0.001 ***
ADHD	2.473 (1.098–5.568)	0.029 *			
EP	1.584 (0.678–3.698)	0.288			
NCCs complexity	3.034 (1.431–6.433)	0.004 **	NCCs complexity	2.671 (1.055–6.764)	0.038 *
HCAs	5.950 (2.721–13.011)	< 0.001 ***	HCAs	2.801 (1.070–7.332)	0.036 *
BMs	5.140 (2.599–10.167)	< 0.001 ***	BMs	3.701 (1.601–8.558)	0.002 **

trio-WES, trio-based whole exome sequencing; LF-ASDs, low-function autism spectrum disorders; GDD/ID, global developmental delay/intellectual disability; NCCs, neurodevelopmental/neurological comorbid conditions; ADHD, attention deficit hyperactivity disorder; EP, epilepsy; HCAs, head circumference abnormalities; BMs, brain malformations; OR (95%CI), odds ratio (95% confidence interval). **p* < 0.05; ***p* < 0.01; ****p* < 0.001.

TABLE 3 The collinearity diagnostic analysis of indicators for predicting diagnostic efficacy of using trio-WES in LF-ASDs children in training cohort.

Candidate variables	Tolerance	VIF
GDD/ID severity	0.822	1.216
NCCs complexity	0.967	1.034
HCAs	0.823	1.215
BMs	0.891	1.123

Trio-WES, trio-based whole exome sequencing; LF-ASDs, low-function autism spectrum disorder; GDD/ID, global developmental delay/intellectual disability; ADHD, attention deficit hyperactivity disorder; NCCs, neurodevelopmental/neurological comorbid conditions; HCAs, head circumference abnormalities; BMs, brain malformations; VIF, variance inflation factor.

achieved a total score of approximately 140.5 points; the corresponding probability of obtaining a genetic diagnosis through trio-WES was approximately 80%. Indeed, this girl was diagnosed with “Developmental and Epileptic Encephalopathy 2 (OMIM#300672)” through trio-WES (Figure 2B). Another subject with LF-ASDs, a 9-year-old boy (subject No. 84 in the training cohort) with mild-moderate ID and multiple BMs detected via brain MRI, including basal ganglia lesions, pituitary dysplasia, and cerebellar atrophy, but without HCAs, ADHD, and epilepsy (simple NCCs), scored approximately 54 points, with the corresponding probability of a positive genetic diagnosis via trio-WES being approximately 35%. This boy has yet to receive a genetic diagnosis despite his multiple brain anomalies following comprehensive trio-WES analysis and re-evaluation of the trio-WES data (Figure 2C).

Predictive model visualization and nomogram usage

Figure 2 shows the nomogram plot based on the established logistic regression model. This nomogram provides an estimate of the individual probability of diagnosis via trio-WES through a score-contribution system and is designed for a child with LF-ASDs at the time of initial admission or during the pre-diagnosis stage. The methodology for this score-contribution system utilizes the coefficients (β) from the logistic regression model. The calculated score assigned to each indicator is proportional to its β value, mapped to a 0 to 100-point scale through linear transformation. Specifically, we assigned a value of 100 to the indicator with the maximum β value (β_{\max}), from which the scores for the other predictive indicators can be calculated using the formula: $\text{calculated score}_x = 100 \times \beta_x \div \beta_{\max}$. Detailed parameters related to this regression model and the corresponding model scores for each independent predictive variable are provided in Table 4.

As indicated in Figure 2A, LF-ASDs with severe-profound GDD/ID had the greatest influence on the probability of receiving a diagnosis through trio-WES, followed by BMs, HCAs, and complicated NCCs (NCC with coexisting ADHD and/or epilepsy). For example, a 4.5-year-old girl with LF-ASDs (subject no. 89 in the training group) with severe-profound GDD and comorbid ADHD and epilepsy (complicated NCCs), but without BMs and HCAs,

Assessment and internal validation of the nomogram performance in the training set

First, a calibration curve with Hosmer-Lemeshow testing was used to evaluate the fitness of the nomogram model within the training cohort. As demonstrated in Figure 3A, the calibration analysis indicated a good fit between the observed and model-predicted diagnostic probabilities ($\chi^2 = 4.275$; *p* value = 0.511), indicating satisfactory consistency between the predicted and observed values.

Subsequently, we used ROC curves to evaluate the discriminative ability of the nomogram. Figure 3B demonstrates an AUC of 0.868 (95% CI: 0.811–0.925), indicating good predictive performance in the training cohort. Based on the ROC plot of the training set, the maximal Youden index was 0.677 and was used to establish the optimal cutoff nomogram score (nomoScore) value = 54, generating a confusion matrix that yielded sensitivity, specificity, accuracy, precision, recall, and F1 score values of 85.56, 82.05, 83.93, 84.62, 85.56%, and 0.85, respectively, in the training set (Figure 4A; Table 5). These findings further underscore the nomogram’s promising capability in predicting the diagnostic probability of applying trio-WES in the diagnostic strategy of LF-ASDs in children.

Additionally, we used DCA and the CIC to assess the clinical usefulness of the nomogram model in the training set. As demonstrated in Figures 3C–D, children with LF-ASDs could receive

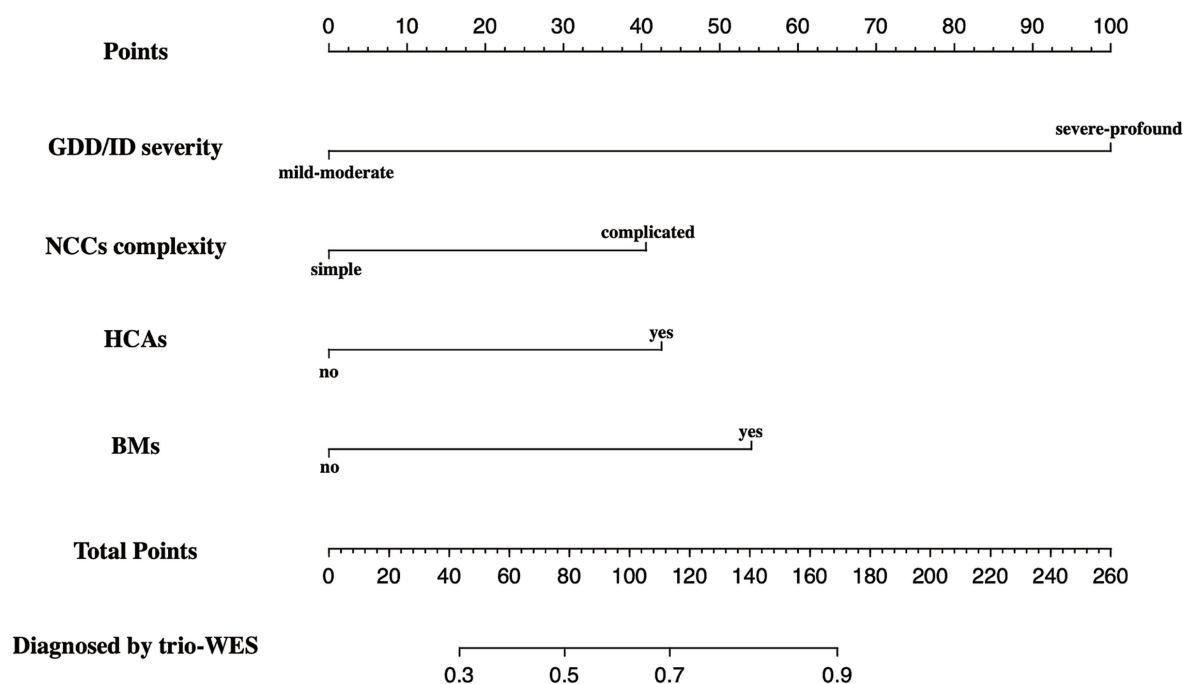
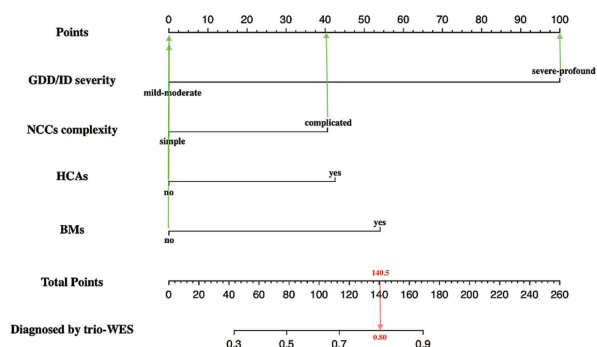
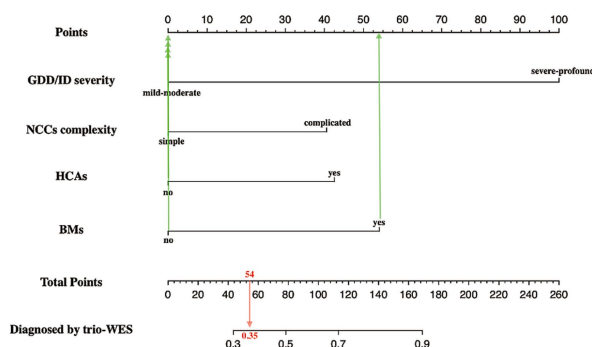
A**B****C**

FIGURE 2

Predictive nomogram for LF-ASDs children in the training set, estimating the probability of receiving a positive genetic diagnosis by applying trio-WES. **(A)** The nomogram plot has two parts: the top portion (from the "Point" section to the last the "BMs" section) is designed to calculate the respective point scores of each incorporated phenotypic indicator. The bottom portion (from the "Total Points" section to the "Diagnosis by trio-WES" section) is used to analyze the probability of having a positive genetic diagnosis via trio-WES for each enrolled LF-ASDs subject. **(B,C)** represent two examples with high and low probabilities of receiving a genetic diagnosis via trio-WES, respectively. The red arrow in **(B)** reveals an LF-ASDs child with an approximate total score of 154, and the matched predicted probability of receiving a genetic diagnosis via trio-WES is approximately 85%, whereas the red arrow in **(C)** indicates an LF-ASDs subject with an approximate total score of 54. The matched predicted probability of receiving a genetic diagnosis via trio-WES is approximately 35%. LF-ASDs, low-function autism spectrum disorders; GDD/ID, global developmental delay/intellectual disability; NCCs, neurodevelopmental/neurological comorbid conditions; HCAs, head circumference abnormalities; BMs, brain malformations; trio-WES, trio-based whole-exome sequencing.

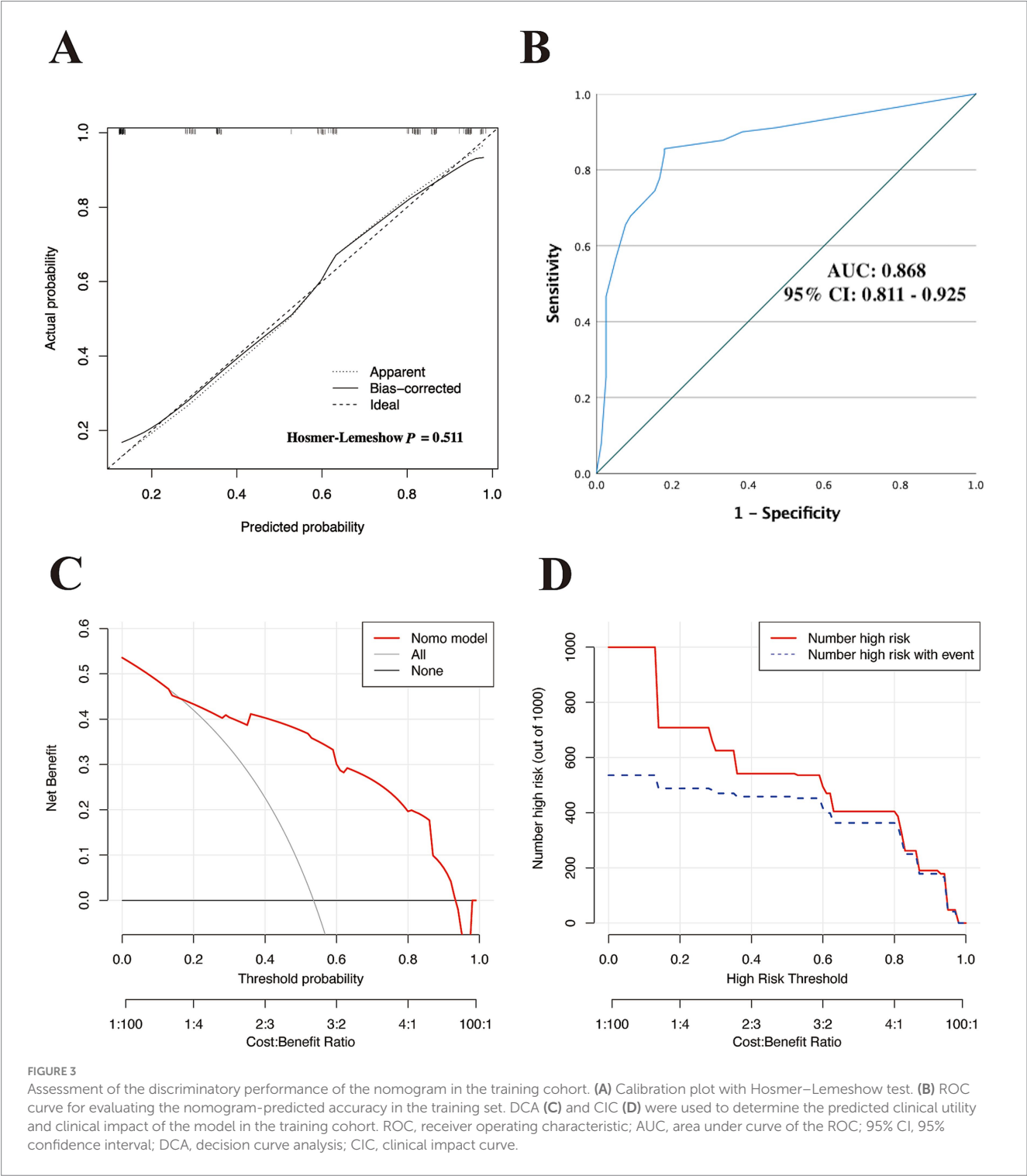
greater net benefits from this nomogram compared to hypothetical treat-none or treat-all scenarios, suggesting that applying this model to predict the diagnostic efficacy of trio-WES for LF-ASDs patients may yield significant benefits.

Finally, we used 10-fold cross-validation and bootstrapping re-for internal validation to determine the generalization performance of the nomogram. As depicted in Figure 5A C-index value following the 10-fold cross-validation was 0.860 (95% CI: 0.785–0.935).

TABLE 4 Coefficients of binary logistic regression for predicting diagnostic efficacy via trio-WES in individuals with LF-ASDs in training set.

Phenotypic variables	B	S. E.	Wald	<i>p</i> value	OR	95% CI for OR	Calculated score
GDD/ID severity	2.422 (β_1)	0.436	30.789	<0.001	11.264	4.788–26.495	100*
NCCs complexity	0.983 (β_2)	0.474	4.297	0.038	2.671	1.055–6.764	40.5 ($100 \times \beta_2 \div \beta_1$)
HCAs	1.030 (β_3)	0.491	4.402	0.036	2.801	1.070–7.332	42.5 ($100 \times \beta_3 \div \beta_1$)
BMs	1.309 (β_4)	0.428	9.364	0.002	3.701	1.601–8.558	54 ($100 \times \beta_4 \div \beta_1$)

Trio-WES, trio-based whole exome sequencing; LF-ASDs, low-function autism spectrum disorders; GDD/ID, global developmental delay/intellectual disability; NCCs, neurodevelopmental/neurological comorbid conditions; HCAs, head circumference abnormalities; BMs, brain malformations; B, β value; S. E., standard error; OR, odds ratio; 95% CI, 95% confidence interval. * The variable showing the max β value and being set 100-point as reference-point for other included variables.



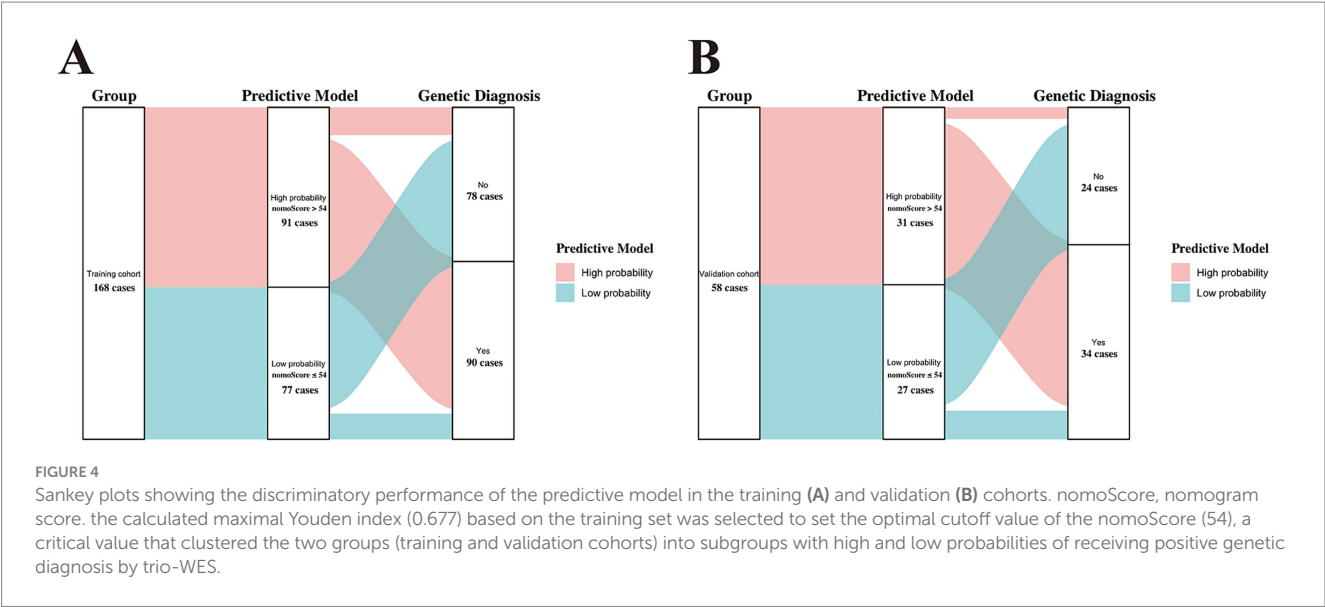


TABLE 5 Predictive performance of the constructed phenotype-driven nomogram model in training and validation sets.

Predictive values	Training set	Validation set
Sensitivity (%)	85.56%	85.29%
Specificity (%)	82.05%	91.67%
Accuracy (%)	83.93%	87.93%
Precision (%)	84.62%	93.55%
Recall (%)	85.56%	85.29%
F1 score	0.85	0.89

Similarly, the bootstrap method with 1,000 resamples yielded a C-index of 0.856 (95% CI: 0.776–0.939; Figure 5B). Collectively, these results indicate that the nomogram exhibits good stability with excellent consistence and no evidence of overfitting.

External validation of the nomogram performance in the independent validation set

Based on the optimal cutoff value (nomoScore = 54), all 58 cases within the independent validation cohort were classified into 31 nomogram-predicted positive diagnostic cases and 27 nomogram-predicted negative diagnostic cases. As illustrated in Figure 6A, the calibration curve with the Hosmer-Lemeshow test demonstrated excellent agreement between the nomogram-predicted values and the actual observed results ($\chi^2 = 1.125$, p value = 0.952) in the transformed external set. Additionally, a ROC plot was used to validate the discriminative performance of the model in the transformed external set, which revealed robust discriminative ability (AUC: 0.941; 95% CI: 0.880–0.998; Figure 6B). The results of DCA and CIC within the transformed external cohort indicated that employing this nomogram to predict the diagnostic efficacy of trio-WES for LF-ASDs children could yield significant net benefits (Figures 6C,D). The confusion matrix results illustrated in the Sankey plot for the external set revealed

sensitivity, specificity, accuracy, precision, and F1 score of the nomogram were 85.29, 91.67, 87.93, 93.55%, and 0.89, respectively (Figure 4B; Table 5). These external validation results suggest that the proposed model demonstrates stable reproducibility and robust repeatability.

Discussion

The rapid development of next-generation sequencing technology has enabled the identification of genetic components in many children with unexplained neurological syndromes, including LF-ASDs (24, 31–33). The application of trio-WES has transformed the landscape of clinical genetics, facilitating a more cost-effective means of obtaining diagnoses for various Mendelian disorders compared to traditional genetic tests, such as target-panel sequencing with family phenotype segregation analysis. This shift has alleviated the “diagnostic odysseys” frequently encountered by affected children and their families (34–37). However, it is essential to recognize the technical limitations of trio-WES and the complexities associated with human genomic disturbances, such as intronic structural variants and non-coding variants, which may hinder its diagnostic effectiveness. Whole-genome sequencing (WGS) could potentially address these limitations by facilitating the identification of intronic structural or non-coding variants. Nevertheless, the high costs associated with WGS significantly restrict its clinical application and widespread use as a routine genetic diagnostic approach is largely restricted (38, 39). To date, trio-WES remains the first-tier genetic diagnosis option globally and constitutes a critical element of subsequent genetic counseling and patient management for many Mendelian disorders (40). Thus, developing effective approaches and tools to analyze the diagnostic rate of trio-WES in clinical contexts related to various idiopathic and complex disorders, including LF-ASDs, remains a pertinent endeavor.

Numerous factors, such as disorder type, disease onset age, and variant capture strategy, may influence the diagnostic efficacy of trio-WES in clinical practice (12, 38). For instance, the diagnostic yield of trio-WES may reach 92% in children with idiopathic dermatological syndromes (12), potentially due to the clear

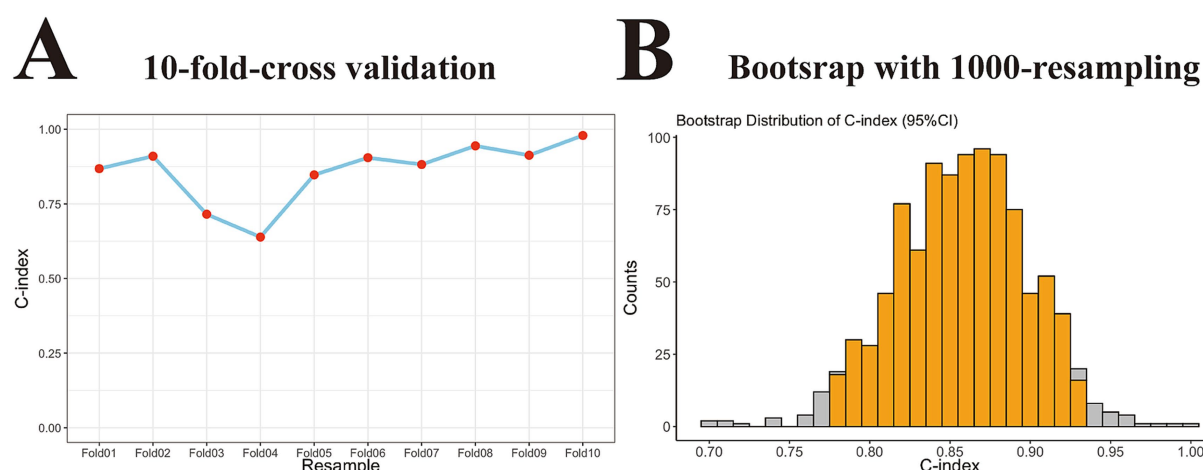


FIGURE 5

Internal validation of the generalization performance of the nomogram model in training cohort. **(A)** Point-fold line chart with 10-fold-cross validation approach showing the nomogram had good stability with excellent consistence in training set (C-index, 0.860 with 95% CI, 0.785–0.935). **(B)** Histogram with 1,000-time resampling bootstrap method revealing the nomogram did not overfit in training set (C-index, 0.856 with 95% CI, 0.776–0.939). C-index, concordance index; 95% CI, 95% confidence interval.

presentation and delineation of phenotypic information evident in those conditions, which likely contributes to the elevated diagnostic rate achievable through trio-WES. Thus, accurate and comprehensive assessments of clinical phenotypes at the pre-diagnosis stage are paramount and pose considerable challenges, necessitating rigorous collection and precise analysis of various phenotypic features for every patient. In this study, we meticulously recorded all associated phenotypic clues that accurately reflect the neurological conditions of each child with LF-ASDs to enhance the diagnostic yield. The global diagnostic yields in the training and external validation groups were 53.6 and 58.6%, respectively; these figures are similar to those documented in a previous WES report involving LF-ASDs children with additional associated conditions (51.3%) (1). Our findings regarding the global diagnostic rate associated with the implementation of trio-WES in children with LF-ASDs further reinforce prior conclusions that complex phenotypic features—encompassing multiple neurological disorders (including ADHD and/or epilepsy), multiple neurological anomalies (such as HCAs and/or BMs), and severe-profound levels of cognitive or developmental impairment—are more likely linked to an exon-level variant within the clinical setting. This observation implies a potential relationship between the enrichment of phenotypic characteristics and the diagnostic yield of trio-WES in children with LF-ASDs (1, 24). We thus propose the possibility of establishing a diagnostic predictive model for the application of trio-WES in LF-ASDs patients by incorporating key phenotypic factors associated with a higher probability of obtaining genetic results. This model could help pediatricians make more appropriate and personalized management for affected children during the pre-diagnosis stage.

The present study identified four key phenotypic indicators (GDD/ID severity, NCC complexity, HCAs, and BMs) as being associated with the possibility of obtaining genetic results through trio-WES in children with LF-ASDs. We hypothesize that severe-profound GDD/ID, BMs, HCAs, and a broad spectrum of NCCs associated with LF-ASDs may share common genetic backgrounds

linked to overlapping genetic factors, which ultimately results in a higher trio-WES diagnostic rate. Over 1,200 genes related to ASDs susceptibility (called ASDs-related genes) have been cataloged in the Simons Foundation Autism Research Initiative⁸ gene dataset (41); the two major gene fall into two functional categories: those involved in gene expression regulation (mainly chromatin modification and transcription regulation) and those involved in neuronal communication (mainly synaptic communication and ion channel regulation) (42, 43). This suggests that dysregulation in gene expression and neuronal communication may significantly contribute to the genetic components underlying syndromic ASDs. We speculate that alterations in the genetic functions pertinent to expression regulation and neuronal communication may be fundamental contributors to the genetic components associated with severe-profound GDD/ID and multiple NCCs inherent in syndromic ASDs. Furthermore, neuronal communication between the craniofacial ectoderm and neural crest cells are vital for craniofacial patterning and morphogenesis during craniofacial development (44). We, therefore, speculate that alterations in these neuronal communication-related genes can cause disruption between the craniofacial ectoderm and neural crest cells, leading to a broad spectrum of craniofacial anomalies, among which BMs and HCAs are prominent phenotypic features. However, these speculations warrant further exploration through in-depth mechanistic experiments.

Additionally, previous research had demonstrated that the four phenotypic features—severe to profound GDD/ID, complicated NCC complexity, the presence of HCAs, and BMs—are strong indicators of rare monogenic neurodevelopmental disorders (24, 45). moreover, variants at exon-level have been recognized as the main cause of rare monogenic neurodevelopmental disorders (46). Given the close relationship between these elements, it is reasonable to infer that

⁸ SFARI, <http://gene.sfari.org/>

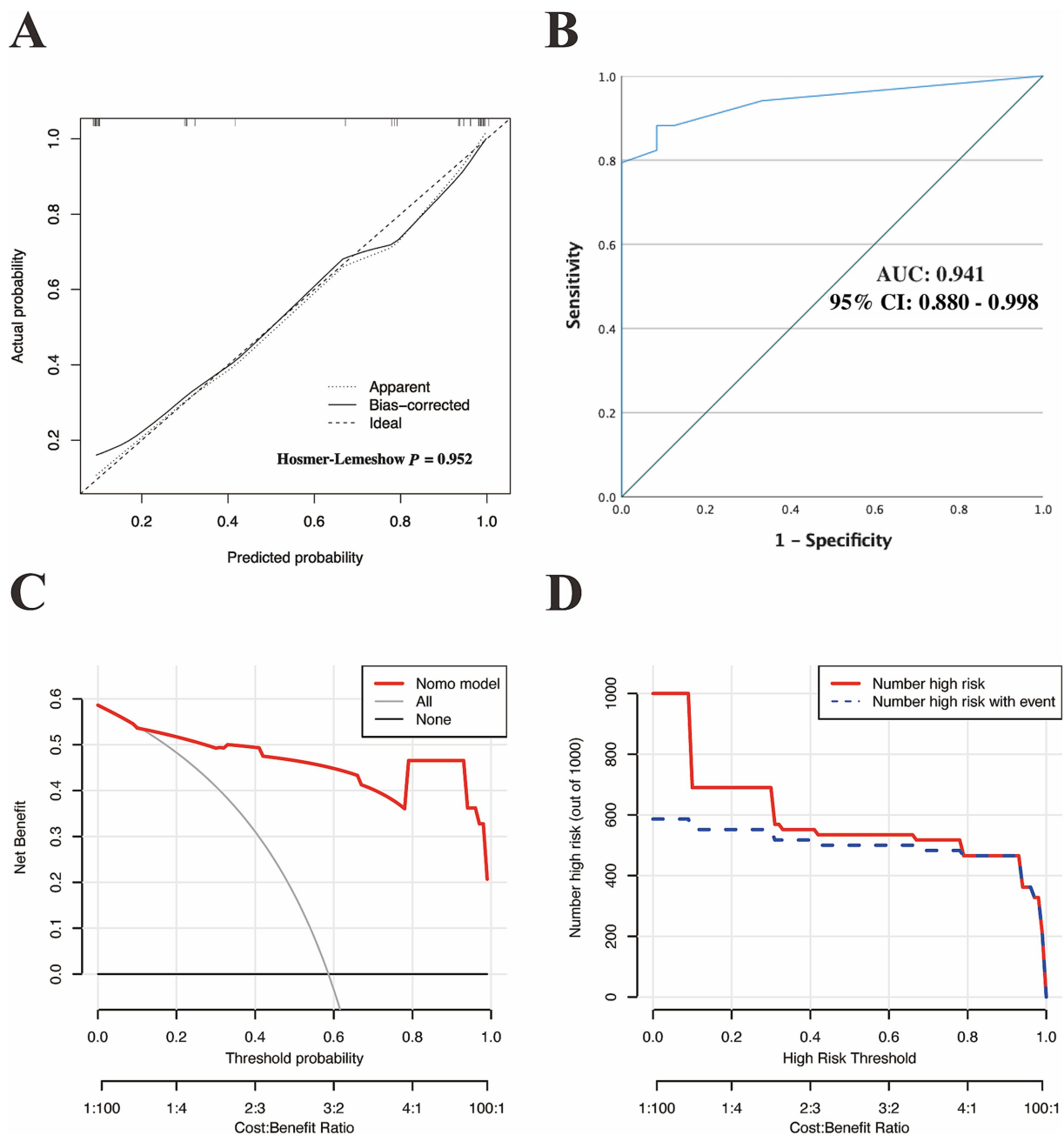


FIGURE 6

External validation of the discriminatory performance of the nomogram model in an independent external cohort. (A) Calibration plot with Hosmer–Lemeshow test. (B) ROC curve verifying the nomogram-predicted accuracy in the external set. DCA (C) and the CIC (D) were used to verify the clinical value of the model in an external cohort. ROC, receiver operating characteristic; AUC, area under curve of the ROC; 95% CI, 95% confidence interval; DCA, decision curve analysis; CIC, clinical impact curve.

individuals with LF-ASDs exhibiting multiple phenotypic indicators associated with rare monogenic disorders may possess greater probabilities of harboring relevant exon-level variants, thereby facilitating a more straightforward diagnosis via trio-WES. However, LF-ASDs is a very complicated disorder with a high heterogeneity of genetic abnormalities including variants at exon-level and at out-exon-level (such as non-coding variation, epigenetics, and polygenic effects); given the technical limitation (exon-level sequencing only) of trio-WES, there is definitely a part of LF-ASDs patients cannot get an exact genetic diagnosis by using trio-WES alone because their genetic

components may lie outside exon regions. Because of the lack of effective diagnostic tools (such as WGS and RNA-seq) applied in current study, we cannot determine whether the specific causes of those cases with negative trio-WES diagnosis in our study are non-coding variants, epigenetics, polygenic effects or other unknown genetic factors. Further analysis of these cases are needed to ascertain their specific etiologies and will be a focus of our future investigations.

Notably, the current study established a logistic regression model based on the four easily obtained phenotypic indicators and showed good calibration and discrimination with high accuracy and precision

within both the training and validation LF-ASDs groups, revealing promising clinical applications. Moreover, considering the diagnostic yield (53.6%) and the number of identified variables (four) in the training cohort, in conjunction with the 10 events per variable, i.e., 10 EPV, the expected total number of cases in the training set should be over 75 ($4 \times 10 \div 0.536$) (47). In actuality, the training cohort comprised 168 cases, which is substantially greater than 75, further affirming the reliability and robustness of the constructed nomogram.

However, the current study has several limitations. First, this study adopted a dual-center design, whereby the case sources for the training set (from a tertiary hospital) and the validation set (from a specialized LF-ASDs rehabilitation center) differed, potentially introducing selection bias. A multicenter study with a consistent case source (where all training and validation cases come from tertiary hospitals) are needed to further validate our predictive model. Second, our nomogram was constructed and validated within professional medical institutions, leaving its performance in primary medical institutions undetermined. Lastly, the study's scope was not comprehensive enough, as it only considered ADHD and epilepsy as common NCCs in children with LF-ASDs. Further improvements, such as applying cutting-edge and Chinese version assessment tools to enable objective assessment of sleep and anxiety disorders in young children, allowing for the incorporation of such phenotypic variables into our predictive model, are required to further enhance the reliability of the model.

Conclusion

We developed and validated an user-friendly nomogram based on common, objective, and easily obtained phenotypic indicators related to neurological conditions to predict the individualized diagnostic probability of trio-WES in children with LF-ASDs. This tool could assist affected children and their families in estimating their personalized diagnostic probability and selecting more suitable diagnostic strategies at the pre-diagnosis stage, ultimately reducing unnecessary financial expenditures. Additionally, this nomogram may enable pediatricians to identify children with LF-ASDs at high risk for relevant genetic factors at the early admission stage, facilitating more individualized patient management and subsequent genetic counseling.

Data availability statement

The original contributions presented in the study are included in the article/[Supplementary material](#), further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by the Ethical Committee of the Sun Yat-sen Memorial Hospital, Sun Yat-sen University. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

RW: Formal analysis, Writing – original draft, Validation, Project administration, Visualization, Data curation, Conceptualization, Software, Investigation. XL: Resources, Writing – review & editing, Supervision. ZH: Writing – review & editing, Resources. ZM: Resources, Writing – review & editing. WT: Writing – review & editing, Resources, Methodology, Supervision. LL: Supervision, Writing – review & editing, Funding acquisition, Resources.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Acknowledgments

We thank all the children and parents for their participation in this study.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fneur.2025.1597588/full#supplementary-material>

References

- Chen S, Xiong J, Chen B, Zhang C, Deng X, He F, et al. Autism spectrum disorder and comorbid neurodevelopmental disorders (ASD-NDDs): clinical and genetic profile of a pediatric cohort. *Clin Chim Acta*. (2022) 524:179–86. doi: 10.1016/j.cca.2021.11.014
- Maenner MJ, Shaw KA, Baio J, Washington A, Patrick M, DiRienzo M, et al. Prevalence of autism spectrum disorder among children aged 8 years - autism and developmental disabilities monitoring network, 11 sites, United States, 2016. *MMWR Surveill Summ*. (2020) 69:1–12. doi: 10.15585/mmwr.ss6904a1
- Xu G, Strathearn L, Liu B, Bao W. Prevalence of autism Spectrum disorder among US children and adolescents, 2014–2016. *JAMA*. (2018) 319:81–2. doi: 10.1001/jama.2017.17812
- Li Q, Li Y, Liu B, Chen Q, Xing X, Xu G, et al. Prevalence of autism Spectrum disorder among children and adolescents in the United States from 2019 to 2020. *JAMA Pediatr*. (2022) 176:943–5. doi: 10.1001/jamapediatrics.2022.1846
- GBD 2019 Mental Disorders Collaborators. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: a systematic analysis for the global burden of disease study 2019. *Lancet Psychiatr*. (2022) 9:137–50. doi: 10.1016/S2215-0366(21)00395-3
- Xu S, Li M, Yang C, Fang X, Ye M, Wei L, et al. Altered functional connectivity in children with low-function autism Spectrum disorders. *Front Neurosci*. (2019) 13:806. doi: 10.3389/fnins.2019.00806
- Srivastava S, Love-Nichols JA, Dies KA, Ledbetter DH, Martin CL, Chung WK, et al. Meta-analysis and multidisciplinary consensus statement: exome sequencing is a first-tier clinical diagnostic test for individuals with neurodevelopmental disorders. *Genet Med*. (2019) 21:2413–21. doi: 10.1038/s41436-019-0554-6
- Wu R, Li X, Chen Z, Shao Q, Zhang X, Tang W, et al. Development and validation of a nomogram based on common biochemical indicators for survival prediction of children with high-risk neuroblastoma: a valuable tool for resource-limited hospitals. *BMC Pediatr*. (2023) 23:426. doi: 10.1186/s12887-023-04228-2
- Gao T, Yang L, Zhou J, Zhang Y, Wang L, Wang Y, et al. Development and validation of a nomogram prediction model for ADHD in children based on individual, family, and social factors. *J Affect Disord*. (2024) 356:483–91. doi: 10.1016/j.jad.2024.04.069
- Costantine MM, Tita ATN, Mele L, Casey BM, Peaceman AM, Varner MW, et al. The association between infant birth weight, head circumference, and neurodevelopmental outcomes. *Am J Perinatol*. (2024) 41:e1313–23. doi: 10.1055/s-0043-1761920
- Parkhurst JT, Vesco AT, Ballard RR, Lavigne JV. Improving diagnostic accuracy: comparison of nomograms and classification tree analyses for predicting the diagnosis of oppositional defiant disorder. *Psychol Serv*. (2023) 20:184–95. doi: 10.1037/ser0000670
- Marinakos NM, Svingou M, Veltra D, Kekou K, Sofocleous C, Tilemis FN, et al. Phenotype-driven variant filtration strategy in exome sequencing toward a high diagnostic yield and identification of 85 novel variants in 400 patients with rare Mendelian disorders. *Am J Med Genet A*. (2021) 185:2561–71. doi: 10.1002/ajmg.a.62338
- First MB. Diagnostic and statistical manual of mental disorders, 5th edition, and clinical utility. *J Nerv Ment Dis*. (2013) 201:727–9. doi: 10.1097/NMD.0b013e3182a2168a
- Haem E, Doostfateme N, Firouzabadi N, Ghazanfari N, Karlsson MO. A longitudinal item response model for aberrant behavior checklist (ABC) data from children with autism. *J Pharmacokinet Pharmacodyn*. (2020) 47:241–53. doi: 10.1007/s10928-020-09686-0
- Chakraborty S, Bhatia T, Antony N, Roy A, Shriharsh V, Sahay A, et al. Comparing the Indian autism screening questionnaire (IASQ) and the Indian scale for assessment of autism (ISAA) with the childhood autism rating scale-second edition (CARS2) in Indian settings. *PLoS One*. (2022) 17:273780. doi: 10.1371/journal.pone.0273780
- Robins DL, Casagrande K, Barton M, Chen CM, Dumont-Mathieu T, Fein D. Validation of the modified checklist for autism in toddlers, revised with follow-up (M-CHAT-R/F). *Pediatrics*. (2014) 133:37–45. doi: 10.1542/peds.2013.1813
- Accardo PJ. 50 years ago in the journal of pediatrics: the predictability of Gesell developmental scales in mongolism. *J Pediatr*. (2013) 162:55. doi: 10.1016/j.jpeds.2012.07.060
- Na SD, Burns TG. Wechsler intelligence scale for children-V: test review. *Appl Neuropsychol Child*. (2016) 5:156–60. doi: 10.1080/21622965.2015.1015337
- Scheffer IE, Berkovic S, Capovilla G, Connolly MB, French J, Guilhoto L, et al. ILAE classification of the epilepsies: position paper of the ILAE Commission for Classification and Terminology. *Epilepsia*. (2017) 58:512–21. doi: 10.1111/epi.13709
- Han P, Wei G, Cai K, Xiang X, Deng WP, Li YB, et al. Identification and functional characterization of mutations in LPL gene causing severe hypertriglyceridaemia and acute pancreatitis. *J Cell Mol Med*. (2020) 24:1286–99. doi: 10.1111/jcmm.14768
- Zhang R, Chen S, Han P, Chen F, Kuang S, Meng Z, et al. Whole exome sequencing identified a homozygous novel variant in CEP290 gene causes Meckel syndrome. *J Cell Mol Med*. (2020) 24:1906–16. doi: 10.1111/jcmm.14887
- Dai Y, Liang S, Dong X, Zhao Y, Ren H, Guan Y, et al. Whole exome sequencing identified a novel DAG1 mutation in a patient with rare, mild and late age of onset muscular dystrophy-dystroglycanopathy. *J Cell Mol Med*. (2019) 23:811–8. doi: 10.1111/jcmm.13979
- Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. (2015) 17:405–24. doi: 10.1038/gim.2015.30
- Wu R, Li X, Meng Z, Li P, He Z, Liang L. Phenotypic and genetic analysis of children with unexplained neurodevelopmental delay and neurodevelopmental comorbidities in a Chinese cohort using trio-based whole-exome sequencing. *Orphanet J Rare Dis*. (2024) 19:205. doi: 10.1186/s13023-024-03214-w
- Riggs ER, Andersen EF, Cherry AM, Kantarci S, Kearney H, Patel A, et al. Technical standards for the interpretation and reporting of constitutional copy-number variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics (ACMG) and the clinical genome resource (ClinGen). *Genet Med*. (2020) 22:245–57. doi: 10.1038/s41436-019-0686-8
- Yuan H, Shangguan S, Li Z, Luo J, Su J, Yao R, et al. CNV profiles of Chinese pediatric patients with developmental disorders. *Genet Med*. (2021) 23:669–78. doi: 10.1038/s41436-020-01048-y
- Xu B, Gao Y, Zhang Q, Li X, Liu X, Du J, et al. Establishment and validation of a multivariate predictive model for the efficacy of oral rehydration salts in children with postural tachycardia syndrome. *EBioMedicine*. (2024) 100:104951. doi: 10.1016/j.ebiom.2023.104951
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. (1977) 33:159–74. doi: 10.2307/2529310
- Tang W, Shao Q, He Z, Zhang X, Li X, Wu R. Clinical significance of nonerythrocytic spectrin Beta 1 (SPTBN1) in human kidney renal clear cell carcinoma and uveal melanoma: a study based on Pan-Cancer analysis. *BMC Cancer*. (2023) 23:303. doi: 10.1186/s12885-023-10789-3
- Wu R, Tang W, Qiu K, Li P, Li Y, Li D, et al. An integrative Pan-Cancer analysis of the prognostic and immunological role of casein kinase 2 alpha protein 1 (CSNK2A1) in human cancers: a study based on bioinformatics and Immunohistochemical analysis. *Int J Gen Med*. (2021) 14:6215–32. doi: 10.2147/IJGM.S330500
- Wu R, Tang W, Li P, Meng Z, Li X, Liang L. Identification of a novel phenotype of external ear deformity related to coffin-Siris syndrome-9 and literature review. *Am J Med Genet A*. (2024):63626. doi: 10.1002/ajmg.a.63626
- Wu RH, Tang WT, Qiu KY, Li XJ, Tang DX, Meng Z, et al. Identification of novel CSNK2A1 variants and the genotype-phenotype relationship in patients with Okur-Chung neurodevelopmental syndrome: a case report and systematic literature review. *J Int Med Res*. (2021) 49:1–16. doi: 10.1177/03000605211017063
- Wu R, Li Y, He Z, Meng Z, Tang W, Liang L. Case report: observation of early-onset high myopia with fundus tessellation changes in coffin-Siris syndrome 9 (CSS9) and literature review. *Front Pediatr*. (2025) 13:1603863. doi: 10.3389/fped.2025.1603863
- Bertier G, Héту M, Joly Y. Unsolved challenges of clinical whole-exome sequencing: a systematic literature review of end-users' views. *BMC Med Genet*. (2016) 9:52. doi: 10.1186/s12920-016-0213-6
- Iglesias A, Anyane-Yeboah K, Wynn J, Wilson A, Truitt Cho M, Guzman E, et al. The usefulness of whole-exome sequencing in routine clinical practice. *Genet Med*. (2014) 16:922–31. doi: 10.1038/gim.2014.58
- Niguidula N, Alamillo C, Shahmirzadi Mowlavi L, Powis Z, Cohen JS, Farwell Hagman KD. Clinical whole-exome sequencing results impact medical management. *Mol Genet Genomic Med*. (2018) 6:1068–78. doi: 10.1002/mggg.3484
- Yang Y, Muzny DM, Reid JG, Bainbridge MN, Willis A, Ward PA, et al. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N Engl J Med*. (2013) 369:1502–11. doi: 10.1056/NEJMoa1306555
- Taylor J, Craft J, Blair E, Wordsworth S, Beeson D, Chandratte S, et al. Implementation of a genomic medicine multi-disciplinary team approach for rare disease in the clinical setting: a prospective exome sequencing case series. *Genome Med*. (2019) 11:46. doi: 10.1186/s13073-019-0651-9
- White SJ, Laros JFJ, Bakker E, Cambon-Thomsen A, Eden M, Leonard S, et al. Critical points for an accurate human genome analysis. *Hum Mutat*. (2017) 38:912–21. doi: 10.1002/humu.23238
- Manickam K, McClain MR, Demmer LA, Biswas S, Kearney HM, Malinowski J, et al. Exome and genome sequencing for pediatric patients with congenital anomalies or intellectual disability: an evidence-based clinical guideline of the American College of Medical Genetics and Genomics (ACMG). *Genet Med*. (2021) 23:2029–37. doi: 10.1038/s41436-021-01242-6
- Banerjee-Basu S, Packer A. SFARI gene: an evolving database for the autism research community. *Dis Model Mech*. (2010) 3:133–5. doi: 10.1242/dmm.005439
- Ye Z, McQuillan L, Poduri A, Green TE, Matsumoto N, Mefford HC, et al. Somatic mutation: the hidden genetics of brain malformations and focal epilepsies. *Epilepsy Res*. (2019) 155:106161. doi: 10.1016/j.eplepsyres.2019.106161
- Satterstrom FK, Kosmicki JA, Wang J, Breen MS, De Rubeis S, An JY, et al. Large-scale exome sequencing study implicates both developmental and functional changes in the neurobiology of autism. *Cell*. (2020) 180:568–584.e23. doi: 10.1016/j.cell.2019.12.036

44. Hiraide T, Yamoto K, Masunaga Y, Asahina M, Endoh Y, Ohkubo Y, et al. Genetic and phenotypic analysis of 101 patients with developmental delay or intellectual disability using whole-exome sequencing. *Clin Genet.* (2021) 100:40–50. doi: 10.1111/cge.13951
45. Russ JB, Stone AC, Maney K, Morris LC, Wright CF, Hurst JH, et al. Cell-specific expression biases in human cortex of genes associated with neurodevelopmental disorders. *Sci Rep.* (2025) 15:23172. doi: 10.1038/s41598-025-05117-7
46. Rauch A, Wieczorek D, Graf E, Wieland T, Ende S, Schwarzmayr T, et al. Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *Lancet.* (2012) 380:1674–82. doi: 10.1016/S0140-6736(12)61480-9
47. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol.* (1996) 49:1373–9. doi: 10.1016/S0895-4356(96)00236-3