

---

# Integrating Genomics and Deep Phenotyping for Diagnosing Rare Pediatric Neurological Diseases: Potential for Sustainable Healthcare

---

[Nigara YerKhojayeveva](#)\*, [Nazira Zharkinbekova](#), [Sovet Azhayev](#), [Ainash Oshibayeva](#), [Gulnaz Nuskabayeva](#), [Rauan Kaiyrzhanov](#)

Posted Date: 30 June 2025

doi: 10.20944/preprints202506.2423.v1

Keywords: rare pediatric neurological diseases; deep phenotyping; Human Phenotype Ontology; exome sequencing; diagnostic odyssey



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

# Integrating Genomics and Deep Phenotyping for Diagnosing Rare Pediatric Neurological Diseases: Potential for Sustainable Healthcare

Nigara YerKhojayeva <sup>1,\*</sup>, Nazira Zharkinbekova <sup>2</sup>, Sovet Azhayev <sup>1</sup>, Ainash Oshibayeva <sup>1</sup>, Gulnaz Nuskabayeva <sup>1</sup> and Rauan Kaiyrzhanov <sup>3</sup>

<sup>1</sup> Khoja Akhmet Yassawi International Kazakh-Turkish University, Faculty of Medicine, Turkestan 161200, Kazakhstan

<sup>2</sup> South Kazakhstan Medical Academy, Department of Neurology, Psychiatry, Rehabilitation and Neurosurgery, Shymkent 160019, Kazakhstan

<sup>3</sup> University College London, Institute of Neurology, Department of Neuromuscular Diseases, London WC1N 3BG, United Kingdom

\* Correspondence: nigara.erkhodzhaeva@ayu.edu.kz; Tel.: +7 702 882 04 00

## Abstract

**Background:** Rare pediatric neurological diseases (RPND) often elude timely diagnosis, resulting in prolonged and costly diagnostic odysseys. Integration of Human Phenotype Ontology (HPO)-based deep phenotyping with exome sequencing (ES) and reverse phenotyping may improve diagnostic yield and efficiency, especially in resource-limited settings. **Objectives:** To assess the diagnostic yield and clinical impact of an integrated approach combining deep phenotyping, ES, and reverse phenotyping in children with suspected RPND in a multi-center, resource-limited setting. **Methods:** Eighty-one children with suspected RPND from 11 hospitals in South Kazakhstan were enrolled via the Central Asian and Transcaucasian Rare Pediatric Neurological Diseases Genomic Consortium (CAT-RPND). ES was performed by 3Billion (South Korea). All patients underwent HPO-based deep phenotyping, with variant classification according to American College of Medical Genetics and Genomics (ACMG) guidelines. Reverse phenotyping and interdisciplinary case discussions supported interpretation. **Results:** Molecular diagnoses were achieved in 43 of 81 patients (53%), including 18 pathogenic, 12 likely pathogenic, and 9 variants of uncertain significance (VUS). Reverse phenotyping refined or expanded phenotypes in 33% of diagnosed cases and supported likely pathogenicity in 8 of 9 VUS. The integrated approach reduced the median diagnostic odyssey from 72 to <5 months and the median number of procedures from 20 to 2 (Wilcoxon  $p = 1.91 \times 10^{-6}$ ; Cohen's  $d = 2.43$ ). **Conclusions:** Combining deep phenotyping, ES, and reverse phenotyping improved diagnostic outcomes and shortened the diagnostic journey. This approach minimizes unnecessary procedures and delays, offering scalable value for sustainable healthcare in resource-limited settings.

**Keywords:** rare pediatric neurological diseases; deep phenotyping; Human Phenotype Ontology; exome sequencing; diagnostic odyssey

## 1. Introduction

Rare pediatric neurological diseases (RPND) represent one of the most complex challenges in modern medicine. It is known that more than 7,000 rare diseases have been registered, approximately 80% of which have a genetic basis and may first manifest during childhood, including various neurological pathologies. Although each individual diagnosis is rare (approximately 1 in 10,000–40,000 newborns), rare diseases collectively affect 6–8% of the population, corresponding to over 300 million people worldwide. These figures are supported by international organizations such as EURORDIS (<https://www.eurordis.org>) and Orphanet (<https://www.orpha.net>) [1].

RPND are characterized by high clinical complexity, heterogeneity of presentation, often unclear etiology, and progressive course, leading to severe disability, reduced quality of life, and increased mortality in children. Traditional diagnostic methods often result in prolonged diagnostic processes involving numerous examinations, consultations, and repeated testing. According to Orphanet, approximately 70% of rare diseases begin in childhood, and the National Organization for Rare Disorders (NORD) reports that the diagnostic odyssey may last from 3 to 7 years (<https://rarediseases.org/rare-disease-information/what-are-rare-diseases/>). A study published in *Genetics in Medicine* showed that more than 60% of patients with rare genetic diseases underwent over five standard diagnostic procedures before receiving an accurate diagnosis, which significantly increases the financial burden on the healthcare system and exerts substantial emotional stress on families [2,3]. These data highlight the need to adopt new methodological approaches capable of reducing diagnostic time and resource expenditures.

In recent years, significant progress has been made in the fields of deep phenotyping and genomics, offering innovative opportunities for diagnostics. Deep phenotyping enables the identification of even subtle clinical features that may be overlooked during standard clinical assessments [4]. Robinson (2012) emphasized that detailed phenotypic descriptions are a cornerstone of implementing precision medicine principles [5]. When integrated with genomic technologies such as exome sequencing (ES) or genome sequencing (GS), it becomes possible to accurately correlate phenotypic manifestations with genetic mutations, thereby substantially increasing diagnostic accuracy [6].

In several regions, access to advanced next-generation sequencing (NGS) technologies remains limited [7]. In settings where the use of costly sequencing platforms is constrained, deep phenotyping can serve as an effective pre-laboratory method for patient selection for subsequent genetic testing. This approach narrows the range of suspected genetic conditions based on clinical presentation, contributing to more rational resource utilization and reducing the overall cost of molecular genetic diagnostics.

The integration of deep phenotyping and genomics is crucial for sustainable healthcare development. Accurate and early diagnosis optimizes the use of medical resources, reduces the need for additional testing, and ensures equitable access to advanced technologies, which is particularly relevant in resource-limited settings. The application of this integrated approach facilitates the timely initiation of targeted treatment, improves clinical outcomes, and alleviates both emotional and financial burdens on patient families. This approach aligns with the United Nations Sustainable Development Goals (SDGs), particularly SDG 3 (Good Health and Well-being), SDG 9 (Industry, Innovation and Infrastructure), and SDG 10 (Reduced Inequalities) (<https://sdgs.un.org/goals>) [8].

**The Study Objective** is to develop and implement an integrated diagnostic approach combining deep phenotyping using the Human Phenotype Ontology (HPO), ES with laboratory interpretation following American College of Medical Genetics and Genomics (ACMG) guidelines, and reverse phenotyping based on OMIM data [9]. This approach aims to shorten the time to accurate diagnosis, optimize resource utilization, and ensure equitable access to modern diagnostic methods for rare pediatric neurological diseases. Ultimately, it is expected that combining deep and reverse phenotyping with ES will increase the proportion of genetically confirmed diagnoses and reduce the duration and number of diagnostic procedures.

## 2. Materials and Methods

### Study Design

This study is best classified as a retrospective observational study. Although patients were identified and deeply phenotyped according to a standardized prospective protocol, the key clinical and procedural variables—such as time to diagnosis, number of diagnostic interventions prior to ES, and turnaround time—were collected retrospectively from medical records. This design allowed for the evaluation of real-world diagnostic trajectories in children with suspected rare neurological diseases. Such a design allowed for a comprehensive assessment of the effectiveness of a diagnostic

algorithm based on deep phenotyping and ES, with a focus on reducing the time to diagnosis and minimising the diagnostic burden in children suspected of having rare neurological diseases. The choice of this design is justified by its effectiveness in studying rare diseases, as it enables the implementation of an integrated diagnostic approach with rational use of time and resources [10].

Patients were referred from 11 healthcare institutions in South Kazakhstan between July 1 and December 20, 2023. Key stages of the study—deep phenotyping and biological sample collection—were conducted at the Clinical and Diagnostic Center of the Khoja Akhmet Yassawi International Kazakh-Turkish University.

### **Preliminary Phase: Physician Training**

Before patient recruitment, individual and group training sessions were conducted for outpatient physicians. These sessions focused on recognizing rare neurological diseases with suspected genetic etiology. The training covered key phenotypic features of RPND and aimed to develop physicians' skills in the early identification of patients who may benefit from deep phenotyping and genetic testing. Upon completion, pediatric neurologists and general practitioners referred patients who met the established criteria, ensuring high-quality patient selection for further investigation.

### **Study Population**

Initially, 250 children under the age of 18 with suspected RPND were considered for inclusion. The final study sample was determined based on strict inclusion and exclusion criteria.

**Inclusion Criteria:** complex neurological phenotype involving a combination of syndromes (e.g., epilepsy, cognitive impairment, movement disorders, neurodegeneration, psychiatric symptoms, peripheral neuropathy), clinical suspicion of hereditary etiology, progressive or relapsing disease course with unknown etiology, lack of definitive diagnosis following clinical and instrumental investigations.

**Exclusion Criteria:** established non-genetic causes (infectious, toxic, vascular brain lesions), neuromuscular diseases confirmed by electromyography, muscle biopsy, or molecular genetic testing, specific nosologies (e.g., isolated epilepsy, confirmed chromosomal abnormalities), significant structural brain abnormalities detected by MRI or CT that do not require genetic confirmation.

After applying these criteria, the final study sample included 81 patients. This sample size is sufficient for the objectives of the study, given the epidemiological rarity of the conditions under investigation and the stringent inclusion/exclusion criteria.

### **Deep Phenotyping**

Each patient underwent a comprehensive clinical assessment including detailed history-taking, neurological examination, and review of instrumental findings [4]. The goal of this phase was to standardize and systematize clinical data for subsequent correlation with genetic findings. Deep phenotyping was performed using the HPO system, allowing for structured and unified documentation of clinical manifestations.

### **Stages of Deep Phenotyping:**

1. **Clinical Examination and History Taking:** Each patient underwent a detailed physical and neurological examination, assessing both central and peripheral nervous systems. History taking focused on identifying neurological and somatic symptoms, their onset, progression, and dynamics.
2. **HPO-Based Coding of Clinical Features:** Identified clinical manifestations were converted into standardized terms using the HPO system. Each symptom, its severity, and specificity were recorded using appropriate HPO codes. Special attention was paid to the correct term hierarchy: parent terms define general clinical categories, while child terms specify particular manifestations within those categories. For example, if a patient had seizures, the parent term "Seizure" (HP:0001250) was used. If generalized seizures were present, the child term "Generalized clonic seizure" (HP:0011169) provided further detail. Similarly, "Severe muscular hypotonia" was recorded as HP:0006829.

3. **Phenotypic Profile Formation:** Based on the coded HPO terms, an individual phenotypic profile was created for each patient. This profile served as an integrated clinical map, reflecting all detected symptoms and their characteristics. For instance, if a patient presented with generalized clonic seizures and severe muscular hypotonia, their profile would include: HP:0011169, and HP:0006829. This coding ensured a standardized and precise clinical description, which is critical for subsequent genetic interpretation.
4. **Phenotype Verification:** At this stage, the phenotypic profiles were analyzed to eliminate redundant or non-informative data. Each phenotypic feature was assessed for clinical relevance and consistency with known disease patterns. Results were discussed during interdisciplinary case reviews. All verified data were uploaded to the 3Billion laboratory portal to maintain accuracy and support the precise interpretation of genetic findings.

### **Molecular Genetic Analysis**

To identify the genetic causes of the disorders, ES was performed at the 3Billion laboratory in Seoul, South Korea. Genomic DNA was extracted from peripheral blood or dried blood spots using standard protocols. Exome capture was performed using the xGen Exome Research Panel v2 (Integrated DNA Technologies), supplemented with mitochondrial and custom panels. Sequencing was carried out on the NovaSeq X (Illumina, USA), achieving an average depth of 140× and ≥20× coverage for 99.6% of the targeted regions. Sequencing quality met clinical-grade standards (CAP, CLIA certified).

Raw data were processed using 3Billion's bioinformatics pipeline EVIDENCE v4.2, which incorporates GATK (v4.4.0) for SNV/INDEL calling, Manta for structural variant detection, 3bCNV for copy number variation analysis, and ExpansionHunter, MELT, and AutoMap for repeat expansions, mobile element insertions, and regions of homozygosity, respectively. Variant annotation was performed using Ensembl's Variant Effect Predictor (VEP v104.2) [11].

Variants were filtered using a multi-parameter approach. First, allele frequency thresholds were applied, excluding variants with a minor allele frequency (MAF) ≥1% in the gnomAD v4.1.0 database. Second, predicted pathogenicity was assessed, with prioritization given to protein-truncating variants such as frameshift, nonsense, and canonical splice-site changes, as well as to missense variants predicted to be deleterious by multiple in silico tools, including SIFT, PolyPhen-2, and Combined Annotation Dependent Depletion (CADD). Third, inheritance patterns were considered, taking into account family history and zygosity (e.g., homozygous, compound heterozygous, de novo). Finally, only high-confidence variants with a sequencing depth of at least 20× were retained for analysis; low-quality or ambiguous variants were subjected to orthogonal validation by Sanger sequencing.

All variants were classified according to ACMG guidelines (Richards et al., 2015), integrating evidence from population databases, computational tools, prior literature, and gene-disease relationships. Importantly, each clinically significant or uncertain variant was reviewed by a multidisciplinary team, including pediatric neurologists and certified clinical geneticists. This manual validation step ensured the clinical accuracy and contextual relevance of the automated classification output from EVIDENCE.

### **Reverse Phenotyping**

Reverse phenotyping—the process of correlating identified genetic variants with a patient's clinical manifestations to clarify their significance—was conducted using the OMIM database [12]. The clinical profile of each patient was compared with phenotypic descriptions associated with the identified genetic variants. The genetic findings were analyzed during interdisciplinary case discussions involving neurologists and neurogeneticists, during which their diagnostic relevance and interpretation were assessed, including for variants of uncertain significance (VUS).

### **Evaluation of the Reduction in Diagnostic Odyssey**

This study assessed the effectiveness of the diagnostic approach in terms of reducing the so-called “diagnostic odyssey”—the period from the appearance of the first symptoms to the establishment of a final diagnosis.

The analysis was based on two key parameters:

1. Duration of the Diagnostic Process - the time interval from the onset of the first symptoms to the patient's inclusion in the study, as well as from the time of inclusion to the final diagnosis, was examined. This allowed the evaluation of whether the new diagnostic algorithm accelerated the diagnostic process.
2. Diagnostic Burden - the total number of diagnostic tests (e.g., specialist consultations, MRI, biopsies, etc.) conducted before and after the patient's inclusion in the study was analyzed. This parameter enabled the assessment of whether the new diagnostic intervention reduced the number of additional investigations or, conversely, required more.

The obtained results will help determine the extent to which the proposed approach contributes to diagnostic optimization, reduction of time and resource expenditures, and improvement in the quality of medical care for patients.

### Statistical Analysis

Statistical analysis was performed using R software (version 4.3.1) and JASP (version 0.18.1). Descriptive statistics for quantitative variables were presented as median, interquartile range (IQR), range, and standard deviation, depending on the distribution of the data. Categorical variables were described using absolute numbers and percentages.

Normality of distribution was assessed using the Shapiro–Wilk test and visual inspection of Q–Q plots. Due to deviations from normal distribution, Wilcoxon signed-rank test was used to compare paired dependent samples.

To quantify the effect size, Cohen's d coefficient (adapted for paired samples) was calculated. The level of statistical significance was set at  $p < 0.05$  for two-tailed tests. All calculations were conducted with 95% confidence intervals.

## 3. Results

### Results of Deep Phenotyping

The study included 81 patients with neurological symptoms, with a predominance of males (64.2%, male-to-female ratio: 52:29). Age at the time of evaluation ranged from 6 months to 17 years (median: 6 years; interquartile range: 4–11.5 years).

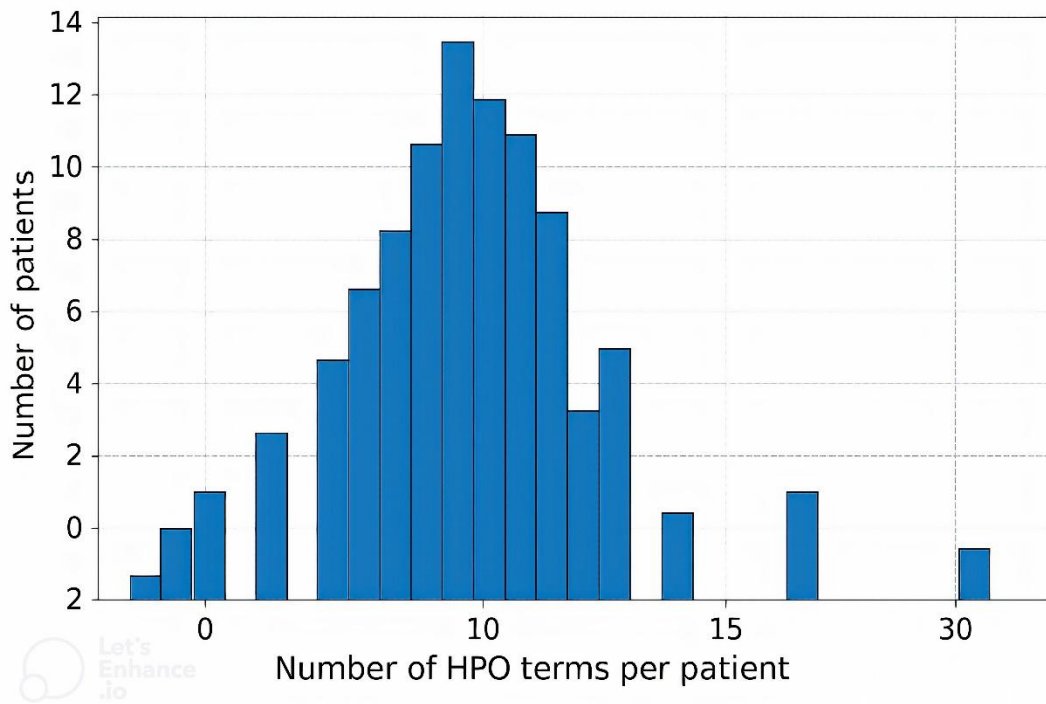
Almost all patients exhibited psychomotor developmental delay and intellectual disability: 78 out of 81 children (96.3%) had varying degrees of cognitive impairment. Severe or profound intellectual disability was observed in the majority of cases (approximately 78%), moderate in 13%, and mild in only 6%; only 2 patients demonstrated age-appropriate cognitive development.

Epileptic seizures were reported in 56 patients (67%). All participants had some degree of motor impairment: approximately 49% were unable to walk independently and had severe motor dysfunction corresponding to Gross Motor Function Classification System (GMFCS) levels IV–V, while 13% retained relatively preserved motor function (GMFCS level I).

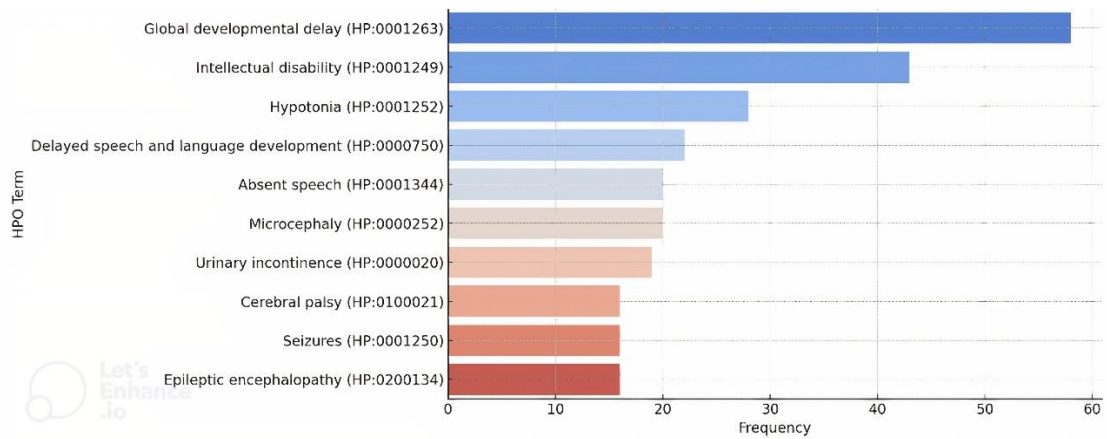
Each patient's clinical phenotype was standardized and structured using HPO terms. On average, 12 HPO terms were assigned per patient (median: 11; range: 2–30), reflecting the spectrum of clinical manifestations. The most common neurological features included global developmental delay/intellectual disability, epilepsy, infantile hypotonia, and spastic paresis, among others. Nearly all patients exhibited complex phenotypes with multiple (two or more) manifestations, and in over half of the cases, 10 to 15 HPO terms were used to describe the phenotype.

The development of such detailed phenotypic profiles allowed for a standardized clinical description, which is essential for the interpretation of subsequent genetic testing results.

Figure 1 presents the distribution of the number of HPO terms per patient. The most frequently observed HPO terms are presented in Figure 2. A detailed clinical characterization of the probands is presented in Table 1 and Table S1.



**Figure 1.** Number of HPO Terms Assigned per Patient in the Study Population. HPO - Human Phenotype Ontology.



**Figure 2.** Top 10 Most Frequent HPO Terms in the Study Population.

**Table 1.** Detailed Characteristics of the Study Cohort.

Case	Age (y)	Sex	GMFS	ID	Epilepsy	HPO Terms Count	Diagnostic Odyssey Before Research (months)	Total diagnostic procedures before participation	Total diagnostic procedures during the study	Diagnosis time (months) in the study	Clinical diagnosis refinement via reverse phenotyping	Genetic Diagnosis
1	9	Male	I	Severe	Yes	13	106	7	2	5	Yes	RAI1
2	7	Female	II	Moderate	Yes	9	84	20	2	5	Yes	KCNQ2
3	10	Male	I	Moderate	Yes	13	108	21	2	5	Yes	KDM6B
4	7 months	Female	V	Severe	Yes	10	7	7	2	5	NA	
5	9	Female	II	Profound	Yes	9	108	15	2	5	NA	
6	3	Female	V	Severe	Yes	11	30	14	2	5	Yes	SCN8A
7	5	Female	II	Severe	Yes	10	60	10	2	5	NA	
8	10	Female	II	Severe	Yes	23	114	25	2	5	Yes	CDKL5
9	14	Female	I	Moderate	Yes	9	159	18	2	5	Yes	UBAP2L
10	17	Male	I	Severe	Yes	10	140	20	2	5	Yes	deletion 3q29
11	4	Male	II	Severe	Yes	9	48	12	2	5	NA	
12	3	Male	II	Moderate	Yes	9	33	25	2	5	Yes	SCN1A
13	14	Male	V	Severe	Yes	10	168	14	2	5	NA	
14	12	Male	I	No	No	7	126	18	2	5	NA	
15	10	Female	V	Severe	No	10	120	26	2	5	NA	
16	5	Female	II	Severe	Yes	16	59	30	2	5	Yes	TSC2
17	4	Female	V	Profound	Yes	10	48	30	2	5	NA	
18	13	Male	II	Moderate	Yes	13	156	28	2	5	No	9p24.3p22.2
19	16	Male	V	Severe	Yes	10	192	19	2	5	NA	
20	10	Male	IV	Severe	Yes	5	120	25	2	5	Yes	FAR1
21	5	Male	I	легкая	Yes	10	40	26	2	5	Yes	TSC1
22	4	Male	V	Severe	No	4	48	25	2	5	NA	
23	17	Female	III	Mild	Yes	10	204	32	2	5	NA	

24	17	Male	V	Severe	Yes	10	204	27	2	5	NA	
25	4	Male	V	Severe	Yes	14	38	20	2	5	No	ANO3
26	3	Female	V	Severe	No	10	36	26	2	5	Yes	BRAT1
27	5	Male	II	Moderate	Yes	13	60	16	2	5	Yes	KLHL20
28	17	Female	I	No	Yes	7	72	28	2	5	Yes	TSC2
29	2	Female	V	Severe	Yes	11	24	12	2	5	NA	
30	14	Male	III	Severe	Yes	18	168	20	2	5	No	SCN1A
31	3	Male	V	Severe	Yes	11	36	20	2	5	NA	
32	4	Male	V	Severe	Yes	12	48	20	2	5	NA	
33	14	Male	II	Moderate	No	10	168	23	2	5	NA	
34	1	Male	V	Severe	No	14	12	15	2	5	NA	
35	4	Female	II	Severe	Yes	7	48	15	2	5	Yes	deletion 17q12.
36	5	Male	V	Severe	No	9	60	14	2	5	Yes	PRICKLE2
37	6	Female	V	Severe	No	11	72	32	2	5	Yes	SLC17A5
38	14	Female	II	Mild	No	9	168	39	2	5	Yes	MTHFS
39	6	Female	II	Severe	Yes	9	72	15	2	5	Yes	KCND2
40	3	Male	V	Severe	No	20	36	20	2	5	NA	
41	7	Female	V	Severe	Yes	9	84	40	2	5	Yes	FRA10AC1
42	3	Male	V	Severe	Yes	15	36	20	2	5	NA	
43	4	Male	V	Profound	Yes	8	48	19	2	5	NA	
44	13	Male	IV	Moderate	Yes	12	156	29	2	5	NA	
45	3	Male	V	Severe	No	9	36	20	2	5	NA	
46	12	Male	II	Moderate	No	15	144	28	2	5	Yes	LAMA1
47	6	Female	V	Severe	Yes	18	72	20	2	5	Yes	TRIO
48	4	Male	II	Severe	Yes	14	48	15	2	5	NA	
49	5	Male	V	Severe	Yes	30	60	20	2	5	Yes	CUL4B, GRIK2
50	4	Female	I	Severe	Yes	15	48	15	2	5	Yes	DYRK1A
51	7	Male	II	Severe	Yes	10	84	20	2	5	NA	

52	10	Male	II	Severe	No	20	120	38	2	5	Yes	VPS13B
53	3	Female	IV	Severe	Yes	19	36	20	2	5	Yes	ARID1B
54	3	Male	V	Severe	No	16	36	18	2	5	NA	
55	4	Male	V	Severe	Yes	5	48	20	2	5	NA	
56	16	Male	I	Mild	No	2	90	14	2	5	NA	
57	7	Female	II	Severe	Yes	12	84	25	2	5	Yes	duplication 9p24.3p21.2
58	17	Male	V	Severe	Yes	11	204	20	2	5	NA	
59	4	Male	III	Severe	Yes	12	48	18	2	5	NA	
60	13	Male	II	Severe	No	10	156	28	2	4	Yes	ITPR1
61	15	Female	II	Moderate	Yes	5	180	38	2	4	Yes	L2HGDH
62	4	Male	V	Severe	No	13	48	45	2	4	Yes	OCRL
63	1	Female	V	Severe	No	24	12	18	2	4	Yes	HRAS
64	3	Male	V	Severe	No	14	36	20	2	4	NA	
65	5	Male	II	Severe	No	18	60	30	2	4	Yes	IDS, MKKS
66	5	Male	V	Severe	Yes	8	60	20	2	4	Yes	MTTL1
67	5	Male	V	Severe	No	19	60	15	2	4	NA	
68	6	Male	IV	Severe	Yes	20	72	18	2	4	NA	
69	8	Male	II	Moderate	Yes	23	96	25	2	4	Yes	FGFR2
70	7	Female	II	Severe	Yes	17	84	14	2	4	Yes	KMT2D
71	10	Male	V	Severe	Yes	13	120	20	2	4	NA	
72	5	Male	II	Severe	No	18	60	28	2	3	Yes	deletion 5p15.33p14.3
73	3	Male	IV	Severe	No	11	36	15	2	3	NA	
74	14	Male	I	Mild	No	6	168	10	2	3	NA	
75	8	Female	II	Severe	Yes	23	96	25	2	3	NA	
76	6	Female	II	Severe	No	22	72	20	2	3	Yes	ARID1B
77	4	Male	II	Severe	No	15	48	18	2	3	NA	

78	6	Male	II	Severe	Yes	17	72	20	2	3	NA	
79	11	Female	IV	Severe	Yes	13	132	20	2	3	Yes	WDFY3
80	12	Male	I	Severe	Yes	7	72	18	2	3	Yes	GRIN2A
81	6 months	Male	V	Profound	Yes	18	6	12	2	3	Yes	WFOX

GMFCS - Gross Motor Function Classification System, HPO – Human Phenotype Ontology, ID – Intellectual disability, NA - Not applicable, y - years.

## ES Results

In the study group of 81 patients, genomic variations were identified in 43 individuals (54.2%), including single nucleotide variants (SNV) in 38 patients (88.9%) and copy number variations (CNV) in 5 patients (11.1%).

Among the SNVs, a number of pathogenic / likely pathogenic and potentially significant genetic variants were identified (Table 2, Table S2). It should be noted that in one case, two SNVs were found – IDS and MKKS. A total of 39 rare variants were detected, of which, according to ACMG criteria, 18 were classified as pathogenic, 12 as likely pathogenic, and 9 as VUS. Mutations were found in 39 different genes, with the most frequently affected being SCN1A, TSC2, and ARID1B (two different variants identified in each gene in different patients).

**Table 2.** Single nucleotide variants were detected in the studied cohort by ES analysis.

Case	Gene	Zygoty	Phenotype	ACMG criteria
1	RAI1	Heterozygous	Smith-Magenis syndrome	Pathogenic
2	KCNQ2	Heterozygous	Developmental and epileptic encephalopathy 7	Likely pathogenic
6	SCN8A	Heterozygous	Developmental and epileptic encephalopathy 13	VUS
8	CDKL5	Heterozygous	'Developmental and epileptic encephalopathy 2	Likely pathogenic
12	SCN1A	Heterozygous	Dravet syndrome	Likely pathogenic
3	KDM6B	Heterozygous	Neurodevelopmental disorder with coarse facies and mild distal skeletal abnormalities	Likely pathogenic
9	UBAP2L	Heterozygous	Neurodevelopmental disorder with impaired language, behavioral abnormalities, and dysmorphic facies	Likely pathogenic
16	TSC2	Heterozygous	Tuberous sclerosis-2	Pathogenic
20	FAR1	Heterozygous	Cataracts, spastic paraparesis, and speech delay	Pathogenic
21	TSC1	Heterozygous	Tuberous sclerosis-1	Pathogenic
27	KLHL20	Heterozygous	KLHL20-related disorder	Likely pathogenic
25	ANO3	Heterozygous	Dystonia 24	Pathogenic
26	BRAT1	Heterozygous	Neurodevelopmental disorder with cerebellar atrophy and with or without seizures	Likely pathogenic
28	TSC2	Heterozygous	Tuberous sclerosis-2	VUS
30	SCN1A	Heterozygous	SCN1A-related disorder	VUS
36	PRICKLE2	Heterozygous	PRICKLE2-related neurodevelopmental disorder	VUS
37	SLC17A5	Homozygous	Salla disease	Pathogenic

38	MTHFS	Heterozygous	Neurodevelopmental disorder with microcephaly, epilepsy, and hypomyelination	Likely pathogenic
39	KCND2	Heterozygous	KCND2-related neurodevelopmental disorder	Likely pathogenic
41	FRA10AC1	Homozygous	Neurodevelopmental disorder with growth retardation, dysmorphic facies, and corpus callosum abnormalities	Pathogenic
46	LAMA1	Heterozygous	Poretti-Boltshauser syndrome	Likely pathogenic
47	TRIO	Heterozygous	Intellectual developmental disorder, autosomal dominant 44, with microcephaly	VUS
49	CUL4B	Hemizygous	Intellectual developmental disorder, X-linked syndromic, Cabezas type	VUS
50	DYRK1A	Heterozygous	Intellectual developmental disorder, autosomal dominant 7	Pathogenic
53	ARID1B	Heterozygous	Coffin-Siris syndrome 1	Likely pathogenic
52	VPS13B	Heterozygous	Cohen syndrome	Pathogenic
60	ITPR1	Heterozygous	Gillespie syndrome	VUS
61	L2HGDH	Homozygous	L2-hydroxyglutaric aciduria	Pathogenic
62	OCRL	Hemizygous	Lowe syndrome	VUS
63	HRAS	Heterozygous	Costello syndrome	Pathogenic
65	MKKS	Homozygous	BARDET-BIEDL SYNDROME 6	Pathogenic
	IDS	Hemizygous	Mucopolysaccharidosis II	Pathogenic
66	MTTL1	Heteroplasmic	Mitochondrial myopathy, encephalopathy, lactic acidosis, and stroke-like episodes	Pathogenic
69	FGFR2	Heterozygous	Crouzon syndrome	Pathogenic
70	KMT2D	Heterozygous	Kabuki syndrome 1	Pathogenic
76	ARID1B	Heterozygous	Coffin-Siris syndrome 1	Pathogenic
79	WDFY3	Heterozygous	Microcephaly 18, primary, autosomal dominant	VUS
80	GRIN2A	Heterozygous	Epilepsy, focal, with speech disorder and with or without impaired intellectual development	Pathogenic
81	WWOX	Homozygous	Developmental and epileptic encephalopathy 28	Likely pathogenic

VUS - Variant of uncertain significance, ACMG - American College of Medical Genetics and Genomics.

A significant portion of the identified variants was associated with developmental and epileptic encephalopathies (DEE). Pathogenic or likely pathogenic mutations were identified in genes associated with DEE, including *KCNQ2* (DEE7), *SCN1A* (Dravet syndrome, DEE type 6), *SCN8A* (DEE13), *CDKL5* (DEE2), *WWOX* (DEE28), *DNM1* (DEE31A), and *GRIN2A* (focal epilepsy with speech disorder and intellectual disability). Altogether, variants in DEE-related genes accounted for approximately 20% of all identified mutations.

Another large group of findings (35.5%) involved genes associated with neurodevelopmental disorders (intellectual disability, developmental and behavioral impairments). These included, for example, *KDM6B* and *KCND2* (both associated with autosomal dominant forms of developmental delay), *UBAP2L* and *BRAT1* (autosomal recessive neurodevelopmental syndromes), as well as *TRIO*, *CUL4B*, *DYRK1A*, and others.

In addition, a number of identified variants were linked to hereditary syndromic disorders: in our cohort, mutations were found that are associated with tuberous sclerosis (*TSC1/TSC2*), Costello syndrome (*HRAS*), Kabuki syndrome (*KMT2D*), Smith–Magenis syndrome (*RAI1*), Lowe syndrome (*OCRL*), Bardet–Biedl syndrome (*MKKS*), and other monogenic disorders.

Of particular note are the variants related to inherited metabolic diseases: in particular, pathogenic mutations were identified that cause Salla disease (*SLC17A5*) and L2-hydroxyglutaric aciduria (*L2HGDH*), as well as a mitochondrial mutation associated with MELAS syndrome (*MTTL1*).

The identified CNVs included deletions and duplications of various chromosomal regions associated with specific phenotypic manifestations (Table 3). In Case 18, a duplication in 9p24.3p22.2 was detected in the patient, but it was not associated with any known syndromes. However, there have been reports of duplications partially or completely involving this genomic region, in which individuals exhibited phenotypic similarities to this patient [13–15].

**Table 3.** Identified CNVs in the Study Cohort.

Case	CNV	Phenotype
10	deletion 3q29	Chromosome 3q29 microdeletion syndrome
18	duplication 9p24.3p22.2	N/A
35	deletion 17q12.	Chromosome 17q12 deletion syndrome
57	duplication 9p24.3p21.2	Trisomy 9p
72	deletion 5p15.33p14.3	Cri-du-chat syndrome

CNV - Copy Number Variations.

### Reverse Phenotyping

In the group of 30 patients with identified pathogenic and likely pathogenic variants, reverse phenotyping allowed for refinement of the clinical diagnosis in 100% of cases. In these cases, the patients' clinical features were fully consistent with the identified genetic variants.

Reverse phenotyping also contributed to the identification of additional phenotypic features in 10 out of 30 patients (33.3%). During targeted re-evaluation, previously undocumented symptoms were discovered. Newly identified phenotypic manifestations included metabolic abnormalities, dysmorphic features, clinically insignificant cardiac arrhythmia, and others.

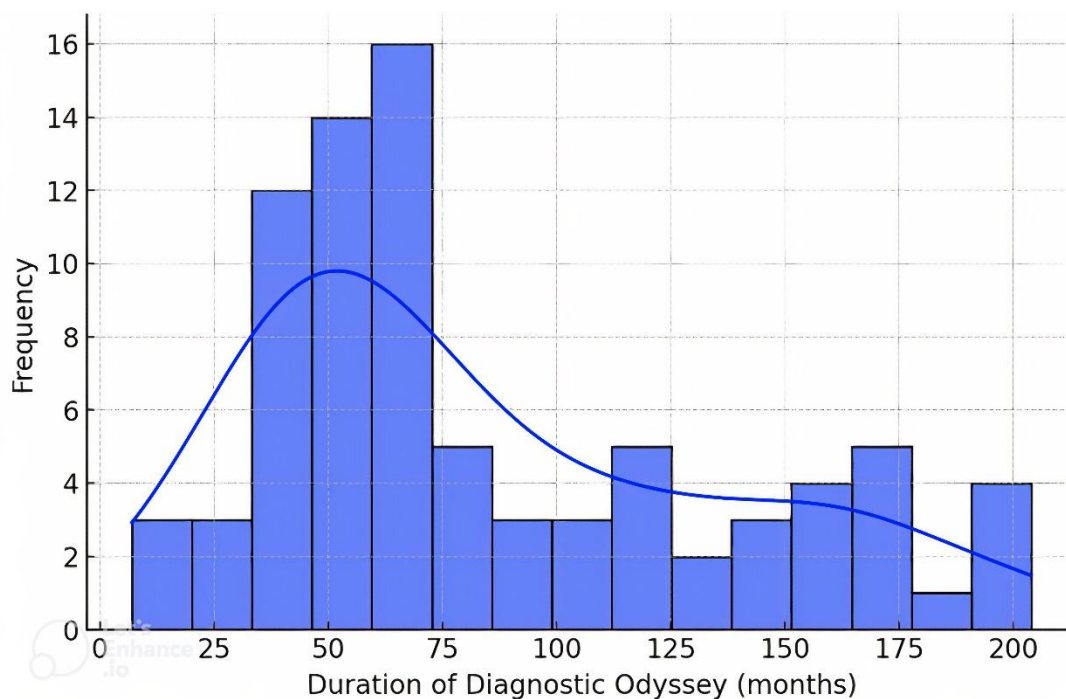
In the group of 9 patients with identified VUS, an in-depth analysis was performed using the reverse phenotyping approach combined with interdisciplinary discussion. In 88.9% of cases (8 out of 9 patients), the refined phenotypic profile was consistent with clinical features described in the literature for diseases associated with the respective gene. This enabled the formulation of a well-supported assumption regarding the likely pathogenic role of the identified variant in the patient's clinical presentation.

When formulating the presumptive genetic diagnosis, the following factors were considered: concordance between the phenotype and known disease manifestations; results of repeated

phenotypic assessment using HPO terms; and conclusions from the interdisciplinary case review. Although the formal classification of the variant remained as VUS, the obtained data strengthened its clinical relevance in the context of the individual case.

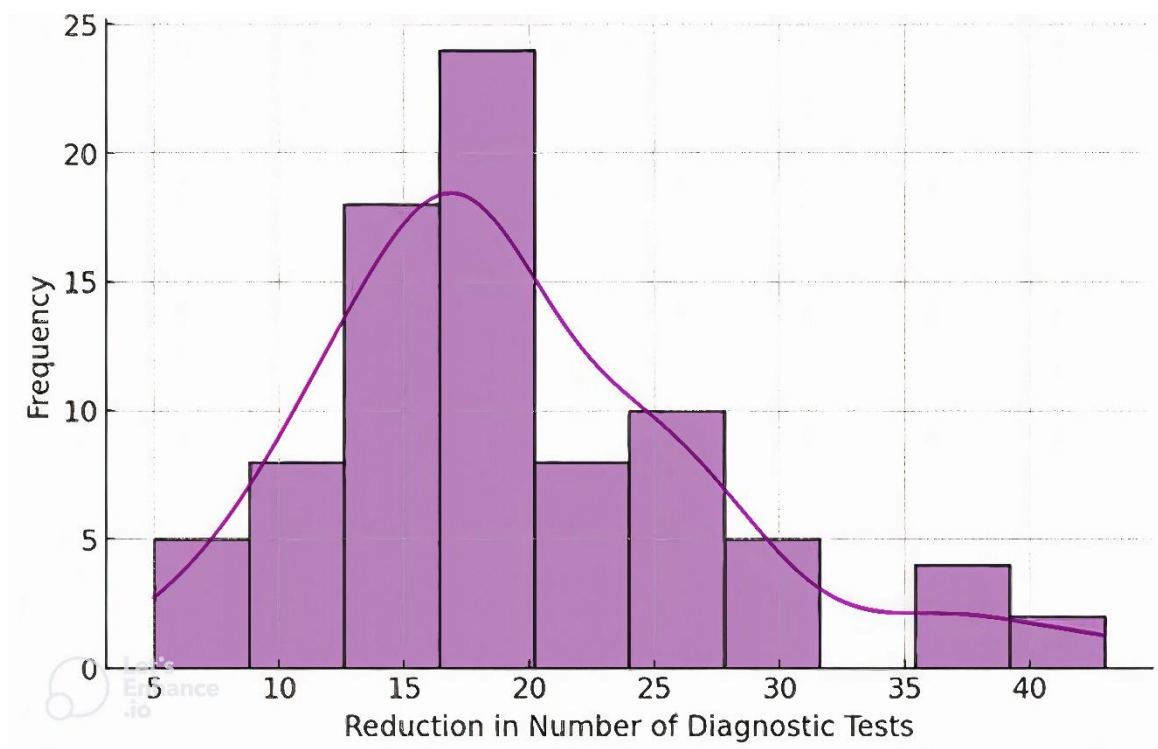
#### Evaluation of the Reduction in Diagnostic Odyssey

Prior to inclusion in the study, the median time interval from the onset of the first symptoms to participation in the study was 72 months (range: 6–204 months) (Figure 3). Following participation in the study and the initiation of molecular genetic testing, this period was reduced to up to 5 months.



**Figure 3.** Distribution of Diagnostic Odyssey Duration.

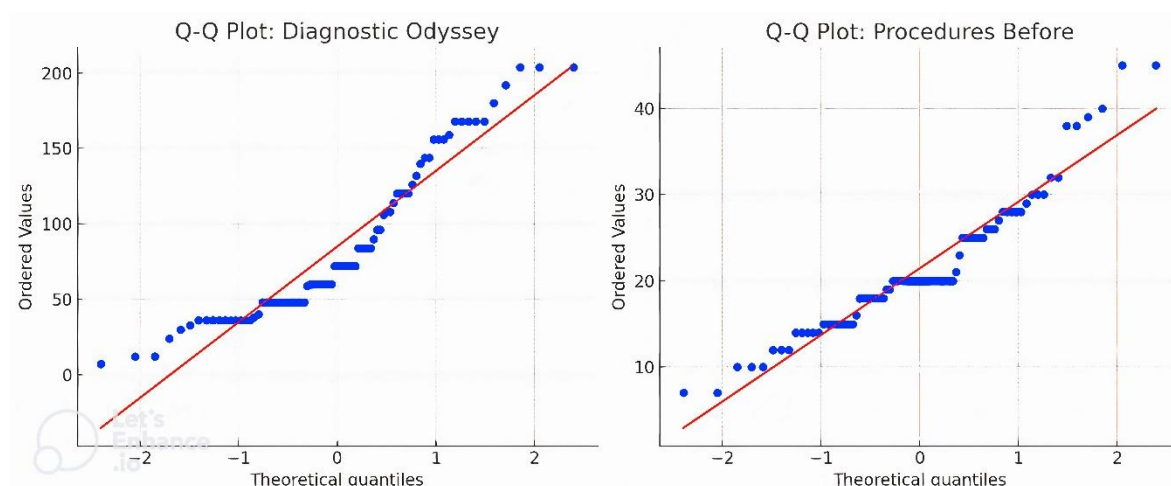
Before undergoing genetic testing, patients underwent an average of 20 different diagnostic procedures (Figure 4). After the implementation of genetic testing and targeted confirmatory investigations, the median number of additional tests was 2 (Table 1).



**Figure 4.** Distribution of Reduction in Diagnostic Tests.

A comparison of these indicators before and after study participation demonstrates a marked reduction in the diagnostic odyssey. As a result of applying the new diagnostic approach, the average time to diagnosis was reduced by approximately 19-fold compared to the initial period, and the number of required diagnostic procedures decreased more than tenfold.

The Q-Q plots demonstrate deviations from normality in both the duration of the diagnostic odyssey and the number of diagnostic procedures before genetic testing, indicating non-normal data distribution (Figure 5).



**Figure 5.** Q-Q Plots Assessing Normality of Diagnostic Odyssey Duration and Number of Procedures Before Study.

To assess the significance of differences in the duration of the diagnostic process before and after inclusion in the study, statistical data analysis was performed. The Shapiro–Wilk test revealed a deviation from normal distribution ( $p < 0.05$ ) in both groups, which justified the use of a non-

parametric comparison method. The Wilcoxon signed-rank test for paired samples was applied to assess statistical differences.

The results showed that the median duration of the diagnostic odyssey was significantly reduced after implementation of the proposed diagnostic algorithm ( $p = 1.91 \times 10^{-6}$ ). The main statistical indicators are presented in Table 4. Effect Size Estimation: To quantitatively assess the magnitude of change, effect size (Cohen's  $d$ ) was calculated. The resulting value of  $d = 2.43$  corresponds to a very strong effect in reducing diagnostic time.

**Table 4.** Comparative analysis of diagnostic process duration before and after the study.

Indicator	Before Study Participation	After Genetic Investigation
Mean diagnostic duration (months)	102.6	4.2
Standard deviation (months)	57.2	0.77
Median diagnostic duration (months)	72	5
Range (months)	6 – 204	3 – 5
Number of diagnostic procedures (median)	20 [7–45]	2 [2–2]
Number of confirmed genetic diagnoses	0 (pre-genetic stage)	43 out of 81 (53%)
Proportion of variants of uncertain significance	–	3 out of 43 (7% of identified variants)

#### 4. Discussion

The results of our study demonstrate the high effectiveness of integrating deep phenotyping and genomic sequencing in the diagnosis of rare RPND. The diagnostic yield in our cohort reached 53% (43 of 81 patients received a genetically confirmed diagnosis), which is notably higher than that reported in studies employing genomic testing in the absence of standardized deep phenotyping approaches, such as Human Phenotype Ontology-based annotation. Our findings are comparable to leading international outcomes—for instance, a recent genome GS study reported a diagnostic yield of 61%, underscoring the value of comprehensive clinical assessment and reverse phenotyping [16]. Notably, to our knowledge, no similar large-scale studies integrating structured phenotyping and genomic diagnostics have been conducted in Central Asia, highlighting the novelty and regional significance of our work.

All patients in our cohort underwent systematic deep phenotyping, including thorough medical history, neurological examination, and standardized documentation using HPO. This structured approach enabled precise alignment between clinical manifestations and ES results. Reverse phenotyping further uncovered previously unrecognized features in a substantial subset of patients, thereby strengthening genotype–phenotype correlations. Such methodologies are increasingly endorsed in the current literature as effective strategies for interpreting VUS [16].

From a healthcare delivery perspective, this integrated diagnostic model resulted in a marked reduction in the diagnostic odyssey. Prior to inclusion, the median duration of the diagnostic process was 72 months, with a median of 20 diagnostic procedures per patient. Following the implementation of our protocol, diagnoses were achieved within 4–5 months and required only two confirmatory investigations. These results align with global initiatives aimed at streamlining diagnostics for rare diseases. While genomic testing incurs initial costs, recent analyses show that early deployment of such technologies can substantially reduce cumulative diagnostic expenses [17]. Our data reinforce this notion, demonstrating that a single, well-targeted sequencing test can replace years of fragmented, inconclusive diagnostics.

Beyond economic considerations, timely diagnosis confers significant intangible benefits—alleviating familial anxiety and enabling earlier clinical management. Thus, the integrated approach presented here not only enhances diagnostic efficacy but also supports the broader goal of sustainable healthcare by promoting more rational resource allocation and improving the lived experiences of affected families.

Comparison with other cohorts reveals that our diagnostic performance is consistent with or surpasses global benchmarks. The proportion of genetically confirmed diagnoses in neurodevelopmental disorders is reported to range between 30% and 50%, depending on clinical severity and sequencing methodology [18]. Our study focused on patients with severe phenotypes—combinations of epilepsy, intellectual disability, and complex syndromic features—many of which are known to be monogenic, likely contributing to the high diagnostic yield.

Our findings illustrate a broad genetic spectrum, including well-established pathogenic variants (e.g., *RAI1* in Smith–Magenis syndrome, *KCNQ2* in early infantile epileptic encephalopathy) and less characterized gene-disease associations (e.g., a novel *KDM6B* variant likely linked to developmental delay). Notably, several molecular diagnoses diverged from the initial clinical hypotheses, highlighting the diagnostic value of ES in cases with atypical or overlapping phenotypes.

Despite methodological similarities to prior studies, our work is unique in its demonstration of a structured integration of deep and reverse phenotyping into routine clinical pathways in a resource-limited region, offering a reproducible model for other low-resource healthcare systems.

#### Limitations and Future Directions

Despite these strengths, our study has several limitations. Approximately 47% of cases remained without a molecular diagnosis. This may be partly attributed to the inherent limitations of ES, which does not provide full coverage of the genome and may miss pathogenic variants in deep intronic regions or structural rearrangements. However, other pathogenic mechanisms—such as repeat expansions, epigenetic modifications, and somatic mosaicism—are also likely contributors and are poorly detected even by genome GS. Emerging technologies such as methylation profiling, long-read sequencing, and single-cell approaches may help close this diagnostic gap in the future. While GS is increasingly adopted in high-income countries as a first-line diagnostic tool, neither GS nor ES are routinely available in many low- and middle-income regions due to the absence of local infrastructure and accredited sequencing facilities. In such contexts, genomic analyses must be outsourced to international laboratories. Among these options, ES remains relatively more accessible owing to its lower cost, making it the more practical choice for genomic diagnostics in resource-limited settings. Second, although VUS re-evaluation was supported by reverse phenotyping and multidisciplinary discussion, confirmatory functional studies were not performed, and formal reclassification under ACMG guidelines was not feasible within the study period. Third, segregation analysis—which could significantly aid in variant interpretation—was not systematically conducted and should be incorporated into future diagnostic workflows.

In light of these limitations, we plan to expand our diagnostic platform to include GS, chromosomal microarray analysis for copy number variants, epigenetic testing, long-read sequencing, and functional validation studies. These efforts aim to increase diagnostic yield among currently undiagnosed patients and to further refine genotype-phenotype relationships.

## 5. Conclusions

To our knowledge, this is one of the first multi-center studies in Central Asia to demonstrate the successful integration of deep phenotyping and advanced genomic diagnostics in pediatric neurology. Unlike isolated case reports, we present a reproducible and scalable model built on interdisciplinary collaboration, standardized clinical data collection, phenotype-driven variant prioritization, and reverse phenotyping. This diagnostic framework can serve as a prototype for establishing rare disease centers in other resource-limited settings.

In settings where families affected by rare disorders face significant emotional and financial strain, shortening the diagnostic journey from years to months represents a major advancement in care. Our findings not only validate a practical diagnostic approach but also emphasize the need for continued genomic innovation in unresolved cases. Early, accurate diagnosis enables targeted management, improves outcomes, and aligns with the principles of sustainable healthcare.

In conclusion, the integration of deep phenotyping, first-line ES, and reverse phenotyping in children with rare neurological disorders yielded a high diagnostic rate (53%), well above

conventional methods. This approach drastically reduced diagnostic delays and minimized unnecessary investigations, improving both patient quality of life and healthcare system efficiency. Our results support the incorporation of this model into routine practice and advocate for the broader adoption of genomics-informed, resource-conscious diagnostic strategies in pediatric neurology

## 6. Patents

This work resulted in the development of a diagnostic algorithm for complex neurological phenotypes using deep phenotyping methods. The algorithm is registered as a scientific intellectual property object with the Copyright Certificate No. 56915, issued on April 17, 2025, by the Ministry of Justice of the Republic of Kazakhstan.

**Supplementary Materials:** The following supporting information can be downloaded at the website of this paper posted on Preprints.org.

**Author Contributions:** **Author Contributions.** Conceptualization, N.Y. and R.K.; methodology, N.Y., R.K. and N.Zh.; software, R.K.; validation, N.Zh. and R.K.; formal analysis, N.Y. and A.O.; investigation, N.Y. and R.K.; resources, N.Zh. and G.N.; data curation, N.Y. and A.O.; writing—original draft preparation, N.Y. and S.A.; writing—review and editing, N.Y., S.A. and R.K.; visualization, N.Y. and G.N.; supervision, N.Zh. and R.K.; project administration, A.O.; funding acquisition, R.K., G.N. and A.O. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by 3Billion Inc. (South Korea), the Central Asian and Transcaucasian Rare Pediatric Neurological Diseases Genomic Consortium. (<https://www.cat-genomics.com/>), and the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. BR24992814).

**Institutional Review Board Statement:** The study protocol was reviewed and approved by the Ethics Committee of Khoja Akhmet Yassawi International Kazakh-Turkish University (Protocol №16, dated June 8, 2023). All procedures were conducted in accordance with the Declaration of Helsinki (2013 revision), the Convention on the Rights of the Child, and national regulations governing research involving minors.

**Informed Consent Statement:** Prior to inclusion in the study, all participants and their legal guardians were thoroughly informed—both verbally and in writing—about the objectives, procedures, potential risks and benefits of the research, as well as the possibility of publication of anonymized clinical and genetic data. Written informed consent for participation and publication was obtained from all legal guardians. When appropriate, verbal and written assent was also obtained from the minor participants in accordance with their age and level of understanding.

**Data Availability Statement:** De-identified clinical and genomic datasets generated during this study are available from the corresponding author upon reasonable request, subject to ethical and institutional review board approval.

**Acknowledgments:** The authors sincerely thank all the families who participated in this study for their trust, time, and cooperation. We gratefully acknowledge the contributions of pediatricians, pediatric neurologists, and clinical specialists involved in patient evaluation, recruitment, and phenotypic data collection.

We thank 3Billion Inc. (South Korea) for providing ES services and variant interpretation. We also acknowledge the collaborative and scientific support of the Central Asian and Transcaucasian Rare Pediatric Neurological Diseases Genomic Consortium. (CAT-RPND), coordinated by the UCL Queen Square Institute of Neurology (London, UK). We are grateful to the South Kazakhstan Medical Academy (Shymkent, Kazakhstan) for administrative and logistical support, including documentation and international transport of biological samples. The authors thank the Khoja Akhmet Yassawi International Kazakh-Turkish University for its institutional support throughout the study. Finally, we acknowledge the Clinical and Diagnostic Center of Khoja Akhmet Yassawi International Kazakh-Turkish University (Turkestan, Kazakhstan) for providing facilities for clinical assessments, patient visits, and biological sample collection.

**Conflicts of Interest:** The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## Abbreviations

The following abbreviations are used in this manuscript:

ACMG	American College of Medical Genetics and Genomics
CAT-RPND	Central Asia and Transcaucasia Rare Pediatric Neurological Diseases Genomic Consortium
GMFCS	Gross Motor Function Classification System
CNV	Copy Number Variations
DEE	Developmental and epileptic encephalopathy
ES	Exome Sequencing
GS	Genome Sequencing
HPO	Human Phenotype Ontology
NGS	Next-Generation Sequencing
NORD	National Organisation for Rare Disorders
RPND	Rare Paediatric Neurological Diseases
SDG	Sustainable Development Goals
SNV	Single Nucleotide Variants
VUS	Variant of uncertain significance

## References

1. Ferreira CR. The burden of rare diseases. *Am J Med Genet A*. 2019 Jun;179(6):885-892. doi: 10.1002/ajmg.a.61124.
2. Stark Z, Tan TY, Chong B *et al*. A prospective evaluation of whole-exome sequencing as a first-tier molecular test in infants with suspected monogenic disorders. *Genet Med*. 2016 Nov;18(11):1090-1096. doi: 10.1038/gim.2016.1.
3. Makarova EV, Krysanov IS, Valilyeva TP, Vasiliev MD, Zinchenko RA. Evaluation of orphan diseases global burden. *Eur J Transl Myol*. 2021 May 14;31(2):9610. doi: 10.4081/ejtm.2021.9610.
4. Köhler S, Schulz MH, Krawitz P *et al*. Clinical diagnostics in human genetics with semantic similarity searches in ontologies. *Am J Hum Genet*. 2009 Oct;85(4):457-64. doi: 10.1016/j.ajhg.2009.09.003.
5. Robinson PN. Deep phenotyping for precision medicine. *Hum Mutat*. 2012 May;33(5):777-80. doi: 10.1002/humu.22080.
6. Yang Y, Muzny DM, Reid JG *et al*. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N Engl J Med*. 2013 Oct 17;369(16):1502-11. doi: 10.1056/NEJMoa1306555. Epub 2013 Oct 2.
7. Kaiyrzhanov R, Zharkinbekova N, Guliyeva U *et al*. Elucidating the genomic basis of rare pediatric neurological diseases in Central Asia and Transcaucasia. *Nat Genet*. 2024 Dec;56(12):2582-2584. doi: 10.1038/s41588-024-02016-x. PMID: 39578646.
8. Strong K, Noor A, Aponte J *et al*. Monitoring the status of selected health related sustainable development goals: methods and projections to 2030. *Glob Health Action*. 2020 Dec 31;13(1):1846903. doi: 10.1080/16549716.2020.1846903.
9. Richards S, Aziz N, Bale S *et al*.; ACMG Laboratory Quality Assurance Committee. 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 May;17(5):405-24. doi: 10.1038/gim.2015.30.
10. Kernohan KD, Boycott KM. The expanding diagnostic toolbox for rare genetic diseases. *Nat Rev Genet*. 2024 Jun;25(6):401-415. doi: 10.1038/s41576-023-00683-w.

11. Seo GH, Kim T, Choi IH *et al.* Diagnostic yield and clinical utility of whole exome sequencing using an automated variant prioritization system, EVIDENCE. *Clin Genet.* 2020 Dec;98(6):562-570. doi: 10.1111/cge.13848.
12. Wilczewski CM, Obasohan J, Paschall JE *et al.* Genotype first: Clinical genomics research through a reverse phenotyping approach. *Am J Hum Genet.* 2023 Jan 5;110(1):3-12. doi: 10.1016/j.ajhg.2022.12.004.
13. Capkova Z, Capkova P, Srovnal J, Adamova K, Prochazka M, Hajduch M. Duplication of 9p24.3 in three unrelated patients and their phenotypes, considering affected genes, and similar recurrent variants. *Mol Genet Genomic Med.* 2021 Mar;9(3):e1592. doi: 10.1002/mgg3.1592.
14. Glessner JT, Li J, Wang D *et al.* Copy number variation meta-analysis reveals a novel duplication at 9p24 associated with multiple neurodevelopmental disorders. *Genome Med.* 2017 Nov 30;9(1):106. doi: 10.1186/s13073-017-0494-1.
15. Guilherme RS, Meloni VA, Perez *et al.* AB Duplication 9p and their implication to phenotype. *BMC Med Genet.* 2014 Dec 20;15:142. doi:10.1186/s12881-014-0142-1.
16. Akgun-Dogan O, Tuc Bengur E, Ay B *et al.* Impact of deep phenotyping: high diagnostic yield in a diverse pediatric population of 172 patients through clinical whole-genome sequencing at a single center. *Front Genet.* 2024 Mar 15;15:1347474. doi: 10.3389/fgene.2024.1347474.
17. Runheim H, Pettersson M, Hammarsjö A *et al.* The cost-effectiveness of whole genome sequencing in neurodevelopmental disorders. *Sci Rep.* 2023 Apr 27;13(1):6904. doi: 10.1038/s41598-023-33787-8.
18. Wang Q, Tang X, Yang K, Huo X, Zhang H, Ding K, Liao S. Deep phenotyping and whole-exome sequencing improved the diagnostic yield for nuclear pedigrees with neurodevelopmental disorders. *Mol Genet Genomic Med.* 2022 May;10(5):e1918. doi: 10.1002/mgg3.1918.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.